THE RELATIONSHIP BETWEEN INDIVIDUAL DIFFERENCES IN WORKING MEMORY AND FILTERING TASK-IRRELEVANT INFORMATION, IN CHILDREN AND ADULTS

by

MARIA KHARITONOVA

B. S., DEPAUL UNIVERSITY, 2003

M.A., UNIVERSITY OF COLORADO AT BOULDER, 2006

A thesis submitted to the

Faculty of the Graduate School of the

University of Colorado in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

2010

This thesis entitled:

The relationship between individual differences in working memory and filtering task-irrelevant information, in children and adults

written by Maria Kharitonova has been approved for the Department of Psychology and Neuroscience

Yuko Munakata (chair)

Tim Curran

Date: September 8th, 2010

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above-mentioned discipline.

IRB protocol #0806.25

ABSTRACT

Kharitonova, Maria. (Ph.D., Psychology, Neuroscience, Department of Psychology & Neuroscience, University of Colorado at Boulder)

The relationship between individual differences in working memory and filtering task-irrelevant information, in children and adults

Dissertation directed by Professor Yuko Munakata

We are constantly bombarded with myriad pieces of information, of which only a portion is directly relevant to our immediate experiences. What determines how much irrelevant information we filter, and how quickly we can adjust the amount of filtering? This dissertation explores the possibility that working memory (WM), an ability to actively maintain task-relevant information, plays a critical role in dynamically adjusting filtering strategy, based on task demands. High filtering could result from upregulating top-down control, leading to processing only the task-relevant information. Low filtering could result from loosening this control, to allow for a larger amount of information to be processed. This account was tested in three different experiments, with both adults and six-year-old children. Experiment 1 showed that high WM could support both high and low filtering, within the same adult participants. High WM was associated with high filtering in a paradigm where filtering task-irrelevant information was advantageous, because filtering the distractors reduced WM demand. In contrast, high WM was associated with low filtering in a task-switching paradigm, where low filtering was advantageous, because currently irrelevant features became relevant on subsequent switch trials. Experiment 2 modified the high-filtering-demand task of Experiment 1 to have both high- and low-filtering-demand versions, in order to manipulate filtering demand within the same paradigm. Results were difficult to interpret definitively due to participants' poor compliance with the stated instructions; nonetheless, they point to sensitivity of early attention to subtle changes in experimental setup. Experiment 3 tested filtering across two additional tasks in both adults and children, in order to assess the role of working memory in filtering in more robust paradigms, to test for dynamic changes in filtering strategy within the context of the same task, and to test the developmental origins of the relationship between WM and filtering. Results from both children and adults were mostly consistent with the dynamic filtering theory, but several important caveats are discussed. Despite these shortcomings, results from these experiments provide an important advance in understanding the role of WM in dynamically adjusting filtering strategies based on task demands.

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my advisor, Dr. Yuko Munakata for her tremendous support, continuous guidance, and intellectual challenge throughout graduate school. Her encouragement has been instrumental to my progress, and her sharp thinking and clear writing has been a constant source of inspiration. I owe many thanks to Dr. Tim Curran, who has been more available for support, feedback, and discussions than I could have reasonably expected him to. I am grateful to Dr. Eliana Colunga for her continuous encouragement and for many insightful discussions throughout these years. I thank Dr. Akira Miyake and Dr. Mike Mozer for helpful input during my dissertation proposal. I thank all of the aforementioned faculty for serving as members of my thesis committee.

I am grateful to the following undergraduate Research Assistants for helping with data collection: Chelsea Cerny, Emily Little, Danielle Robinson, Amy Turner, and all other assistants at the Cognitive Development Center and the Curran lab. I would also like to thank Hiromi Sumiya, the lab coordinator at the CDC, and Chris Bird, Brent Young, and Megan Freeman, the Professional Research Assistants in the Curran lab for coordinating and managing the pragmatic aspects of running these studies. I also extend many thanks to the participants of these studies, without whom this research would not be possible. Finally, the graduate students of the Cognitive Development Center and the Psychology department have contributed to both my sanity and to my intellectual growth. Katye Blackwell, Chris Chatham, and Hannah Snyder are thanked for all the ideas, feedback and fascinating discussions over the years. Chris' enthusiasm for all things science served as continual inspiration. Special thanks also goes to Hannah for hosting “grad dinners”, with fantastic food and conversation, and perhaps too much wine. Finally, I am grateful to James Kole for last-minute statistics advising.

Finally, none of this work would be possible without family and friends. I thank my parents and grandparents for inspiring intellectual curiosity way back when, and for unconditionally supporting all manners of intellectual pursuits, including this one. I owe more than I can express to Erika Nyhus, who has been not only a great colleague, but also a fantastic friend through the thick and thin of graduate school. Most of all, I would like to thank my husband, Dmitriy Dligach, for introducing me to the magic of Python and MATLAB, but more importantly, for being an infinite source of patience and encouragement throughout all these years.

The research presented here was supported by NICHD (R01 HD37163 and P50-MH079485).

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION: THE RELATIONSHIP BETWEEN WORKING MEMORY AND FILTERING TASK-IRRELEVANT INFORMATION

We are constantly bombarded with myriad pieces of information, of which only a portion is directly relevant to our immediate experiences. How do we choose what to focus our attention on, in light of the behaviors we are involved in currently and the behaviors we are envisioning in our immediate future? Sometimes focusing narrowly on the most relevant pieces of information is most appropriate, as when you are driving and need to keep your eyes (and attention) focused exclusively on the road ahead. In other situations, it might be more appropriate to focus broadly, by scanning the environment to determine the best outcome. For example, while teaching, it is important to be continuously scanning the classroom to check if students look confused and perhaps even raise their hands to ask a question, instead of only narrowly focusing on the slides.

What determines how broadly or narrowly we focus our attention and how quickly we can shift from attending narrowly to attending more broadly, and vice versa? Research has shown that chronic media multitasking in daily life (i.e. simultaneous consumption of multiple media sources) is associated with enhanced attention and memory for task-*irrelevant* distractors, suggesting that a broad focus in media consumption is mirrored by a broad focus in cognitive control measures (Ophir, Nass & Wagner, 2009). However, this research did not examine the possibility that the scope of attention can be malleable and shift to accommodate changing task demands. This dissertation will propose that working memory (WM), or the ability to actively maintain task-relevant information across delays and interferences (Miller & Cohen, 2001;

O'Reilly & Frank, 2006) and thus bias downstream processing in a controlled, top-down way (Kane & Engle, 2003), plays a critical role in determining the scope of attention. This dissertation will suggest that strong WM can dynamically modulate attentional scope to alternate between focusing only on the immediately relevant pieces of information (high filtering) and focusing more broadly and processing information that may not be immediately relevant (low filtering).

There is an apparent discrepancy in the existing literature regarding the way in which WM affects filtering. A number of studies report a positive link between WM and filtering, such that strong WM supports high filtering of task-irrelevant information (e.g. Fukuda & Vogel, 2009; Kane, Brown, McVay, Silvia, Myin-Germeys, & Kwapil, 2007; Vogel, McCollough & Machizawa, 2005). Several other studies, however, report a *negative* relationship between WM and filtering, such that strong WM is associated with the ability to maintain a broad attentional focus and take in a large amount of available information (e.g. Just & Carpenter, 1992; Waring, Payne, Schacter, & Kensinger, 2009). One way to reconcile these seemingly conflicting findings is by positing that high WM capacity can enable the most *efficient* allocation of top-down control to support either high or low filtering of task-irrelevant information, based on the immediate task demands. The hypothesis that high WM can support dynamic adjustment of filtering strategy based on current task demands is examined in detail in the three experiments that comprise this dissertation. Before discussing the experiments conducted to test this theory, I will first review the evidence suggesting both positive and negative relationships between WM and filtering, followed by a review of literature that explains *how* WM can support dynamic adjustment of filtering strategy, based on task demands. The need to examine the relationship between WM and filtering in development will also be discussed. The introduction will end with an overview of the experiments and the summary of the data, which are described in detail in subsequent chapters.

Positive relationship between WM and filtering

Most studies investigating individual differences in working memory and filtering have reported a *positive* relationship between WM and filtering task-irrelevant information. In terms of group-level effects, a number of studies show that overloading WM impairs the ability to successfully filter task-irrelevant information (e.g. Lavie, Hirst, de Fockert & Viding, 2004) and conversely, the process of attending to irrelevant distractors also impairs WM (e.g. Zanto & Gazzaley, 2009).

In terms of individual differences, high WM is often associated with more focused attention and greater ability to filter task-irrelevant information, with the converse pattern for low WM individuals. For example, participants with high WM treat visual displays with targets and distractors as if there were no distractors, whereas low WM participants process the distracting information (Fukuda & Vogel, 2009; Vogel et al., 2005). Participants with low WM are also slowed down more by distracting incongruent color words, when needing to report only the ink color of a word in the Stroop task than participants with high WM (Long & Prat, 2002). Prefrontal patients who are often impaired at WM tasks (e.g. Kane & Engle, 2002) are also impaired at ignoring task-irrelevant, distracting information (Chao & Knight, 1995; Kane & Engle, 2002). More generally, high WM has been associated with less mind wandering during challenging daily tasks (Kane et al., 2007), and with generally being less prone to distraction (Kane & Engle, 2003).

These findings suggest that high WM is related to the ability to selectively attend to the relevant, and filter the task-irrelevant information. However, in all of these situations, the high filtering strategy was task-advantageous, as it allowed participants to complete the task in the most optimal way, for example, by lowering the number of items that ultimately need to be maintained in WM (as in Vogel et al., 2005).

Several researchers discuss the positive link between WM and the ability to focus exclusively on the task-relevant information in terms of inhibitory processes, e.g., by suggesting that WM-related increase in filtering is achieved by actively inhibiting task-irrelevant information (e.g. Zacks & Hasher, 1994). However, as described below, filtering task-irrelevant information does not need to rely on inhibitory processes, and instead is argued to stem directly from the PFC-supported WM system that provides top-down control for processing the task-relevant information. The experiments in this dissertation were not designed to reconcile the inhibitory and the non-inhibitory explanations for filtering task-irrelevant information. Rather, the goal was to investigate whether high WM can dynamically support both high and low filtering strategies, based on task demands.

Negative relationship between WM and filtering

Although most research to date has focused on examining the relationship between WM and the narrow attentional focus (i.e. high filtering), a few studies have reported that WM can also be associated with a broad attentional focus. Specifically, despite everyone showing some deficit in memory for the neutral element (e.g. a forest) than the emotional element (e.g. a snake) of an overall emotional image (e.g. the snake in the forest), high WM participants showed a smaller deficit, and were more likely to remember the non-emotional element than the low WM participants (Waring et al., 2009). This suggests that high WM can be associated with a broader attentional focus in certain situations. Similarly, when viewing temporarily syntactically ambiguous sentences, such as "The experienced soldiers warned about the dangers before the midnight raid" ("warned" could be interpreted as either the main verb, such that soldiers did the warning, or as a past participle, such that "warned" qualifies soldiers), high WM participants were more likely than low WM participants to maintain both interpretations of the ambiguous word ("warned") long enough for the final word of the sentence to disambiguate the meaning (Just & Carpenter, 1992). Only high WM participants showed an effect of ambiguity, which was measured as the difference in RTs between reading an ambiguous sentence and an unambiguous sentence (e.g. "The experienced soldiers *spoke* about the dangers before the midnight raid"). This ambiguity effect was only apparent during the last few words of the sentence, which disambiguate its meaning. This finding suggests that high WM participants were strongly aware of the ambiguity, and maintained both meanings of the critical word, instead of narrowly focusing on the more dominant interpretation.

In both of these situations, the low filtering strategy was task-advantageous, as it allowed participants to ultimately choose the right answer in the task. Thus, in situations where low filtering is appropriate, high working memory can support *less* filtering, resulting in maintenance of a larger amount of the available information.

How can WM support dynamic adjustment of filtering?

The research described above suggests that in situations in which filtering of task-irrelevant information is advantageous for task performance, high WM is associated with greater filtering. However, when filtering is *dis*advantageous, high working memory can be associated with less filtering. These findings are not necessarily contradictory, despite seeming so on the surface. The relationship between working memory and filtering task-irrelevant information might be non-linear: high WM might support dynamic updating of the filtering strategy, to accommodate the current task demands in the most efficient manner.

How can WM support both high and low filtering, and how can WM support adjustment of filtering strategy based on task demands? Conceptualizing WM not in terms of static storage (e.g. Baddeley, 2003; Daneman & Carpenter, 1980), but rather in terms of an ability to exert top-down biasing for maintaining *task-relevant* information across delays and interferences (Blackwell, Cepeda, & Munakata, 2009; Frank & O'Reilly, 2006; Kane & Engle, 2002; Miller & Cohen, 2001) suggests that WM and filtering should be positively related. Stronger WM should lead to stronger top-down biasing to process only the relevant information, thus achieving a high level of filtering. Consistent with this idea, filtering irrelevant (and frequently incongruent) faces in the backgrounds, while judging whether the written names belong to pop stars or politicians, is impaired when this task is performed under high relative to low WM load (deFockert et al., 2001). Similarly, ignoring incongruent flankers on the periphery is more difficult under high than low WM load (Lavie et al., 2004), and irrelevant colors capture attention more in a visual search paradigm under high WM load (Lavie & deFockert, 2005). These results demonstrate that under high demand for filtering, WM is strongly related to the ability to filtering task-irrelevant information.

On the other hand, high WM should allow for maintenance of more items than low WM, if maintaining more items is deemed task appropriate. Thus, participants with high WM might be capable of maintaining both the immediately and the potentially relevant information, if the low filtering strategy appears to be most appropriate. For example, stronger WM can support maintenance of both the more and the less dominant interpretations of the syntactically ambiguous sentences until the disambiguating information comes in at the end of the sentence, whereas weaker WM might support maintenance of only the dominant (and often incorrect) interpretation (Just & Carpenter, 1992).

Moreover, high WM might be able to support flexible, *dynamic* adjustment of filtering, based on the current task demands. If WM is associated with top-down *control* of attentional focus (Kane & Engle, 2003), then greater WM should be associated with greater control of the focus of attention, allowing high WM participants to focus narrowly or broadly, depending on current task goals (Colflesh & Conway, 2007; Cowan, 2005). Top-down attention, for processing only the task-relevant information, and bottom-up attention, for automatically capturing all salient (but not necessarily task-relevant) information work hand-in-hand to process incoming stimuli, possibly via synchronous firing of prefrontal and posterior parietal regions (Buschman & Miller, 2007). It is therefore possible that high WM participants may be better at orchestrating this relationship, to allow for top-down attention in situations that require high filtering, and to support bottom-up attention when processing other salient features is necessary for optimal task performance. The control signal that governs access to WM may stem from preparatory activity in the prefrontal cortex (PFC) and the basal ganglia (BG), and this activity is weaker in low WM individuals (McNab & Klingberg, 2008), resulting in irrelevant items unnecessarily entering WM. Thus, low WM should be associated with less control, leading low WM participants to have a relatively constant and low level of filtering, and biasing these individuals toward automatic bottom-up attention, which processes all salient stimuli regardless of what is currently task-relevant (Fukuda & Vogel, 2009; Zanto & Gazzaley, 2009).

However, WM is not conceptualized here as a panacea for improving performance across *all* cognitive tasks. The advantages conferred by WM should be limited to behaviors that the PFC is known to support, including top-down control for maintenance of task-relevant information (e.g. Buschman & Miller, 2007) and categorical, discrete or abstract processing (e.g. Badre, Kayser, & D'Esposito, 2010; Green, Fugelsang, Kraemer, Shamosh & Dunbar, 2006). Tasks that require strong attention to graded perceptual details should be supported by posterior cortical regions (e.g. Schacter & Buckner, 1998; Wiggs & Martin, 1998), and should, therefore, not be affected by PFC functioning. Thus, low WM participants and children (whose PFC is still underdeveloped) should not show detriments on such tasks, and might even show enhanced performance, if there is a dissociation between PFC-supported active, and abstract representations and posterior-cortical stimulus-specific representations (e.g. Kharitonova, Hulings, & Munakata, in prep).

The idea of high WM supporting both high and low filtering strategies was supported by experiments that found high WM being associated with both greater *selective* attention and greater *divided* attention (Colflesh & Conway, 2007). In the *selective* attention task, participants needed to focus on the auditory stream presented to one ear, while ignoring salient information, such as the participant's name that was presented simultaneously to the other ear; therefore, high filtering of task-irrelevant information (that allowed processing information exclusively from the task-relevant channel) was required. In the *divided* attention task participants needed to focus on two streams of information at once, and shadow information presented to one ear, while pressing a key every time they heard their name in the unshadowed message, presented to the other ear; therefore, low filtering (that allowed processing information from both channels) was the more task-advantageous strategy. These findings provide the first piece of evidence suggesting that high WM can be associated with both high and low filtering strategies, based on current task demands.

However, participants were explicitly told what to do in each case. Thus, it is not clear whether high WM is associated with the ability to dynamically shift filtering strategy, within the context of changing demands in a given task, or whether WM is associated with the ability to preemptively set the focus of attention based on the explicitly stated task goals. Moreover, it is not clear whether this effect is only limited to dichotic listening tasks used in Colflesh & Conway (2007), or whether it extends to other paradigms where filtering task-irrelevant information is required. Therefore, additional studies are needed to explicitly address these possibilities.

Thus, the current experiments were designed to expand on the single finding that suggests that high WM can support both high and low filtering. The experiments in this dissertation were designed to test (1) whether same participants can show both high and low filtering, based on task demands; (2) whether the change in filtering can dynamically occur within the context of a task; and (3) what the developmental roots of this effect are. Because high WM participants can be better than low WM participants at following explicit directions (e.g. via stronger goal maintenance abilities) (Engle, Carullo, & Collins, 1991), the most advantageous strategy for a given task (to filter or not) was never stated in the experiments comprising this dissertation, to test whether high WM participants could spontaneously adjust their filtering strategy, based on changing demands. In addition, several tasks were designed in a way that allowed testing for dynamic, *within-task* adjustment of filtering strategy, to test the strongest version of this theory. Finally, to my knowledge this is the first investigation of the developmental origins of the relationship between filtering and WM.

Examining the relationship between WM and filtering in development

Why is it important to investigate the developmental origins of the relationship between WM and filtering? Understanding developmental trajectories of human behavior can inform our understanding of the underlying mechanisms, given that children often rely on mechanisms that later in development become less apparent due to adults' ability to compensate with other abilities (e.g. Diamond & Kirkham, 2005; Wang & Spelke, 2002). Moreover, understanding the developmental process is imperative for properly understanding the mechanisms supporting the mature system (Karmiloff-Smith, 1998; Karmiloff-Smith et al., 2004), given the complex, interactive nature of the developmental process itself, where small changes may give rise to large (and often unpredictable) outcomes in adulthood (e.g. Elsabbagh & Johnson, 2009).

Studying the development of WM and filtering abilities is particularly fruitful, given their extensive developmental progressions. Both working memory and filtering abilities develop very slowly, with even adolescents showing age-related improvements in working memory tasks (Conklin, Luciana, Hooper & Yarger, 2007). In a developmental visual working memory change detection task, based on the adult version described above and used in the experiments described below (Luck & Vogel, 1997), five-year-old children's capacity (1.5 items) is less than half of the typical number of items remembered by adults (3.8 items), with 10-year-olds' capacity (2.9 items) being in the middle (Riggs, McTaggart, Simpson & Freeman, 2006).

Filtering abilities also develop slowly in children. Filtering improves significantly between four and six years of age (Rueda et al., 2005), but even ten-year-old children take longer than adults to respond when a central stimulus is flanked by incongruent stimuli (Rueda et al., 2004); moreover, children as old as twelve also show deficits on tasks that necessitate filtering of irrelevant information on the Flanker task (Bunge, Dudokovic, Thomason, Vaidya, & Gabrieli, 2002). Even 13-year-old adolescents are more susceptible to processing visual distractors relative to adults (Olesen, Macoveanu, Tegner & Klingberg, 2007).

However, very few studies to date have investigated the relationship between working memory and filtering in young children. One recent study reports that filtering abilities of seven-year-old children are at the adult levels, unless working memory is heavily taxed (Cowan, Morey, AuBuchon, Zwilling, & Gilchrist, 2009), suggesting a non-linear relationship between working memory and filtering in childhood. However, this finding contradicts others that show protracted development of filtering abilities (Bunge et al., 2002; Rueda et al., 2004; Rueda et al., 2005; Olessen et al., 2007); thus, more work is needed to reconcile these findings and better understand the developmental origins of the relationship between WM and filtering task-irrelevant information. Understanding how this relationship develops could inform understanding of the nature of the relationship between the two processes in the mature system, and should thus be studied systematically.

This dissertation will serve as the first step in the direction of examining the developmental trajectory of the relationship between WM and filtering, by testing only two age groups: adults and six-year-old children. Six-year-old children were selected because children at this age have been shown to have somewhat developed, but still imperfect WM (Blackwell, Cepeda, & Munakata, 2009; Riggs et al, 2006). Moreover, six-year-olds are most likely at the point of transitioning to the adult-like proactive WM strategy (Chatham & Munakata, in prep.). Finally, six-year-olds have been shown to perform above floor levels and below ceiling levels on many of the tasks included in this dissertation (Baudouin et al., 2008; Riggs et al., 2006; Rueda et al., 2004). Thus, the six-year-old age group was selected as the first group in which to examine the developmental origins of the relationship between WM and filtering. Child and adult versions of the task were designed to be as similar as possible to each other, to maximize the ability to compare performance across ages. Both younger and older children should be examined in subsequent experiments, with maximally similar tasks, to obtain a fuller picture of the developmental trajectory of the relationship between WM and filtering.

Overview of Experiments

Three experiments were conducted to examine the relationship between WM and filtering. These experiments were designed to test the prediction that high WM can support dynamic adjustment of filtering strategy based on current task demands, and to investigate the developmental origins of this relationship.

Experiment 1 investigated the dynamic filtering account within the same participants, while building closely on established paradigms. In the high-filtering-demand task, participants needed to make same/different judgments regarding the spatial orientation of target items, while needing to ignore distractors. Filtering these distractors was task-advantageous because it decreased WM load. In contrast, in the low-filtering-demand task, participants needed to switch between attending to the shape of a presented item to its color, and vice versa. In some instances, the currently irrelevant feature became relevant on subsequent trials, thus making low filtering (i.e. attending to both color and shape) the more advantageous strategy. Results were consistent with the dynamic filtering account, such that high WM participants showed high filtering in the high-filtering-demand task and low filtering in the low-filtering-demand task.

In Experiment 2, filtering demand was manipulated within a single paradigm. Both Experiment 1 and the previous study exploring this question (Colfesh & Conway, 2007) manipulated filtering demand across different paradigms, thus limiting the ability to interpret the obtained patterns of results solely in terms of filtering demand, since other factors were also varied. Moreover, it was important to avoid explicitly telling participants whether to use high or low strategy in the given task, to test whether high and low WM participants differ in their abilities to determine and adopt the optimal strategy dynamically, within the context of the task. To test these ideas, the previously high filtering demand task was modified to produce both high- and low-filtering-demand versions. In the high-filtering-demand version, paying attention to distractors was disadvantageous because the orientation of distractors did not inform the correct answer, whereas in the low-filtering-demand version of the task, attending to distractors was advantageous because they provided the correct answer on 100% of trials. As in Experiment 1, ERPs were measured during this task. A small change in the design of Experiment 2 had inadvertent effects, which included drawing participants' attention away from the task-relevant location and thus making the ERP results difficult to interpret. RT-based results were somewhat consistent with the predictions, but there were also important caveats to consider in interpreting RTs in relation to filtering. Nevertheless, results from this experiment point to sensitivity of early attention to individual differences in WM, and to subtle changes in experimental setup.

Experiment 3 explored whether the shift in filtering strategy could occur *dynamically*, within the context of performing a task. It also tested for changes in filtering within participants, who were both adults and six-year-old children. In order to assess the role of working memory in filtering in more robust paradigms and to test for dynamic changes in filtering strategy within the context of the same task, experiment 3 tested filtering across two additional tasks, in both children and adults, to test the developmental origins of the relationship between WM and filtering. Results for both kids and adults were mostly consistent with the dynamic filtering account, although several important caveats are discussed.

EXPERIMENT 1: MANIPULATING FILTERING DEMAND WITHIN ADULTS PARTICIPANTS

Experiment 1 was designed to test the prediction that high WM can support dynamic adjusting of filtering strategy, such that it can support high filtering in situations where focusing narrowly is task-advantageous, and low filtering in situations where having a broad attentional scope that allows taking in all of the available information is most appropriate. This was the first attempt to examine whether high WM can support both high and low filtering within the *same* participants, while building closely on existing paradigms. For this first step, high and low filtering demands were manipulated in different paradigms. The high-filtering-demand task was based on the visual WM task developed by Vogel et al. (2005), which required participants to maintain arrays of visual stimuli over a delay. Comparison of arrays with distractors to those without distractors can inform the amount of filtering performed for each participant. Specifically, using high filtering strategy should result in similar memory profiles for the condition with just a few targets and no distractors trials as for the condition with the few targets plus distractors. In contrast, low filtering should result in similar profiles for the condition with many targets and no distractors trials as for the condition with only a few targets plus distractors (since low filtering implies treating distractors as targets in the memory task). The demand for filtering is high in this task because filtering the distractors reduces WM demand.

The low-filtering-demand task was based on the task-switching paradigm (participants needed to switch between attending to the shape of the stimulus and to its color, and vice versa), in which two types of situations were contrasted: (1) the irrelevant feature of a stimulus (e.g. the purple color in the shape-based trial) becomes relevant on the subsequent switch trial, and (2) the irrelevant feature of the stimulus never becomes relevant on the subsequent trial. High filtering (i.e. representing only the currently relevant feature) should not differentiate between the two conditions, because conditions differ only in subsequent relevance of the currently *ir*relevant feature. Alternatively, in the extreme case, high filtering should result should in longer reaction times in the case where the irrelevant feature becomes relevant, relative to the situation where the irrelevant feature never becomes immediately important (e.g. due to negative priming of the currently irrelevant feature). Low filtering (i.e. representing both the relevant and the irrelevant features of a stimulus) should result in faster reaction times when the irrelevant feature becomes relevant on the next trial, compared to when the relevant irrelevant feature never becomes subsequently relevant. This task has a low demand for filtering because representing both the relevant and the irrelevant features of the stimulus helps in the situations when the irrelevant features become relevant on subsequent trials.

Methods

*Participant*s

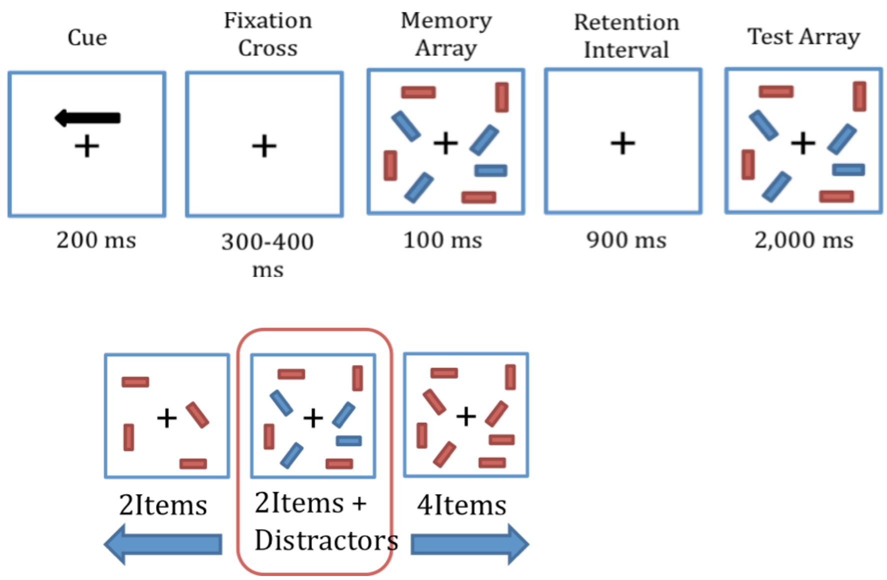
Forty right-handed University of Colorado undergraduate students (23 female) completed this 2-session study, for course credit. In Session 1, EEG was recorded as participants completed the high-filtering-demand task. In Session 2, conducted 4-10 days after the first session, participants completed a behavioral low-filtering-demand task. Three participants failed to return for the second session. Eleven participants were excluded from ERP analyses due to excessive eye movements, indicated by large differences in polarity for the attend-left and attend-right conditions, at the eye channels[1]. One additional participant was excluded from all analyses due to a neurological condition (ADHD).

Materials and Procedure

High-Filtering-Demand Task.

The design of this task was identical to that used in Vogel et al. (2005, see Figure 1A). First, an arrow pointing left or right indicated the side of screen that participants needed to attend (200 ms duration), and was then followed by a random duration fixation cross (300-400 ms time window). Subsequently, arrays of rectangles were presented on both sides of the screen in different spatial orientations (vertical, horizontal or diagonal) for 100 ms, and participants had to maintain the orientation of only the *red* (task-relevant) rectangles on the side of the screen indicated by the direction of the arrow over the 900 ms delay interval. Following the delay, the test array was presented and participants were required to make same/different judgments regarding the spatial orientation of only the task-relevant (red) rectangles on the side of the screen the arrow pointed to. One task-relevant (red) rectangle changed spatial orientation between the sample and test arrays on 50% of the trials. The irrelevant (blue) rectangles and all the rectangles on the opposite side of the arrow never changed from sample to test array.

[1] Eye movements are particularly problematic for this task. Specifically, this task takes advantage of the lateralized organization of the visual system and explores contralateral and ipsilateral activity, relative to the direction of the arrow. Thus, if a participant foveates on the lateralized memory array (e.g. right side of screen), instead of on the central fixation, the notion of contralateral and ipsilateral activity becomes difficult to assess.



(A)

(B)

Figure 1. A: Graphical depiction of the high-filtering-demand task setup. B: Three different trial types used in the high-filtering-demand task. The 2Items+Distractors condition groups with the 2Items condition under high filtering, and with the 4Items condition under low filtering.

Critically, there were three different trial types (Figure 1B): (1) two relevant items in each hemifield (2Items); (2) four relevant items in each hemifield (4Items); and two relevant and two distractor items in each hemifield (2Items+Distractors). This task had a high demand for filtering because ignoring the task-irrelevant (blue) rectangles attenuates working memory load (two vs. four items to remember). Thus, high filtering of the task-irrelevant distractor items is indicated by similar responses in the 2Items and 2Items+Distractors conditions. Low filtering is indicated by similar responses in the 4Items and 2Items+Distractors conditions. Filtering efficiency (FE) can be calculated as the ratio between the difference of activity in the 4Items and 2Items+Distractors condition and the difference of 4Items and 2Items conditions (FE = (4Items -

2Items+Distractors)/(4Items – 2Items). This term approaches zero under low filtering (since 4Items condition is treated just like the 2Items + Distractors, thus producing a difference of zero for the numerator of this ratio), and the term approaches one under high filtering (since 2Items condition is treated like the 2Items + Distractors condition, thus producing equivalent numerator and denominator in this ratio).

As an index of activity for each of the three trial types I used both:

(1). The contralateral delay activity (CDA), measured between 300 and 700 ms after the onset of the delay (most sensitive time window according to McCollough et al., 2006). This ERP component is an index of how much information is maintained in working memory (Vogel & Machizawa, 2004; Vogel et al., 2005).

(2) Reaction times (RT), which I propose should also index activity for each of the trial types, and could thus be used to calculate filtering. The extent to which the RTs from the 2Items + Distractors condition are closer to those from the 2Items condition than the 4Items condition should indicate high filtering.

To estimate WM capacity, we used both:

(1). A behavioral estimate, using Cowan's K (Cowan, 2001) formula: K = S * (H-F), where K is the WM capacity, S is the size of the array, H and F are the observed hit and false alarm rates, respectively. WM capacity (K) was calculated exclusively using performance on the 4Items condition, since performance is nearly perfect for arrays of less than 3 items (Luck & Vogel, 1997) and depending on filtering efficiency, WM demand for the 2Items + Distractors condition is ambiguous.

(2). An ERP estimate of WM (Vogel & Machizawa, 2004), which is calculated as the difference between CDA amplitude for the 4Items, relative to the 2Items conditions. Because the

CDA amplitudes plateau at WM capacity limit of each individual participant (Vogel & Machizawa, 2004), a greater CDA difference between the 4Items and the 2Items conditions is associated with greater WM capacity.

Participants completed two practice blocks (20 trials each) and then completed 720 trials (240 trials for each of the three trial types; intermixed) of the task in a 2-hour-long session. EEG was recorded during the entire task, but was stopped during “blink breaks”, administered every 15 trials, during which participants could rest their eyes. Scalp voltages were collected with a 128-channel HydroCel Geodesic Sensor NetTM connected to an AC-coupled, 128-channel, high-input impedance amplifier (200 MΩ, Net Amps TM, Electrical Geodesics Inc., Eugene, OR). Amplified analog voltages (0.1–100 Hz bandpass) were digitized at 250 Hz. Individual sensors were adjusted until impedances were less than 50 kΩ.

The EEG was baseline corrected to a 200 ms pre-stimulus recording interval and digitally low-pass filtered at 40 Hz. Individual channels were replaced on a trial-by-trial basis with a spherical spline algorithm (Srinivasan et al., 1996). Trials were discarded from analysis if accuracy was incorrect or more than 20% of the channels were bad (average amplitude over 100 μV or transit amplitude over 50 μV). EEG was measured with respect to a vertex reference (Cz), but re-referenced offline to the algebraic average of the left and right mastoids, to be consistent with the previous work (McCollough et al., 2007; Vogel et al., 2005). As in Vogel et al, (2005), CDA was measured at the posterior parietal, lateral occipital and posterior temporal electrode sites (see Figure 2). CDA was calculated as the difference between contralateral and ipsilateral waveforms was analyzed, measured 300-700 ms after the onset of the memory array, which was the most sensitive window in McCollough et al. (2007).

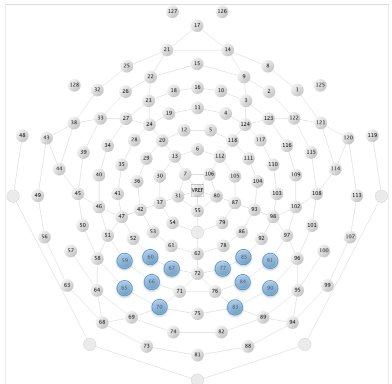


Figure 2. Electrode positions (marked in blue) for CDA amplitudes.

Low-Filtering-Demand Task.

A cued task-switching paradigm was used as a low-filtering-demand task (see Figure 2). Participants responded to either the color (teal, magenta, or yellow) or the shape (cross, diamond, or start) of a centrally presented image, as directed by a cue (C for color or S for shape). On *repeat* trials, participants had to respond to the same dimension (color or shape) as in the previous trial. On *switch* trials, participants had to switch the critical dimension (color or shape, or vice versa) from the previous trial. This task consisted of two types of blocks, which differed in terms of the switch trials; the repeat trials were identical across the two block types. In the *Overlapping* blocks (Figure 3A), the feature of the object that was irrelevant in the previous trial became relevant on the current trial (e.g. Trial 1: shape trial, *yellow* triangle. Trial 2: color trial, *yellow* circle). Thus, filtering the currently irrelevant feature (the *yellow* color) could be disadvantageous in the Overlapping block because that feature became relevant on the subsequent trial. If one ignored this currently irrelevant feature, it might have become more difficult to attend to this feature on the subsequent trial. In contrast, in the *Non-overlapping* blocks (Figure 3B), the irrelevant feature never became relevant on the subsequent trial. Hence, filtering the currently irrelevant information was not disadvantageous in the non-overlapping block.

Filtering efficiency (FE) was thus calculated as the difference in switch costs across the two block types (i.e. overlapping switch cost minus non-overlapping switch cost). Low FE was inferred if participants showed a *smaller* switch cost on the Overlapping blocks, relative to the Non-overlapping ones (i.e. a *negative* difference between the two switch costs). Fast performance on the Overlapping blocks, relative to the Non-overlapping blocks could be achieved if one represented both the relevant (e.g. yellow) and the irrelevant (e.g. triangle) features of a stimulus. Such low filtering should help participants in the Overlapping blocks, since the irrelevant feature becomes relevant on a subsequent switch trial. However, low filtering strategy should not help in the Non-overlapping blocks, since the irrelevant feature does not become relevant on subsequent switch trials. In contrast, high FE profile could be inferred if there was no difference between switch costs for the Overlapping and Non-overlapping blocks, due to these blocks only differing in their treatment of the currently irrelevant feature. In addition, high filtering profile could also be manifested by participants showing a *larger* switch cost for the Overlapping, relative to the Non-overlapping block (i.e. a *positive* difference between the two switch costs). Participants should be slowed down in the Overlapping blocks relatively the Non-overlapping blocks if they fail to actively represent the irrelevant feature (hence, high filtering), which subsequently becomes relevant, in the Overlapping block.

Participants first completed 90 practice trials; they needed to answer at least 85% of trials correctly in order to move on to the next set. In the first 30 practice trials, participants needed to respond only to the shape of the stimulus; in the second 30 practice trials they had to respond only to the color of the stimulus, and in the third 30 practice trials they had to respond to either the shape or the color of the stimulus, based on the cue. The cue appeared 800 ms before the onset of each stimulus; each stimulus remained on the screen until the participant made a response. Feedback was given for incorrect answers ("Incorrect" printed in red ink in the middle of the screen) to encourage participants to maximize their accuracy. Trials were organized in four blocks of 52 trials each (26 switch and 26 repeat trials); the blocks proceeded in the fixed order: Non-overlapping, Overlapping, Non-overlapping and Overlapping. Before the start of each block, the experimenter reminder the participants to "prepare as much as possible and try to answer as accurately as possible on every trial".

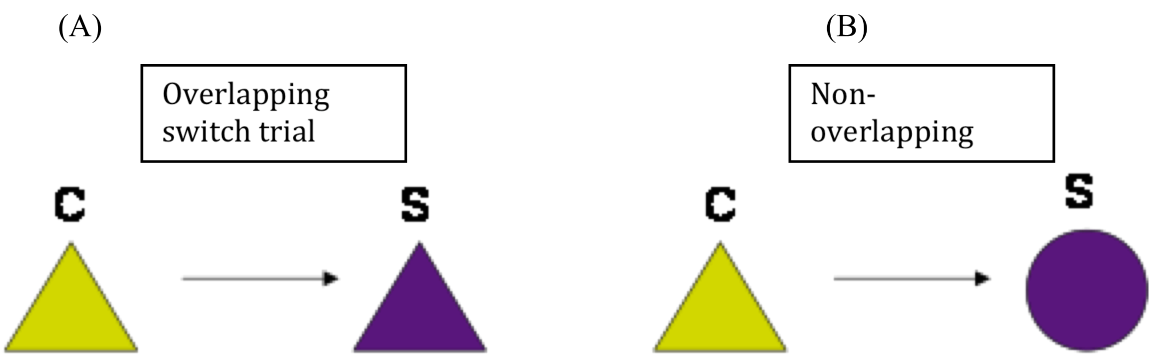


Figure 3. A: Switch trial in the overlapping block. B: Switch trial in the non-overlapping block. FE = switch cost on the overlapping block minus switch costs on the non-overlapping block. FE > 0: High filtering. FE < 0: Low filtering.

Data Trimming

For all reaction time data (including filtering RT measure and task-switching filtering measure), several steps were implemented to trim the data. First, all incorrect trials, plus the trial immediately following the incorrect trial were removed. Then, trials with RTs shorter than 100 ms and longer than 3000 ms were removed. Finally, RTs exceeding 3 standard deviations for each condition and each participant were removed. In addition, in the task-switching filtering measure, the first trial of each block and the two trials following it were removed.

For each linear regression we removed potential outliers (influential points) by running the regression through the Standardized DfFit procedures once, followed by a second run only in the analysis investigating the relationship between ERP filtering and WM, since the data were particularly noisy for the ERP filtering measure. The first round in that analysis eliminated only very extreme values (such as 7, when most values fall in the 0 to 1 range), but not less extreme but nevertheless influential data points. Each run of the Standardized DfFit procedure determined an influence statistic for each participant (i.e. how much each data point influenced the resulting fit of the regression line). According the SPSS guidelines, I eliminated participants from the regression whose Standardized DfFit values exceed the absolute value of 2 times the square root of P * N, where P is the number of parameters in the model and N is the number of participants.

Results

Consistent with our predictions, for the high-filtering-demand (ERP) task a positive correlation was observed between WM and filtering (based on both ERP and RT measures). In contrast, a negative correlation between WM and filtering was observed for the low-filtering-demand (task-switching) task.

High-filtering-demand task

Behavioral results

Participants performed well, with overall accuracy of 87% (similar to McCollough et al., 2006) (see Table 1) and a WM capacity estimate of 2.1 items in the 4Items condition. WM capacity estimates ranged from 0.8 to 3.7 items. High and Low WM capacity participants were identified using a median split for the WM (K) behavioral estimate. The median value for all 40 participants was 1.9 items, and the median for the thirty 30 kept for the ERP analyses was 2.0 items.

	2Items	**2Items + Distractors**	**4Items**
Accuracy (prop. correct)	0.93	0.92	0.76
Mean RT (ms)	777.3	819.5	909.7

Table 1. Group-level behavioral results for the high-filtering-demand task.

ERP results: CDA amplitudes

CDA amplitudes decreased linearly from the 2Items condition (M = -1.1 µV, SD = .6) to the 2Items + Distractors condition (M = -1.4, SD = .7) to the 4Items condition (M = -1.6, SD = .8), $F(1,28) = 29.4$, $p < .001$ (Figure 4).

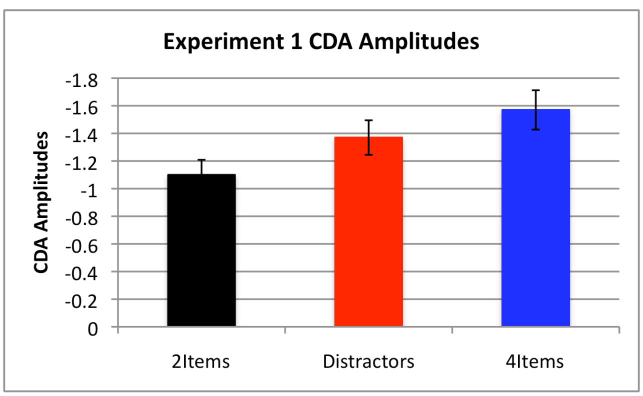


Figure 4. CDA amplitudes for the 3 trial types in Experiment 1.

ERP Results: ERP measure of WM

First, the ERP measure of WM based on CDA amplitudes for the 2Items and 4Items conditions was validated (Vogel & Machizawa, 2004) by replicating the original results. A strong positive correlation (R=.57, p = .01, N = 29) was observed between the behavioral index of WM capacity (K) and the ERP index based on the difference in CDA for the 4Items and 2Items conditions (Figure 5A). Moreover, the difference between 2Items and 4Items CDA waveforms was bigger for the High spans (M=-.66 μV) than for Low spans (M=-.30 μV), t(28) = 2.0, p = .054 (Figure 5B).

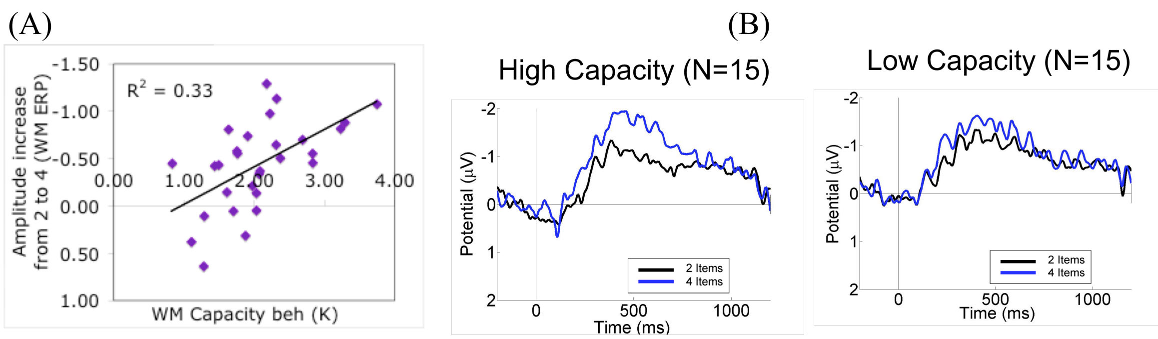


Figure 5. Validating ERP measure of WM capacity, based on CDA differences for the 4Items and 2Items conditions, continuously (A) and in terms of a median split for High and Low WM participants (B).

ERP Results: ERP measure of Filtering

A positive correlation between WM and filtering, under high filtering demand was observed (R = .52, p = .007, N = 25) as in Vogel et al. (2005) (Figure 6A). A median split analysis similarly indicated high filtering in High WM and low filtering in Low WM participants (Figure 6B). Specifically, a significant WM (High vs. Low) by trial type (2Items, 4Items, 2Items + Distractors) interaction (F(2,56) = 3.7, p = .03) revealed that for high spans, all trial types were significantly different from each other (all p's < .02). In contrast, for low spans the 4Items (M = -1.5 μV, SD = .24) and 2Items +Distractors condition (M=-1.5 μV, SD = .20) did not differ from one another (p = .93), despite differences in amplitudes between 2Items and 4Items (p = .06) and between 2Items and 2Items + Distractors (p = .02). This pattern supports the prediction that WM and filtering are positively correlated under high filtering demand[2].

(A) (B)

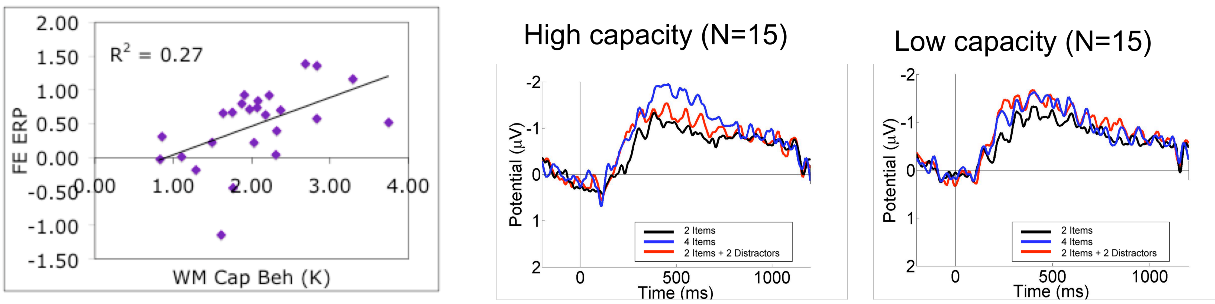


Figure 6. A positive relationship between WM and filtering is observed under high-filtering-demand conditions both continuously (A) and in terms of a median split for High and Low WM participants (B).

RT measure of Filtering

A positive correlation between WM and filtering under high demand for filtering extends to my newly-developed index of filtering based on RTs from correct trials for the three trial types (2Items, 4Items, 2Items + Distractors), using the same formula as for the CDA measure of filtering (Vogel et al., 2005). The RT-based filtering measure was associated with High WM span, when WM was measured behaviorally (R = .36, p = .03, N = 37) (Figure 7A). However, there was no significant correlation between the RT-based measure of filtering and the ERP measure of WM capacity, R = .21, p = .33, N = 24, although the trend was still in the predicted

[2] These results fail to show the pattern of extremely high filtering for high spans that Vogel et al. (2005) demonstrated, such that for high spans, the CDA amplitudes for the 2Items + Distractors condition were identical as for the 2Items condition, while being both significantly lower than the 4Items condition. However, our "high" WM participants were lower in span than high WM participants in Vogel et al.'s study (highest WM capacity was 3.7 vs. 5 items in Vogel et al.); thus, given the prediction that WM and filtering are positively correlated under high filtering demand, it makes sense that our lower capacity "high WM" participants filtered less well.

direction (Figure 7B). Nor was there a relationship between RT-based filtering and ERP-based filtering measures, $R = .08$, $p = .71$, $N = 25$.

(A) (B)

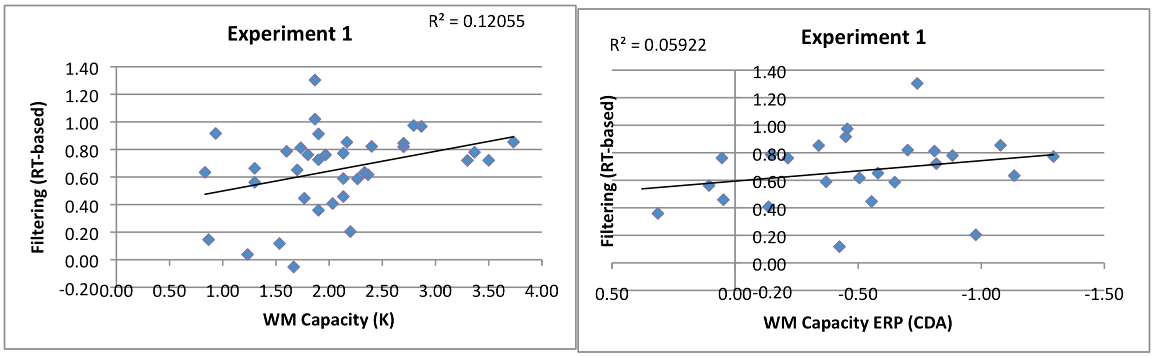


Figure 7. (A): The positive relationship between WM and filtering under high filtering demand extends to RT-based measure of filtering for the behavioral index of WM. (B): The trend in the predicted direction for the RT-based filtering and the ERP-based index of WM.

Low-filtering-demand task

Accuracy

The accuracy for the task-switching filtering task was very high, such that participants answered 94.1% of trials correctly. Accuracy was significantly higher in the repeat (96% correct) than in the switch trials (92% correct), $F(1,36) = 49.7$, $p < .001$. Accuracy did not differ in the Non-overlapping (94.0% correct) and the Overlapping (94.1% correct) conditions, $F(1,36) = .03$, $p = .86$. There was no interaction between the type of trials (switch vs. repeat) and the type of block (Non-overlapping vs. Overlapping), $F(1,36) = .09$, $p = .77$.

Basic reaction times and switch costs

Reaction-time data were submitted to a 2 (Block: Overlapping vs. Non-overlapping blocks) x 2 (Type of trial: Switch vs. Repeat) x 2 (WM: High vs. Low) ANOVA, which showed no effect of Block ($F(1,35) = .29$, $p = .59$) and no interaction between block and WM ($F(1,35) = 2.9$, $p = .10$. There was a main effect of type of trial, such that switch trials ($M = 809$ ms, $SD = 30$) had longer RTs than repeat trials ($M = 746$ ms, $SD = 27$), $F(1,35) = 33.1$, $p < .001$. There was no interaction between the type of trial and WM, $F(1,35) = .72$, $p = .40$. There was also no three-way interaction between block, type of trial and WM. Results were qualitatively the same when the ERP measure of WM (difference in CDA amplitude for the 4Items and 2Items conditions) was used instead of the behavioral WM measure.

Because switch trials had longer RTs than repeat trials, switch costs (switch RT minus repeat RT) were examined. Switch costs for the Overlapping (O) (M=65 ms, $SD = 14$) and Non-overlapping (N) ($M = 61$ ms, $SD = 15$) blocks did not differ from each other, $F(1,35) = .06$, $p = .81$. There was no interaction between the type of trial and WM, $F(1,35) = 2.0$, $p = .16$. Results were again the same when the ERP measure of WM was used.

Filtering measure, based on difference in switch costs

The filtering measure as defined in the Method section, as the difference in switch costs (the switch cost for O block minus the switch cost for N block) was related to the behavioral measure of WM (K), $R = -.43$, $p = .01$, N =35, such that higher WM span was associated with negative O-N switch cost values, while lower WM span was associated with positive values (Figure 8A). When using the ERP measure of WM, results were the same: high WM capacity (negative CDA differences) was associated with negative O-N switch costs, $R = .57$, $p < .001$, $N = 27$ (Figure 8B).

(A) (B)

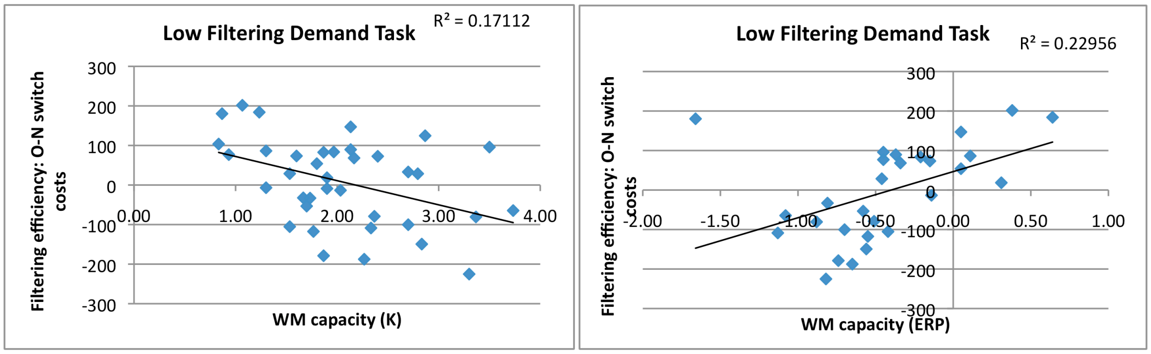


Figure 8. Negative relationships between WM and filtering under low demand for filtering. For both the behavioral index of WM (A) and the ERP index of WM based on difference in CDA for 4Items and 2Items conditions.

These results suggest that under low demand for filtering, the relationship between WM and filtering reverses, such that high WM participants now filter *less* information than low WM participants. Specifically, a negative O-N switch cost for high spans suggests that these participants are faster in the O than the N block, which can only occur if participants represent both the relevant and the irrelevant features of the stimulus, which in turn helps them in the O, relative to the N block.

Discussion

Results from Experiment 1 provide preliminary evidence suggesting that the relationship between WM and filtering depends strongly on filtering demand. Specifically, when the demand for filtering was high, because filtering task-irrelevant items reduced WM demand (as in the ERP filtering task), a positive relationship between WM and filtering was observed. In contrast, when the demand for filtering was low because representing both the currently relevant and the currently irrelevant features of an item may have aided performance on the subsequent trial (as in the task-switching filtering task), the observed relationship between WM and filtering was negative. In the low-filtering-demand task, high WM participants showed a negative difference in switch costs, such that they were faster to switch on the Overlapping, relative to the Non-overlapping blocks, which suggests that high WM participants attended to both the currently relevant and the currently irrelevant features of the stimulus, consistent with the low filtering profile. In contrast, low WM participants showed a positive difference between Overlapping and Non-overlapping switch costs, which is consistent with a high filtering profile.

One possibility is that low spans' attention to the task-relevant features results in negative priming[3] of the currently irrelevant feature, which then becomes relevant on the next trial in the Overlapping block. This possibility may be surprising, given that greater susceptibility to negative priming has previously been associated with high, not low WM (Long & Prat, 2002); however, In the Long & Prat (2002) study, high spans showed a larger susceptibility toward negative priming in the highest conflict condition, where high filtering of the task-irrelevant information was the most task-advantageous strategy. Differences in negative priming susceptibility were not tested in lower conflict conditions; therefore, we do not know whether the effect is due to high spans changing their filtering strategy. Thus, the current results may be due to high spans dynamically adjusting their filtering strategy in the task-switching paradigm to attend to both currently relevant and the irrelevant features, thus appearing to be less susceptible

[3] The reference to negative priming here refers to the *phenomenon* of slower RTs in response to previously ignored features or dimensions, and not the predominant inhibition-based *explanation* of this phenomenon. Directed inhibition does not need to be invoked to explain the phenomenon (MacLeod, 2007), and given that this experiment was not designed to reconcile the inhibition and the non-inhibition accounts of negative priming, I will remain agnostic regarding the nature of this finding.

to negative priming than low spans. In summary, results from Experiment 1 suggest dynamic updating of filtering strategy, based on the current task demands, consistent with the predictions.

However, the substantial differences in the two filtering paradigms prevent definitive conclusions regarding the nature of the relationship between WM and filtering. Many factors other than the different demands for filtering could have contributed to the obtained results. For example, the amount of perceptual information present in each task, and the varying memory requirements could be affecting the amount of filtering observed in each task. Specifically, high perceptual load facilitates filtering irrelevant information, while high working memory load impairs filtering (Lavie, 2005; Yadon, Bugg, Kisley, & Davalos, 2009). The task-switching filtering paradigm has both lower perceptual load than the ERP filtering task (one object on the screen vs. 2-4 objects on each side of the two sides of the screen) and lower memory load (no obvious memory requirements since all available information is available on the screen vs. the requirement to maintain information over the 900 ms delay interval). It is unclear exactly how these group-level differences could relate to individual differences in performance on each task, but one possibility is that low WM participants could be more sensitive to the memory aspect of the task and thus show high filtering in the task-switching paradigm because the low working memory demand of that task permits it, but low filtering in the ERP filtering task, where the high working memory demand completely overloads low spans' memory and precludes filtering. In contrast, the high WM participants could be more sensitive to the perceptual differences in tasks, since the memory load differences may not be highly relevant for them. Higher sensitivity to the perceptual aspects of the task predicts low filtering in the task-switching paradigm and high filtering in the ERP task, which is consistent with the findings.

Alternatively, differences in instructions could have also contributed to the obtained results. There were relatively specific instructions to attend to the *red* and ignore the *blue* items in the ERP task and relatively vague instructions to prepare as much as possible and answer as quickly and as accurately as possible in the task-switching paradigm. Nothing was stated about the differences in the Overlapping and the Non-overlapping blocks. High spans might pay more attention to the specificity of instructions (i.e. to do as they are told), resulting in high filtering in the ERP task and low filtering in the task-switching paradigm. In contrast, low spans might disregard (or not remember) the instructions and try to infer the best strategy in the process of performing the task. This may result in low spans' low filtering in the ERP task (where it is difficult to figure out what do without knowing instructions) and relatively good filtering in the task-switching paradigm (where participants receive negative feedback on incorrect trials, thus making it easier to infer the requirements of the task).

In general, a number of differences in the setup of the two filtering tasks preclude strong conclusions about the relationship between working memory and filtering. Thus, to determine whether high WM can support both high and low filtering based on task demands, filtering demand must be manipulated within a single paradigm. Experiment 2 was designed for this purpose.

EXPERIMENT 2: MANIPULATING FILTERING DEMAND WITHIN A SINGLE PARADIGM: ERP STUDY WITH ADULTS

Experiment 2 was designed to build on Experiment 1 findings in order to continue investigating the relationship between WM and filtering task-irrelevant information across both low and high filtering demands. Experiment 1 manipulated filtering demand within participants but across different paradigms. Experiment 2 was designed to manipulate filtering demand within the *same* paradigm, in order to minimize the possibility that factors other than filtering demand were affecting the results. The previously high-filtering-demand task from Experiment 2 was modified to produce both high- and low-filtering-demand versions. If high WM supports dynamic updating of filtering strategy based on task demands, then high WM participants should show a high filtering profile in the high-filtering-demand version and a low filtering profile in the low-filtering-demand version of the task. Low WM participants are expected to show a consistently low filtering profile.

Methods

Participants

Seventy-one right-handed University of Colorado undergraduate students (44 female) participated in this two-session experiment. EEG was recorded during the first session. The second session was administered 4-10 days after the first session and involved several behavioral measures, which are described and analyzed in the context of Experiment 3. Thirty-three students received course credit, while 38 students were paid $15/hour for the EEG session (which lasted approximately 2.5 hours) and $10 for the entire behavioral session (which lasted approximately 2 hours). Of these participants, seventeen were excluded from EEG analyses for excessive eye movements (six in the Orthogonal condition and eleven in the Correlated condition), and one additional participant (Correlated condition) was excluded due to neurological disorders (ADHD) and overall very low accuracy.

Materials and Procedure

The high-filtering-demand task from Experiment 1 was modified such that filtering demand could be manipulated within this single paradigm. The task was administered in the same way as in Experiment 1, with two critical differences:

(1) Filtering demand was now manipulated across two conditions within the same paradigm to create both *high-* and *low*-filtering-demand versions of this task. In both conditions the distractor items changed on 50% of the trials, unlike in Experiment 1, where distractors never changed across the memory and test arrays (Figure 9A).

Critically, in the *Orthogonal* condition, intended to create a high demand for filtering, the distractors changed *orthogonally* to the targets (Figure 9B). Thus, both distractors changed or stayed the same regardless of whether the targets changed or stayed the same. For example, if a target item changed (signifying a different trial), both distractors changed on 50% of the trials, and stayed the same on the remaining 50% of trials. Thus, distractors were not predictive of type of trial (same vs. different) and should thus have been ignored for optimal performance in the Orthogonal condition. In contrast, in the *Correlated* condition, distractors changed *consistently* with targets (Figure 9C). Thus, both distractors stayed the same on 100% of *same* trials (where the target stayed the same) and changed on 100% of *different* trails (where the target changed orientation across the memory and test arrays). Thus, distractors were 100% predictive and thus did not need to be filtered for optimal performance.

Orthogonal and Correlated conditions were manipulated between participants for both theoretical and pragmatic reasons. Because it is not yet clear how rapidly participants can update their filtering strategy (if at all), filtering demand was manipulated between participants to avoid unnecessary carryover across condition. In addition, each condition (correlated and orthogonal) takes 2.5 hours to complete given the large number of trials that are needed to obtain a clean ERP signal; thus, it was impractical to combine both conditions in a single 5-hour session.

In both conditions participants were instructed to attend to the *red* items. They were told that the *blue* may or may not change. The instructions were purposefully vague with respect to dealing with the blue distractors to both capitalize on the individual differences in figuring out the task demands (which could be related to individual differences in WM) and to avoid reducing the distracting nature of the blue items (i.e. if participants in the Correlated condition were told that paying attention to the blue items might be beneficial, the nature of the task might change, as it no longer contained targets *and* distractors). It is informative to determine whether high WM can support a spontaneous use of the high or low filtering strategy, based on current task demands.

If high WM span can supporting dynamic updating of filtering strategy based on task demands, high WM participants should show high filtering in the Orthogonal condition and low filtering in the Correlated condition.

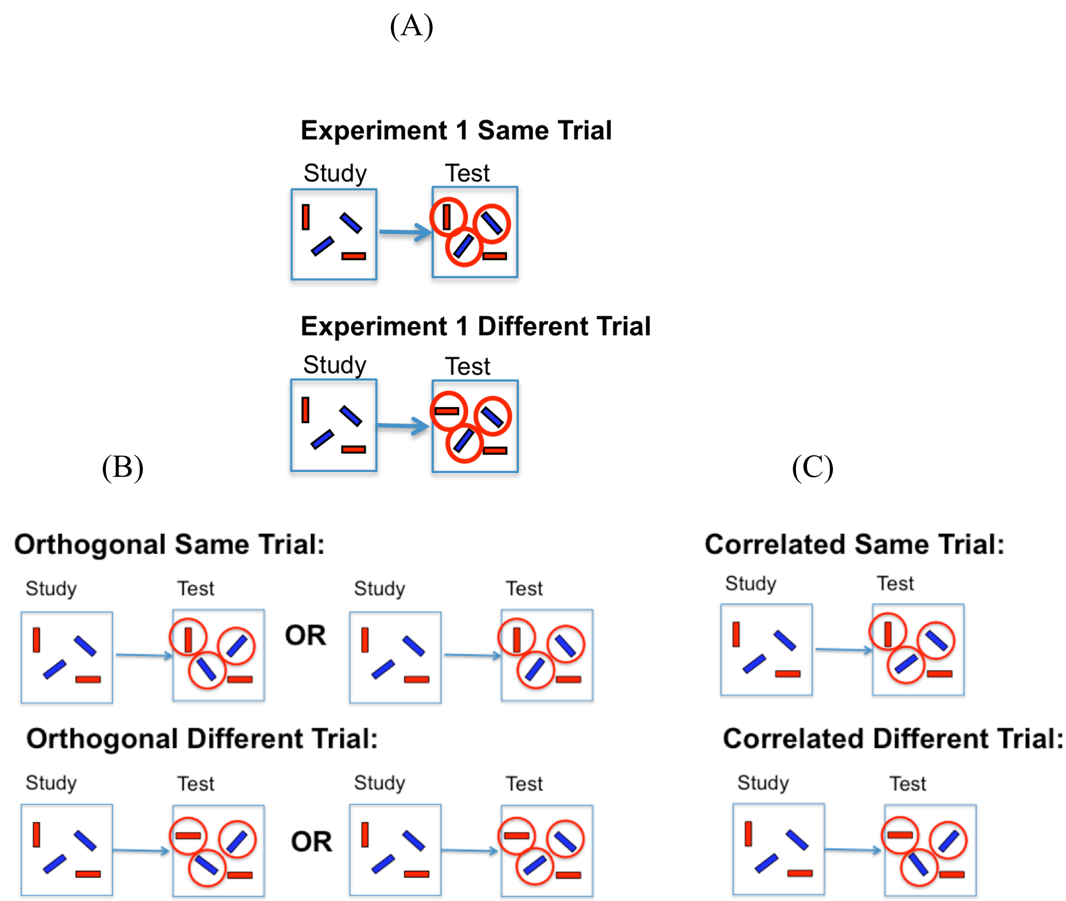


Figure 9. Same and Different trials in the Orthogonal (B) and Correlated (C) conditions, compared to those in Experiment 1 (A), which were the same as Vogel et al. (2005). Circled red items are examples of a target item. Blue items are distractors, and are circled to highlight the fact they changed either orthogonally (in Orthogonal condition) or consistently (in Correlated condition) in Experiment 2. Distractors did not change at all across study and test displays in Experiment 1. For simplicity, only one hemifield is depicted here; the other hemifield contains the same number/types of items, but none of the items changed across memory and test arrays.

(2). The duration of the arrow indicating the side of the screen to which the participants needed to attend was increased from 200 ms in Experiment 1 to 700-1100 ms (random, within this window) in Experiment 2 (see Figure 10). The memory array immediately followed the arrow. The exclusion of the post-arrow fixation (as in Experiment 1) was inadvertent; the duration of the arrow was increased in an attempt to minimize participants' eye movements (see footnote 1 in Experiment 1), which may have increased following the inadvertent exclusion of the post-arrow fixation.

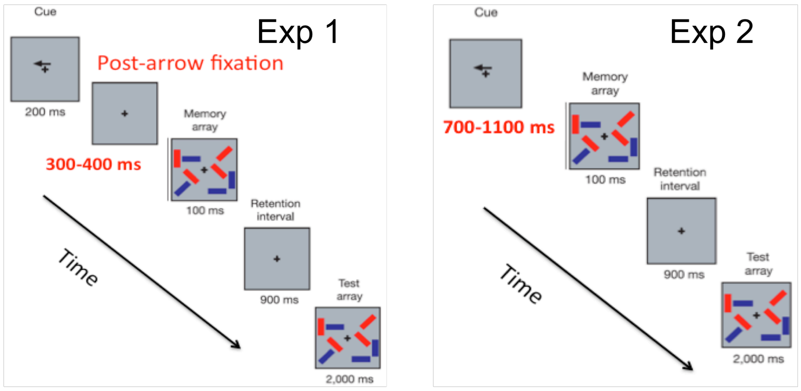


Figure 10. Extended cue duration in Experiment 2.

ERP Analyses

The CDA component was analyzed in the same way as in Experiment 1.

A few additional ERP components were analyzed in the follow-up analyses for this experiment. These included the N2pc component, which shared the same electrodes as the CDA (see Figure 11A) and the P1 and N1 peaks, which were calculated from a slightly different set of posterior electrodes (see Figure 11B, 11C), where P1 and N1 were found to be largest. ERPs were time-locked to the onset of the memory array for the CDA and N2pc components. ERPs were time-locked to the onset of the arrow presentation for the P1 and N1 peaks. The N2pc amplitudes were measured at 220 to 300 ms after the onset of the memory array (consistent with Eimer, 1996). P1 peaks were measured as the maximum amplitude between 0 and 165 ms after the onset of the arrow. N1 peaks were measured as the minimum amplitude between 115 and 205 ms after the onset of the arrow.

(A) (B) (C)

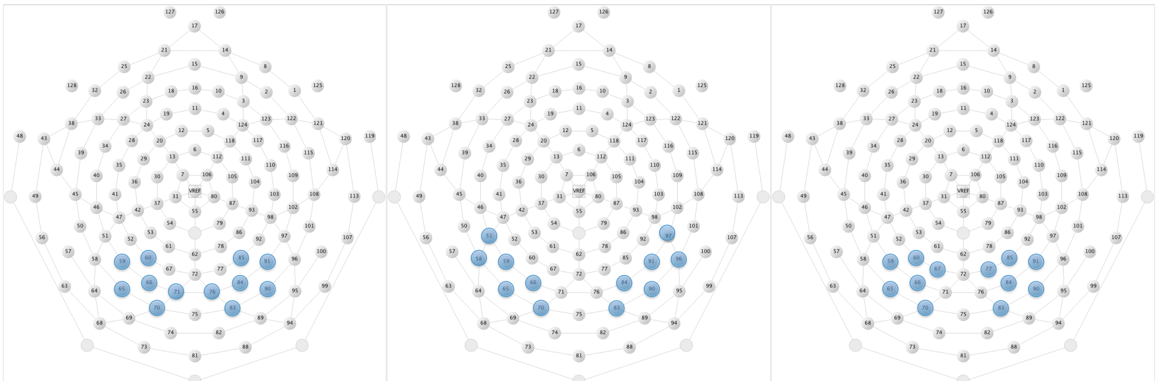


Figure 11. Electrode positions (marked in blue) for P1 (A), N1 (B) and N2pc (C) components.

Data trimming and corrections

Reaction time data were trimmed in the same way as in Experiment 1. For the ERP data, all p-values from repeated-measures ANOVAs were corrected for violations of the sphericity assumption using Geisser and Greenhouse's (1958) method.

Results and Discussion

As elaborated below, the results of Experiment 2 did not support the predictions. In fact, the relationship between the behavioral and the ERP measures of WM, and between WM and filtering did not replicate in this experiment, despite similar behavioral performance. Results for Experiment 2 will be presented alongside relevant results from Experiment 1, in order to compare the patterns and determine the underlying reasons behind the unexpected results. Specifically, the CDA amplitudes overall were significantly lower in Experiment 2 than in Experiment 1, due to greater *ipsilateral* amplitudes, suggesting less lateralized visual memory in Experiment 2. Follow-up analyses indicated less vigilant attention to the direction indicated by the arrow (which told participants which side of screen to attend to for the memory array), more processing of the task-irrelevant side during the arrow presentation, and less overall attention to the relevant side of space in Experiment 2, particularly for the low WM participants. Thus, less compliance with the stated instructions in Experiment 2 (e.g. to attend exclusively to the side indicated by the arrow) made it difficult to definitively test for the predicted patterns in the relationship between WM and filtering task-irrelevant information; nonetheless, results from this experiment point to sensitivity of early attention to individual differences in WM, and to subtle changes in experimental setup.

Behavioral results – Accuracy

Participants performed well, with overall accuracy of 87% in the Orthogonal condition and 87.3% in the Correlated condition (See Figure 12 for comparison of accuracy estimates across experiments and trial types). A 3 (trial type: 2Items, Distractors, 4Items) by 3 (experiment: Experiment 1, Experiment 2 – Orthogonal, Experiment 2 – Correlated) ANOVA revealed no main effect of experiment on accuracy, $F(2,106) = .25$, $p = .78$. There was a main effect of trial type, $F(2,212) = 729.6$, $p < .001$. LSD post-hoc tests revealed that accuracy on the 4Items trials (M = 77.1% correct) was worse than accuracy on both the Distractors trials (M = 92.0%) and the 2Items trials (M = 92.5%), $p < .001$; accuracy on the Distractors trials was only marginally worse than that on the 2Items trials, p = .06. There was no interaction between trial type and experiment, $F(4,212) = 1.9$, $p = .12$.

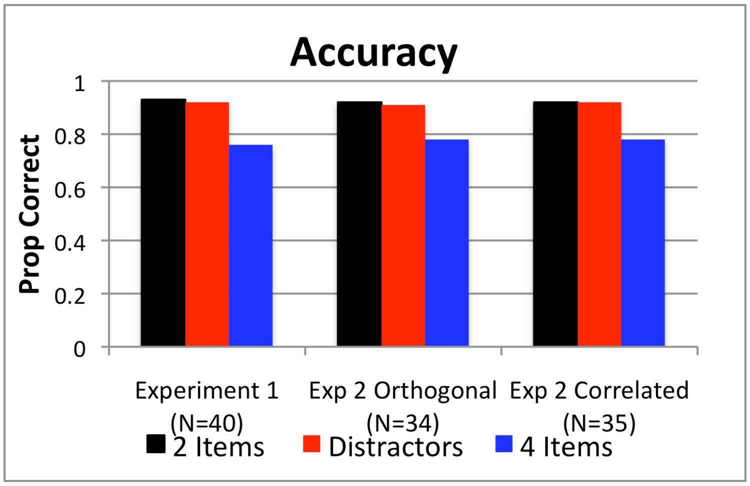


Figure 12. Accuracy across trial types and experiments.

The average WM capacity estimate (based on the 4Items condition accuracy) was 2.23 items in the Correlated condition and 2.21 items in the Orthogonal condition. These estimates were not significantly different from each other (p = .88), and did not differ from the estimate of 2.1 items for Experiment 1 (both p's > .2). The ranges and medians of WM capacity estimates were also fairly similar across conditions and experiments (see Table 2).

	Min K	**Max K**	**Mean K**	**Median K**
Experiment 1	0.83	3.73	2.06	2.0
Experiment 2 – Correlated	1.23	3.0	2.23	2.38
Experiment 2 – Orthogonal	0.47	3.43	2.21	2.18

Table 2. Range of WM (K) estimates across experiments and conditions.

The median WM capacity (K) value from Experiment 1 (M = 2.0) was used to divide participants into High and Low WM participants in order to have consistent criteria across experiments. Basic ERP WM and ERP filtering results were also calculated based on a number of different criteria, which yielded similar results, and are included in Appendix A.

Behavioral results – Reaction Times

Reaction times were also very similar across experiments (see Figure 13). A 3 (trial type: 2Items, Distractors, 4Items) x 3 (experiment: Experiment 1, Experiment 2 – Orthogonal, Experiment 2 – Correlated) ANOVA revealed no main effect of experiment, $F(2,106) = .24$, $p = .79$. There was a main effect of trial type, $F(2,212) = 221.6$, $p < .001$. LSD post-hoc tests revealed that all trial types differed from one another (all p's < .001), such that 2Items trials were performed most quickly ($M = 761$ ms), followed by Distractors trials ($M = 798$ ms), followed by 4Items trials ($M = 889$ ms). There was no interaction between trial type and experiment, $F(4,212) = .78$, $p = .54$.

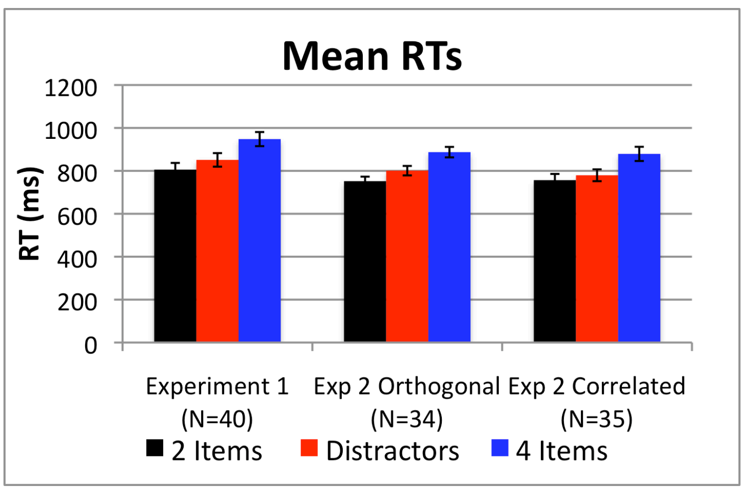


Figure 13. Mean RTs across trial types and experiments.

ERP results: CDA amplitudes

In the Orthogonal condition, the CDA amplitudes decreased linearly, such that they were least negative on the 2Items trials ($M = -.60$ µV, $SD = .57$), more negative on the Distractors trials (M = -.76, SD = .67) and most negative on the 4Items trials ($M = -.81$, $SD = .54$), $F(1,27) = 4.7$, $p = .038$ (Figure 14B). Similarly, in the Correlated condition, the CDA amplitudes decreased linearly from the 2Items trials (*M* = -.59 µV, *SD* = 44) to the Distractors trials (*M* = -.81, *SD* = .59) to the 4Items trials (*M* = -.87, *SD* = .59), $F(1,23) = 10.4$, $p = .004$ (Figure 14C). These overall patterns are consistent with those in Experiment 1 (Figure 14A) and those in Vogel et al. (2005). The differences in CDA amplitudes for the three trial types did not differ by condition (Correlated vs. Orthogonal), $F(2,100) = .16$, $p = .84$.

(A) (B) (C)

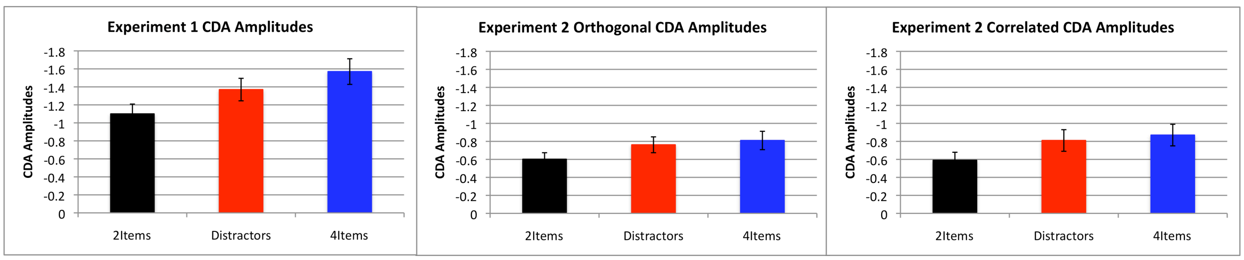


Figure 14. CDA amplitudes for the 3 trial types in the Orthogonal (B) and Correlated (C) conditions of Experiment 2 were smaller than CDA amplitudes in Experiment 1 (A), despite overall similar patterns across trial types.

Even though the pattern of decreasing amplitudes across the three trial types was consistent with Experiment 1, amplitudes overall were significantly lower in Experiment, $F(1,80) = 22.7$, $p < .001$ (Figure 14A). There was a marginal interaction between experiment and trial type, $F(2,160) = 2.6$, $p = .08$, such that the differences between trial type amplitudes were more compressed in Experiment 2 than in Experiment 1 (see Table 3).

	2Items	**Distractors**	**4Items**
Experiment 1 (N=30)	-1.09 (.10)	-1.38 (.12)	-1.57 (.12)
Experiment 2 (N=52)	-0.6 (.08)	-0.79 (.09)	-0.84 (.09)

Table 3. CDA amplitudes (and standard errors in parentheses) for both experiments. Amplitudes were smaller (less negative) in Experiment 2, with less separation across the three trial types.

Since the CDA comprises a difference between contralateral and ipsilateral activity, the smaller CDAs in Experiment 2 suggest less attention to the contralateral (task-relevant) side of space, more attention to the ipsilateral (task-irrelevant) side of space, or both. To determine the reason behind smaller CDA differences in Experiment 2, contralateral and ipsilateral activities were analyzed separately for each study and WM span (see Figure 15).

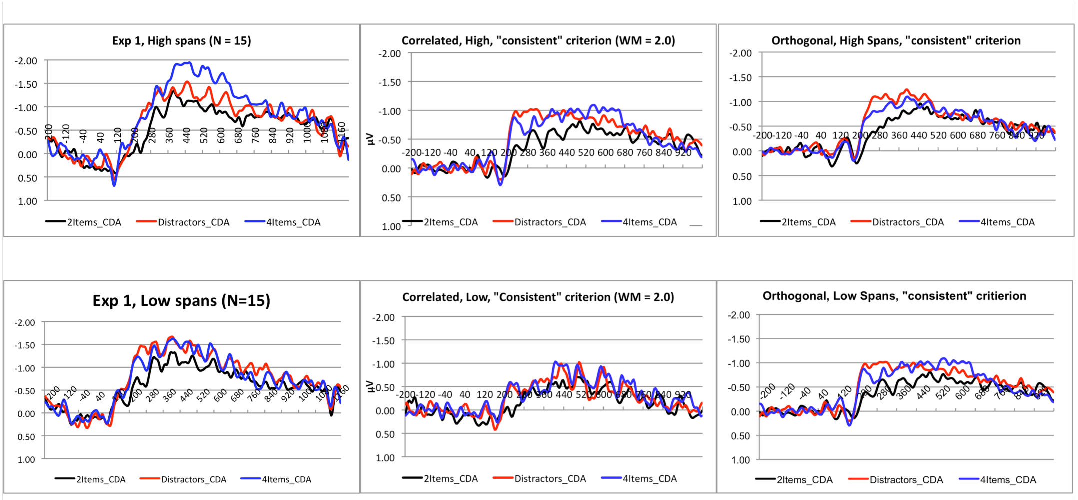
Figure 15. Contralateral and ipsilateral ERP waveforms (A) and mean contralateral and ipsilateral ERP amplitudes (B) in the CDA time window (300-700 ms after onset of memory array) for Experiment 1 and 2, across the three types of trials. Contralateral activity did not change as a function of experiment, whereas ipsilateral activity was larger in Experiment 2. Larger ipsilateral activity and same contralateral activity explains smaller CDA difference amplitudes and suggests less lateralized WM processing in Experiment 2.

In all experiments/conditions, there was a main effect of laterality, such that contralateral activity was less positive than ipsilateral activity, all p's < .001, consistent with the findings of Vogel et al. (2005). Contralateral amplitudes did not change significantly as a function of experiment (Experiment 1 vs. 2), $F(1,80) = .037$, $p = .85$. However, ipsilateral activity was marginally less positive in Experiment 2 ($M = .86$ μV, $SE = .22$) than in Experiment 1 ($M = 1.48$ μV, $SE = .28$), $F(1,80) = 3.04$, $p = .085$. There was no interaction between experiment and trial type (2Items, Distractors, 4Items) on either the contralateral ($F(2,160) = .62$, $p = .52$) or the ipsilateral ($F(2,160) = .03$, $p = .96$) activity. A follow-up ANOVA with Experiment 2 condition (Orthogonal vs. Correlated) as factor revealed no effect of condition, and no interaction between trial type (2Items, Distractors, 4Items) and condition, for both contralateral and ipsilateral activities, all p's > .34. Thus, it appears that visual WM was less lateralized in Experiment 2, with no difference between the Orthogonal and the Correlated conditions, such that ipsilateral activity became similar to contralateral activity in Experiment 2.

ERP results: ERP measure of WM

The ERP measure of WM was calculated as in Experiment 1 and as in Vogel & Machizawa (2004), by subtracting CDA amplitudes for 2Items from CDA amplitudes for 4Items trial types. Unlike in Experiment 1, however, this index did not correlate with the behavioral estimate of WM capacity (see Figure 16, and Appendix A for different high/low WM breakdowns), in either the Orthogonal condition ($R = .15$, $p = .47$, $N = 26$) or the Correlated condition ($R = .19$, $p = .41$, $N = 22$).

(A)



(B)

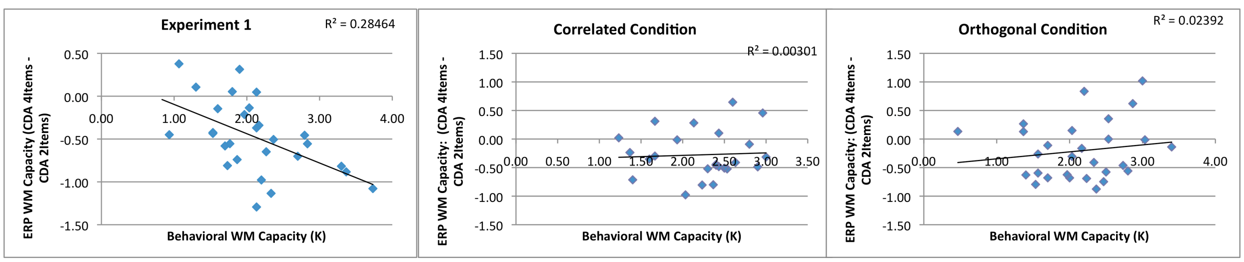


Figure 16. (A): ERP waveforms of the CDA component (relevant time window is 300-700 ms) show a lack of WM modulation in Experiment 2, in contrast to Experiment 1 results. (B) Lack of a relationship between behavioral and ERP indices of WM in Experiment 2, in contrast to the strong relationship in Experiment 1.

ERP Results: Filtering

The dynamic filtering account predicted a positive relationship between WM and filtering for the Orthogonal (high filtering demand) condition and no relationship between WM and filtering for the Correlated (low filtering demand) condition. Filtering was calculated as in Experiment 1 and as in Vogel et al. (2005) by taking a ratio of the difference between CDAs for the 4Items and Distractors trials and the 4Items and 2Items trials. Thus, this measure of filtering was dependent on having sufficient distance between the CDA amplitudes for the 4Items and the 2Items trials in order to have sufficient variability in how close or far the CDA for the Distractors trials falls relative to that for the other two trial types. As the above analyses (and Appendix A) demonstrate, the distance between the 4Items and 2Items trials was overall smaller in Experiment 2, making it unlikely to find meaningful filtering results. However, the ERP filtering measures were examined nevertheless.

Unlike in Experiment 1, there was no correlation between WM and filtering, in either the Orthogonal (R = .099, p = .65, N = 24) or the Correlated conditions (R = .14, p = .56, N = 19) (see Figure 17).

(A) (B) (C)

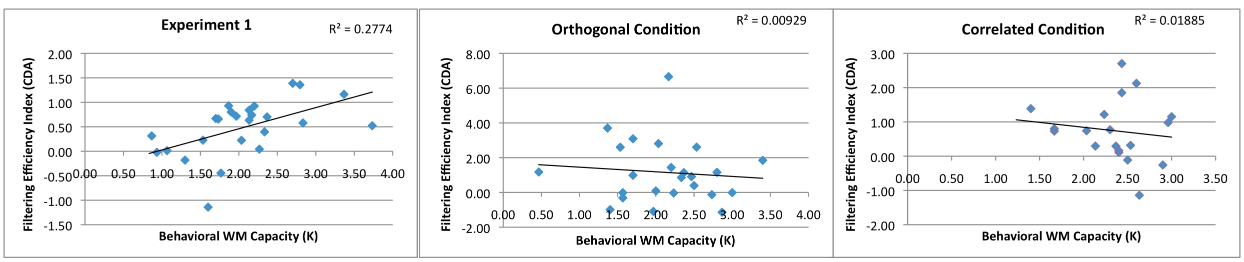


Figure 17. Lack of a relationship between WM and filtering in both Orthogonal (B) and Correlated (C) conditions of Experiment 2, in contrast to the positive relationship between WM and filtering in Experiment 1 (A)

To better understand the reasons underlying the lack of the relationship between WM and filtering in Experment 2, the relationship between WM and each of the three types of trials was analyzed seprately, as a function of the categorical high and low WM span. The analysis of high versus low WM capacity individuals' CDA amplitudes for the three trial types similarly reflected a lack of a relationship between WM span and filtering (see Figure 18 to compare CDA amplitudes for High vs. Low spans for different trial types and experiments; see Appendix A for

ERP waveforms for a number of different high/low WM criteria). A 3 (trialtype: 2Items, Distractors, 4Items) x 2 (WM: high, low) x 2 (condition: Orthogonal, Correlated) repeated measures ANOVA revealed only a main effect of trial type, $F(2,96) = 6.2$, $p = .004$, such that 2Items ($M = -.55$ µV) and Distractor ($M = -.72$ µV) CDAs were significantly different from each other (p = .008), as were the 2Items and the 4Items ($M = -.79$ µV) CDAs ($p = .001$), whereas the Distractor and the 4Items CDA ($M = .65$ µV) were not signficiantly different from each other, $p = .33$, consistent with the low filtering profile overall. There was a significant effect of WM span, $F(1,48) = 4.4$, $p = .04$, such that amplitudes were higher for high WM ($M = -.84$ µV, $SE = .08$) than for low WM ($M = -.54$ µV, $SE = .12$) participants. There was no effect of condition (Orthogonal vs. Correlated), $F(1,48) = .06$, $p = .82$ and no interactions (all p's > .46).

(A) (B)

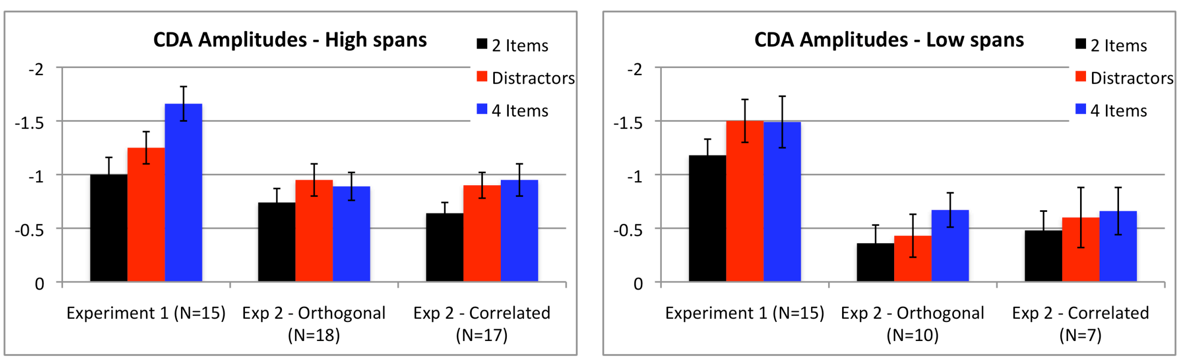


Figure 18. CDA amplitudes for High WM (A) and Low WM (B) participants across the three trial types and the 3 experiments/conditions. Experiment 2 amplitudes were lower overall and there was no relationship between WM and filtering (i.e. the ratio of the difference between 4Items amplitudes and Distractors and the difference between 4Items trials and 2Items amplitudes).

RT-based measure of filtering

As in Experiment 1, an RT-based measure of filtering was calculated to determine the relationship between filtering and WM. Thus, the same formula for calculating the filtering index was used (Vogel et al., 2005), but RTs were used instead of CDA amplitudes. In Experiment 1 there was a positive relationship between WM and both the ERP and the RT-based filtering measures (Figure 19A). Experiment 2 results appeared to be consistent with the predictions (see Figure 19B,C): a positive relationship between WM and filtering for the Orthogonal condition (R = .39, N = 33, p = .02) and no relationship between WM and filtering for the Correlated condition (R = .005, N = 32, p = .98). Categorical examination of high and low WM (based on the median WM estimate of 2.0 items from Experiment 1, for consistency purposes) revealed that in the Orthogonal condition high WM participants (M = .67, SD = .26) filtered marginally more than the low WM participants (M = .46, SD = .39), $t(31)$ = -1.9, p = .07. In contrast, in the Correlated condition, high (M = .63, SD = .44) and low (M = .83, SD = .38) WM participants did not differ in their amount of filtering, $t(30)$ = 1.1, p = .29.

(A) (B) (C)

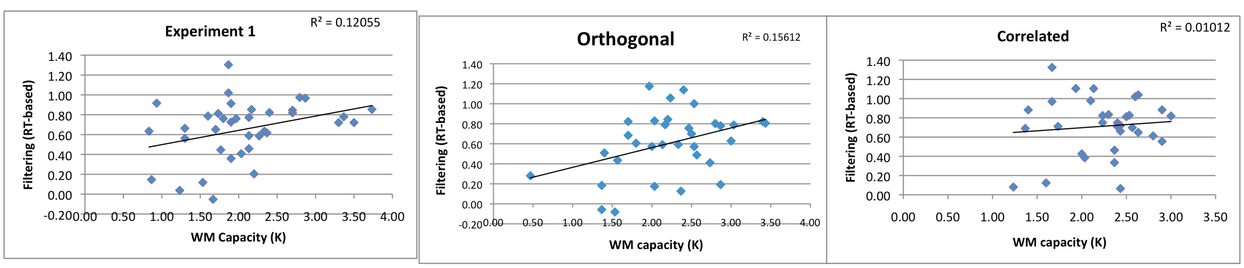


Figure 19. Relationship between WM and RT-based measure of filtering for the Orthogonal and Correlated conditions. There was a positive relationship between WM and RT-based filtering in both Experiment 1 (A) and Orthogonal condition of Experiment 2 (B), but not in the Correlated condition of Experiment 2 (C), consistent with the predictions.

However, a closer look at the data suggests that results do not exactly match predictions. Specifically, the predictions suggested that high WM participants could dynamically change their strategy from high to low, based on task demands. Thus, according to this prediction, in the Orthogonal condition, high spans should show high filtering scores (approaching the score of one with perfect filtering), but in the Correlated condition, high spans should show relatively low filtering scores (approaching the score of zero with no filtering). However, the results show equivalent filtering for high spans across Orthogonal condition (*M* = .67), Correlated condition (*M* = .63), and Experiment 1 (*M* = .70), $F(2,65) = .26$, $p = .78$; all pairwise comparison *p*'s > .5. In contrast, it is the *low* WM participants who seem to be changing their strategy: They appear to be filtering less in the Orthogonal condition (*M* = .46) than in the Correlated condition (*M* = .83), $p = .05$ (neither is significantly different from low span filtering in Experiment 1 (*M* = .61), both *p*'s > .2).

To better understand the relationship between WM and RT-based filtering, the relationships between WM and RTs for individual trial types were examined (see Figure 20). For each experiment, a 3 (trial type: 2Items, Distractors, 4Items) x 2(WM: high, low) repeated measures ANOVA was conducted, to parallel the ERP-based CDA analyses.

In Experiment 1, there was a main effect of trial type, $F(2,76) = 117.6$, $p < .001$, such that RTs for the 2Items trials were shortest (*M* = 773 ms, *SE* = 27.8), RTs, for the 4Items trials were longest (*M* = 901 ms, *SE* = 30.1), and RTs for the Distractors trials were intermediate (*M* = 814 ms, *SE* = 28.0), all pairwise comparison *p*'s < .001. There was no effect of WM on RTs, $F(1,38) = .47$, $p = .50$. However, there was a significant trial type by WM interaction, $F(2,76) = 3.4$, $p = .037$, such that despite RTs for all trial types being significantly different from each other for both high and low spans (all pairwise *p*'s < .001), the difference between 4Items and Distractors RTs was larger for high spans (difference = 108 ms) than for low spans (difference = 68 ms), whereas the differences between 2Items and Distractors were the same (40 ms for high spans and 42 ms for low spans) (see Table 4). This pattern of results is consistent with higher filtering for high WM than low WM participants, because high WM participants treated the Distractors trials more like the 2Items trials than the 4Items trials.

	2Items RT	Distractors RT	4Items RT
High Spans	747 ms (33.8)	787 ms (36.4)	895 ms (32.4)
Low Spans	798 ms (43.2)	840 ms (41.9)	908 ms (49.3)

Table 4. Mean RTs for each trial type for High and Low WM participants for Experiment 1. Standard errors of the mean are in parentheses. RTs for all trial types are significantly different from each other, all p's < .001.

In the Orthogonal condition of Experiment 2, there was also a significant effect of trial type, $F(2,64) = 75.4$, $p < .001$, such that RTs for 2Items trials were fastest (M = 761 ms, SE = 22.3), followed by Distractors trials RTs (M = 811 ms, SE = 23,4), and RTS for the 4Items trials were slowest (M = 887 ms, SE = 26.4), as in Experiment 1. There was no effect of WM, $F(1,32) = .58$, $p = .45$. There was a significant interaction between WM and trial type, $F(2,64) = 5.0$, $p = .01$, such that despite that RTs for all trial types were significantly different from each other for both high and low spans (all pairwise p's < .01), the difference between 4Items and Distractors RTs was larger for high spans (difference = 105 ms) than for low spans (difference = 48 ms), whereas the differences between 2Items and Distractors were the same (49 ms for high spans and 50 ms for low spans) (see Table 5). These patterns are almost identical to those from Experiment 1 and also suggest higher filtering for high than for low WM participants in the Orthogonal condition of Experiment 2.

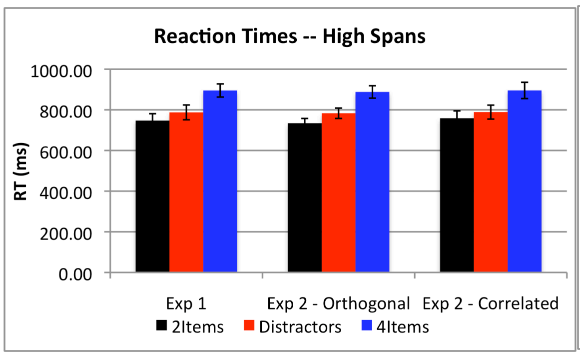
	2Items RT	**Distractors RT**	**4Items RT**
High Spans	734 ms (23.4)	783 ms (25.5)	888 ms (30.4)
Low Spans	788 ms (42.2)	838 ms (42.0)	886 ms (42.5)

Table 5. Mean RTs for each trial type for High and Low WM participants for Experiment 2, Orthogonal condition. Standard errors of the mean are in parentheses. RTs for all trial types are significantly different from each other, all p's $< .01$.

In the Correlated condition of Experiment 2, there was again a main effect of trial type, $F(2,66) = 29.9$, $p < .001$, such that RTs for the 2Items trials were fastest ($M = 755$ ms, $SE = 33.5$), RTs for the Distractors trials only marginally slower ($M = 771$ ms, $SE = 31.9$), $p = .07$, and RTs for the 4Items condition were slowest ($M = 864$ ms, $SE = 38.0$), $p < .001$ for comparing to the other two trial types. There was no effect of WM, $F(1,33) = .28$, $p = .60$. Unlike results for Experiment 1 and for Orthogonal condition of Experiment 2, however, there was no longer an interaction between WM and trial type, $F(2,66) = 1.7$, $p = .19$. Nevertheless, mean RTs and standard errors for each trial type are reported in Table 6. Consistent with the predictions, these RTs suggest that filtering was higher for the *low* spans in the Correlated condition, because low span RTs for 2Items and Distractors trials were the same ($p = .90$), and different from those in the 4Items trials (both p's $= .05$).

	2Items RT	Distractors RT	4Items RT
High Spans	759 ms (36.0)	789 ms (34.3)	895 ms (40.0)
Low Spans	751 ms (41.6)	753 ms (42.6)	832 ms (57.3)

Table 6. Mean RTs for each trial type for High and Low WM participants for Experiment 2, Correlated condition. Standard errors of the mean are in parentheses. For high spans, RTs are all significantly different from each other, all p's < .003. For low spans, 2Items RTs and Distractors RTs were not significantly different from each other ($p = .90$). 4Items RTs were significantly longer than both 2Items and Distractors RTs, both p's = .05.



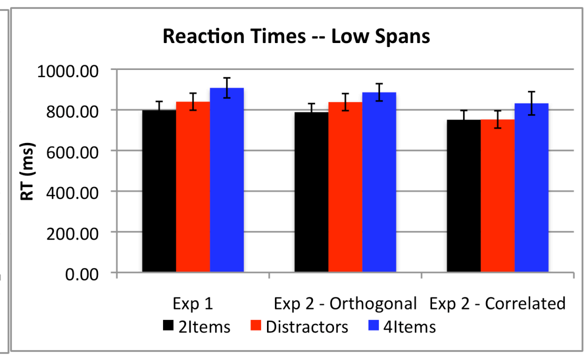


Figure 20. Mean RTs for each of the three trial types across the three experiments/conditions. Error bars represent the standard error of the mean.

However, these results should be interpreted cautiously, given that it is difficult to interpret the precise meaning of fast RTs. For example, fast RTs on the Distractors trials might suggest high filtering, particularly in the Orthogonal condition, where distractors interfere with performance. However, fast RTs on the Distractors trials might also be a result of *low* filtering, particularly in the Correlated condition, where "distractors" were actually task-relevant.

Specifically, attending to distractors in the Correlated condition might result in faster RTs, similar to what is often observed on congruent trials of the Flanker and other similar tasks (Baudouin et al, 2008; Rueda et al., 2004). Thus, interpreting the RTs on the Distractors trials is difficult, particularly in the Correlated condition, where the speedup could indicate high filtering (as a result of ignoring the distractors) *or* low filtering (as a result of the congruency effect).

Interim conclusions for Experiment 2

I predicted a strong positive relationship between WM and filtering in the Orthogonal (high filtering demand) condition, as in Experiment 1, and a lack of (or even a negative) relationship between WM and filtering in the Correlated (low filtering demand) condition. However, the ERP results from Experiment 2 did not support these predictions and instead demonstrated:

- Worse ERP-based filtering overall, and no relationship with WM span across both Orthogonal and Correlated conditions
- Smaller CDA amplitudes overall (i.e. smaller contralateral vs. ipsilateral difference), driven by larger ipsilateral amplitudes in Experiment 2
- Decreased separation between 4Items and 2Items CDA, for high spans
- Same behavioral performance (Acc and RTs) across the two experiments

The RT-based results more closely matched the predictions, such that there was a positive relationship between WM and filtering in the Orthogonal, but not in the Correlated condition. However, due to the potential congruency-based speedup in the Correlated condition, it is difficult to precisely interpret RT-based results. Nevertheless, the finding that the Orthogonal

RT-based results replicated those from Experiment 1 suggests that either (1) the difficulty with the ERP-based results lie with the ERPs themselves (e.g. due to technical problems), and not with the paradigm in general, or (2) subtle processes not detectable by less sensitive RT-based measures were responsible for the unexpected Experiment 2 results. Follow-up experiments, therefore, need to be conducted in order to definitively reconcile these two possibilities. In the meantime, follow-up statistical analyses were conducted in an attempt to better understand the underlying reasons for the unexpected Experiment 2 results.

Follow-up analyses exploring unexpected results of Experiment 2

Differences in experimental setup were examined to determine if these differences could have contributed to the Experiment 2 results, where there was no longer any relationship between ERP-based filtering and WM, contrary to the predictions. These analyses indicated less vigilant attention to the direction indicated by the arrow (which tells participants which side of screen to attend to for the memory array), more processing of the task-irrelevant side during the arrow presentation, and less overall attention to the relevant side of space in Experiment 2, particularly for the low WM participants. Therefore, less compliance with the stated instructions in Experiment 2 (i.e. to attend exclusively to the side indicated by the arrow) made it difficult to test for the predicted patterns in the relationship between WM and filtering task-irrelevant information.

*Changing distractors***.** The purpose of Experiment 2 was to manipulate filtering demand within a single paradigm, to test whether high WM can support dynamic filtering, based on task demands. For this purpose, the distractors (i.e. the blue items) changed (either orthogonally or consistently) with the targets (i.e. the red items).

The change in distractors might have lowered filtering overall in Experiment 2, because quickly changing distractors (across memory and test arrays, separated by only 900 ms) might make the distractors more obvious, making them more difficult to filter.

However, the changing distractors account cannot explain the lack of a sufficient distance between the 2Items and the 4Items conditions (necessary to obtain a reliable ERP index of WM (Vogel & Machizawa, 2004)), since these trial types do not contain any distractors. This account also cannot explain why there were lower CDA amplitudes overall in Experiment 2 relative to Experiment 1. Finally, this account predicts worse behavioral performance, at least for the Distractors condition, whereas behavioral performance was identical across the two experiments.

Increased duration of arrow. In Experiment 2, the duration of the arrow in the beginning of each trial that indicated where the participants needed to attend, was increased from 200 ms (followed by 300-400 ms fixation cross) to 700-1100 ms (and the arrow was followed immediately by the onset of the memory array) (see Figure 2). This was done for the purpose of attempting to reduce eye movements associated with the speeded presentation of arrow in Experiment 1. However, unexpected "side effects" from the increased duration may have contributed to the obtained results of Experiment 2.

Specifically, the increased arrow duration might have lowered overall amount of attention deployed to the direction indicated by the arrow (since participants had more time to process the arrow, their attention to it might have been less focused). Less attention to the direction of the arrow might in turn have increased processing of the irrelevant side (i.e. *ipsilateral* to channel location), at least early in the trial. Increased processing of the irrelevant side might result in lower overall levels of filtering, especially when the filtering measure is based on the differences between contralateral and ipsilateral processing, which are both affected the focus of attention.

To test this account, the following analyses were conducted:

1. Examination of early ERPs (e.g. P1 and N1), time-locked to the arrow onset, in order to test whether early attention to the direction indicated by the arrow differed across experiments, trial types and WM spans.
2. Examination of attentional components (e.g. N2pc), time-locked to the memory array, to test how much attention was deployed to the contralateral vs. ipsilateral hemifield during the trial, and whether there was an interaction with experiment, trial types, and WM span.

Early ERPs to arrow onset

To examine how much attention was deployed to the direction indicated by the arrow, the P1 and N1 components were examined (see Figure 21). The P1 and N1 are early ERP components (~75 to 150 ms post stimulus onset) that reflect initial perceptual processing and have been source-localized to the extrastriate cortical areas (e.g. Hillyard et al., 1998; Drew, McCollough, Horowitz & Vogel, 2009). These components are modulated by spatial attention (Fukuda & Vogel, 2009; Hillyard et al, 1998), and thus both the P1 and N1 amplitudes should be larger when more attention is deployed to processing the arrow. The peak-to-peak distance between the P1 and the N1 components (Bellis, Nicol, & Kraus, 2000) was used here as a measure of early spatial attention. A 3 (condition: Experiment 1, Orthogonal, Correlated) x 2

(WM: high, low) ANOVA revealed a main effect of condition, $F(2,81) = 22.3$ $p < .001$, such that the P1-N1 peak distances were greater in Experiment 1 ($M = 5.5$ μV) than in Experiment 2 ($M = 3.49$ μV in Orthogonal and $M = 3.48$ μV), both p's < .001, suggesting more attention was paid to the arrow in the first than in the second experiment. There was no difference between Orthogonal and Correlated conditions, $p = .97$. There was also no main effect of working memory, $F(1,81) = 1.3$, $p = .26$. There was no interaction between condition and WM, $F(2,81) = .42$, $p = .66$.

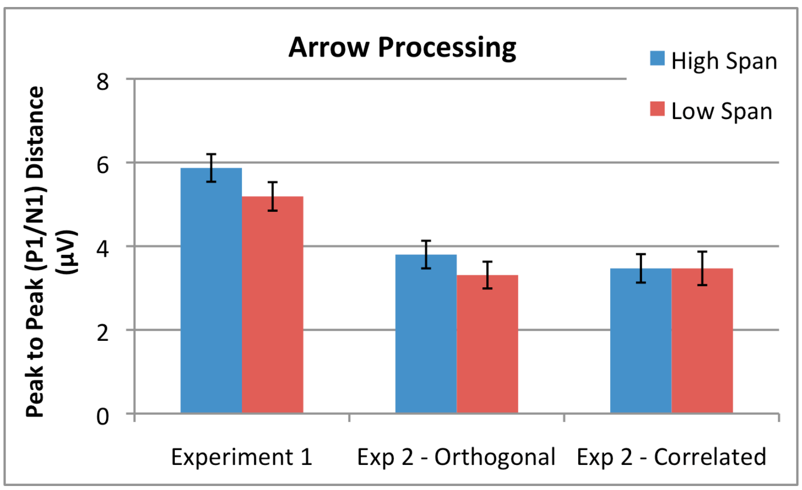


Figure 21. P1/N1 peak distances as a function of experiment and WM span suggest more attention to the arrow in Experiment 1 than in Experiment 2

To determine whether Experiment 2 elicited less attention to the arrow itself, or instead less attention to the *direction* indicated by the arrow, contralateral and ipsilateral P1 peaks were analyzed. The arrow was presented centrally, with only the arrowhead serving as the cue about the relevant direction. Thus, early ERPs should not be lateralized if there are no attention differences to the direction of the arrow. Instead, results from a 2 (laterality: contralateral vs. ipsilateral) by 3 (condition: Experiment 1, Experiment 2 Orthogonal, Experiment 3 Correlated) repeated-measures ANOVA with P1 peaks found a main effect of laterality, $F(1,79) = 9.1$, $p =$

.003, such that contralateral P1 amplitudes (M = 2.52 µV, SE = .11) were *smaller* than ipsilateral P1 amplitudes (M = 2.62 µV, SE = .11). Critically, this effect was tempered with an experiment by laterality interaction, $F(2,79) = 3.6$, $p = .03$ (see Figure 22), such that in Experiment 1, there is no evidence of early laterality (M for contralateral P1 = 3.94 µV, M for ipsilateral P1 = 3.92 µV), $F(1,29) = .14$, $p = .71$. In contrast, in Experiment 2, contralateral P1 amplitudes were *smaller* than ipsilateral P1 amplitudes in both Orthogonal (contralateral M = 1.84 µV, SE = .19; ipsilateral M = 1.99 µV, SE = .19; $F(1,27) = 7.1$, $p = .013$) and Correlated (contralateral M = 1.77 µV, SE = .21; ipsilateral M = 1.94, SE = .20; $F(1,23) = 9.98$, $p = .004$) conditions. These results show that attention was deployed *away* from the direction indicated by the arrow in Experiment 2. In addition, there was a main effect of condition, $F(2,79) = 39.4$, $p < .001$, such that Experiment 1 P1 amplitudes (M = 3.9, SE = .18) were larger than both Experiment 2 Orthogonal P1 amplitudes (M = 1.92, SE = .19) and Experiment 2 Correlated P1 amplitudes (M = 1.86, SE = .20). Adding WM to the ANOVA contributed no effects of WM on laterality of P1 amplitudes and no interactions with WM, all p's > .22, suggesting that WM was not modulating the observed effects. Thus, Experiment 2 elicited both less overall attention to the arrow (suggesting that participants were less vigilant and alert in processing this direction cue) and critically, less attention to the *task-relevant side* of the screen.

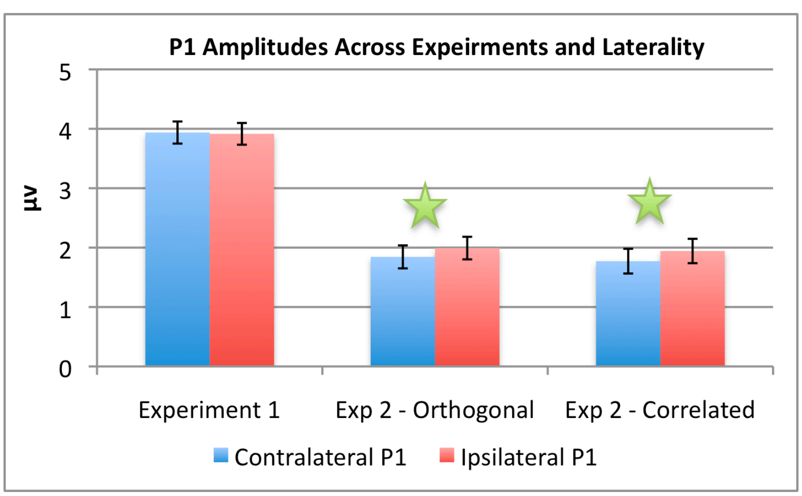


Figure 22. Experiment 2 elicited lower P1 amplitudes overall and lower contralateral, relative to ipsilateral amplitudes, suggesting attention was deployed to the irrelevant side of screen during the early part of the trial in Experiment 2. Stars indicate significant differences between Contralateral and Ipsilateral P1 activity in Experiment 2.

Laterality of Attention (N2pc)

As the P1/N1 results demonstrate, less attention was paid to the direction indicated by the arrow in Experiment 2 than in Experiment 1. To test the implications of the different amounts of attention across the two experiments, the N2pc attentional component was closely examined across the experiments.

The N2pc component (abbreviation for parietal contralateral negativity in the N2 time window) is an increase in negative voltage recorded over a contralateral hemisphere in response to a target (Eimer, 1996; Woodman & Luck, 1999). It differs from the CDA component in both its latency and more medial scalp topography (McCollough, Machizawa & Vogel, 2007). It is interpreted to index the amount of attention paid to the contralateral, relative to the ipsilateral side. Typically N2pc is observed in response to targets embedded with non-targets (distractors) that may need to be filtered in order to allow correct discrimination of the target; thus, it is absent when non-targets can easily rejected or when no filtering is required for the task (Eimer, 1996; Luck & Hillyard, 1994). Therefore, the N2pc component is expected to be largest in response to the Distractors trials.

Thus, to test whether participants paid more attention to the task-irrelevant (i.e. ipsilateral to channel location) side in Experiment 2 than in Experiment 1, first the overall N2pc amplitudes (contralateral minus ipsilateral activity) were examined across the 2 experiments. Results showed a main effect of trial type, $F(2,158) = 44.4$, $p < .001$, such N2pc contralateral minus ipsilateral differences were largest in the Distractors condition ($M = -.86$ µV), as predicted, intermediate in the 4Items condition ($M = -56$ µV), and smallest in the 2Items condition ($M = -.31$ µV), all p's for pair-wise contrasts less than .001[4]. There was also a trial type by condition interaction, $F(4,158) = 5.0$, $p = .001$, such that the differences in amplitudes across trial types were larger in Experiment 1 than in Experiment 2 (see Figure 23).

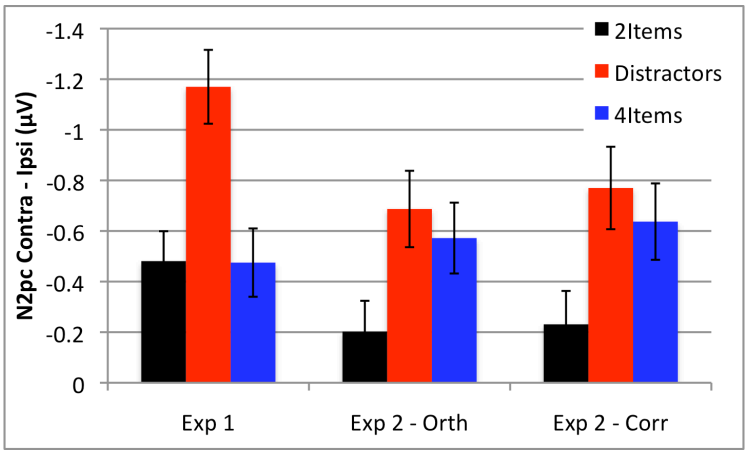


Figure 23. N2pc amplitudes (contralateral minus ipsilateral) across the three trial types in the two experiments

Next, I examined whether the decreased N2pc amplitudes in Experiment 2 were driven by decreased attention to the contralateral side, increased attention to the ipsilateral side, or both. Thus, I separately examined the amplitudes of the contralateral and ipsilateral N2pc amplitudes,

---

[4] The finding that the N2pc amplitudes were largest in the Distractors condition is consistent with the finding of N2pc being sensitive to the number of distractors (Eimer, 1996); thus, all subsequent analyses will focus on the N2pc component for Distractors trial type only, as it provides the most power and sensitivity in finding differences across conditions/experiments.

by experiment, and subsequently by WM span, for the Distractors trials, where the N2pc amplitudes were largest. A 2 (laterality: contralateral, ipsilateral) by 3 (condition: Experiment 1, Exp 2 – Orthogonal, Exp 2 – Correlated) ANOVA revealed a main effect of laterality, $F(1,76) = 78.3$, $p < .001$, such that contralateral N2pc ($M = -.63$ µV) was more negative than ipsilateral N2pc ($M = .16$ µV), $p < .001$ for the pair-wise contrast, suggesting that overall more attention was deployed to the "correct" (contralateral) side of space. There was no effect of WM, $F(1,76) = 1.2$, $p = .28$, no effect of condition, $F(2,76) = .58$, $p = .56$, and no interaction with WM, $F(2,76) = .23$, $p = .8$. However, there was an interaction between laterality and condition, $F(2,76) = 3.7$, $p = .03$ and a significant 3-way interaction between laterality, condition, and WM span, $F(2,76) = 4.0$, $p = .022$.

The 2-way interaction between laterality and condition suggested that attention, indexed by N2pc, was more lateralized in Experiment 1 than in Experiment 2 (with no difference for the Orthogonal and the Correlated conditions of Experiment 2) (see Figure 24).

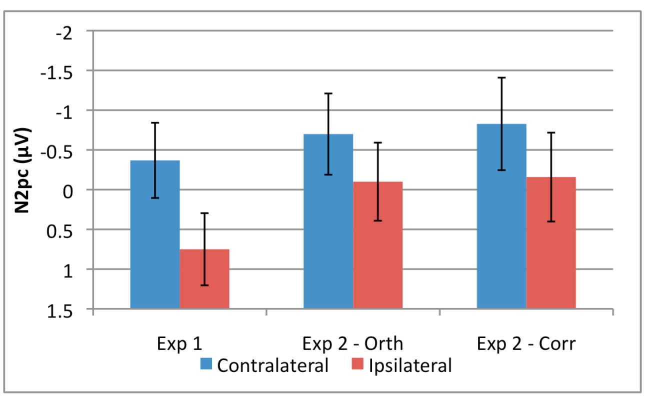


Figure 24. N2pc amplitudes were more lateralized (i.e. there were larger difference between contralateral and ipsilateral activity) in Experiment 1 than in Experiment 2.

Moreover, the 3-way interaction between laterality, experiment and WM span suggested that this decrease in laterality of attention in Experiment 2 (i.e. more attention to the task-irrelevant side) was driven by *low* WM participants, for whom the contralateral-ipsilateral differences were smallest in Experiment 2. (Figure 25).

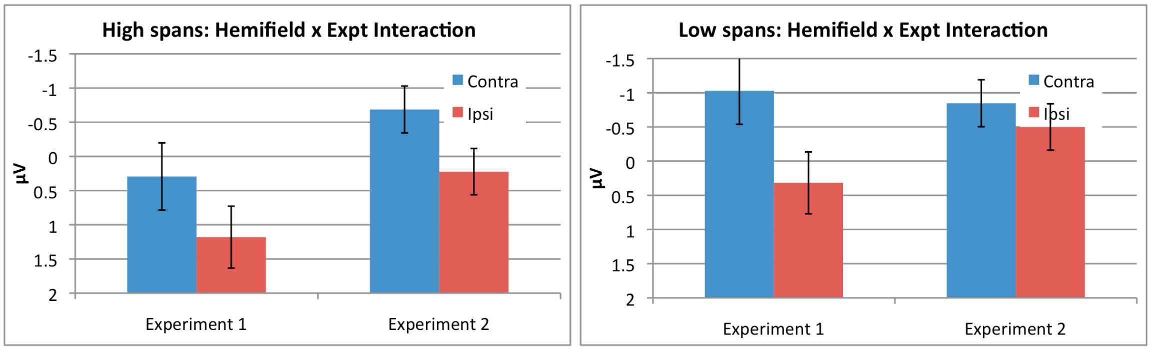


Figure 25. Attention, as indexed by the N2pc component was less lateralized in Experiment 2, particularly for low WM participants.

These follow-up results suggest that in Experiment 2, early on in the trial less attention was paid to the direction indicated by the arrow, and instead participants deployed more attention *away* from the direction indicated by the arrow (since contralateral P1 activity was smaller than ipsilateral activity). When the memory array appeared, participants were able to shift their attention to the task-relevant side of the screen (since contralateral N2pc activity was greater than ipsilateral activity). High WM participants might have been quicker in returning their attention back to the task-relevant side (consistent with this idea, high WM participants showed larger CDA amplitudes overall than low WM participants in Experiment 2), but this was not enough to compensate for the early attention to the irrelevant side. Attention was still less lateralized in Experiment 2, suggesting that this shift in attention was not complete. This was particularly true for low WM participants, who continued to maintain the widest attentional focus (since their ipsilateral N2pc activity was largest).

These follow-up analyses help to explain the reasons behind the unexpected ERP-based filtering results in Experiment 2. Attending away from the task-relevant side of screen and thus having a broader attentional scope could have contributed to lower ERP filtering results in Experiment 2 and lower CDA difference waves, which are based on the difference between contralateral and ipsilateral activities.

Eye movements issue

One important question is whether the unpredicted ERP results in Experiment 2 were the result of an increased rate of eye movements. Specifically, it is possible that the increased duration of the arrow in Experiment 2 made it less likely that participants stayed fixated in the middle of the screen and instead moved their eyes, e.g. in the direction of the arrow. Visually fixating on the relevant stimuli would trivialize the task; therefore, if participants were looking in the direction indicated by the arrow in Experiment 2 (instead of deploying their covert attention there while keeping the eyes centrally fixated), this could explain lower ERP-based measure of filtering, and less distance between voltages for each of the three trial types. However, as explained below, eye movements cannot explain the obtained results, because even the cleanest sample with the least amount of detectable eye movements showed the same pattern of results: a positive relationship between the behavioral and the ERP-based measure of WM, along with a positive relationship between WM and filtering in Experiment 1, and a lack of such relationships in Experiment 2.

Eye movements were quantified by extracting mean amplitudes from the eye channels for the later part of each trial (250-1000 ms after stimulus onset), and contralateral versus ipsilateral differences were calculated, like for the CDA analyses. Eye movements produce electrical activity in opposite polarities on opposite-side eye channels, and therefore, this measure should approach zero if no (or few) eye movements were produced and should have a large magnitude if eye movements are large and frequent. However, because these analyses need to be done on data that is averaged across all trials for any given participant (eye movements could not be detected on individual trial data), saccades to different directions may cancel each other out, producing a score of zero, and erroneously suggesting that no eye movements were made.

Thus, I cannot definitively address the question of how much participants were moving their eyes, given the need to average across trials, such that saccades in one direction cancel looks in the other direction, so a participant could move eyes inconsistently across trials and still appear to have no eye movements. However, the average for each participant at least informs of how much the participant tended to look toward one direction over the other (raw values of eye movement voltages), or the size of such tendencies to look in one direction (absolute values of these eye movement voltages). Critically, differences of interest did not seem to stem from differences in eye movements:

(1) The experiments differed only marginally in the direction participants tend to look, such that eye movements were marginally more likely to be in the direction indicated by the arrow for Experiment 2 (Orthogonal $M = .37$ µV, $SE = .14$; Correlated $M = .33$ µV, $SE = .16$) than in Experiment 1 ($M = -.03$ µV, $SE = .13$), $F(2,76) = .26$, $p = .08$. Moreover, the experiments did not differ in the size of such tendencies to look in one direction (absolute values of eye movements were the same across Experiment 1 ($M = .76$, $SE = .09$), Orthogonal condition ($M = .72$, $SE = .72$) or Correlated condition ($M = .71$, $SE = .11$): $F(2,76) = .07$, $p = .93$.

(2) Focusing on participants with the smallest tendencies to look toward one direction over the other (eye movements not exceeding +/- 1.5 µV), the differences of interest between

experiments were fully replicated. In Experiment 1, there was still a significant relationship between the behavioral measure of WM (K) and the ERP measure of WM based on CDA differences of the 4Items and the 2Items trials, $R = -.58$, $p = .003$, $N = 23$. There was also a significant positive relationship between the behavioral measure of WM and the ERP measure of filtering (outlier filtering values were removed as in previous analyses), as in Vogel et al. (2005), $R = .47$, $p = .05$, $N = 18$. However, in Experiment 2 there was no longer any relationship between behavioral WM measure and the ERP WM measure, $R = .038$, $p = .82$, $N = 38$, and no relationship between the behavioral WM measure and the ERP measure of filtering, $R = .065$, $p = .73$, $N = 31$. As in the analyses reported above, there was a trend for high spans (high/low determined using the consistent criterion across experiments, $M = 2.0$) to filter *less* ($M = .50$, $SD = .84$, $N = 24$) than low spans ($M = .85$, $SD = .51$, $N = 8$), $t(29) = 1.0$, $p = .28$.

Therefore, eye movement did not significantly contribute to the obtained results in Experiment 2, given that the more scrutinous analysis of eye movements produced the same results: meaningful (and predicted) relationships with WM for Experiment 1 and no such relationships with WM for Experiment 2.

General Discussion

The purpose of this experiment was to manipulate filtering demand within a single paradigm in order to test the theory that high WM can support dynamic updating of filtering strategy, based on the current task demands. High- and low-filtering-demand versions of the previously high-filtering-demand task were created for this experiment. However, the results were unexpected such that there was no relationship between WM and the ERP measure of filtering for either the high- or the low-filtering-demand conditions.

Follow-up analyses revealed that the ERP-based measures were not informative in this experiment, most likely due to the way in which this experiment was designed. Specifically, the duration of the arrow indicating the direction participants need to attend to for the memory and test arrays was increased in Experiment 2 for the purpose of trying to alleviate the eye movement problem that was apparent in Experiment 1. The eye movement problem was not alleviated (24% of participants still needed to be excluded), but this increase in duration of arrow produced additional problems: post-hoc analyses demonstrated that in Experiment 2 participants paid less attention to the direction indicated by the arrow during the arrow display and paid less attention to the task-relevant side of the screen during the memory array, resulting in worse differentiation between trial types and worse overall filtering, for both high and low WM participants.

It is possible that the longer duration of the arrow in Experiment 2 resulted in behaviors similar to those observed with Inhibition Of Return (IOR) paradigms (Klein, 2000; McDonald, Hickey, Green, & Whitman 2009; Samuel & Kat, 2003). In IOR paradigms, a peripheral cue signals a target at the same location, but the duration of the cue is varied. With short durations (around 200 ms), target processing in the cued location is facilitated and marked by shorter RTs. However, longer durations (around 700 ms) elicit *longer* RTs to the target at the cued location. This increase in RTs is interpreted as difficulty in returning attention back to the location from which it has already wandered (Klein, 2000). In Experiment 2, the longer duration of the arrow might have encouraged attention wandering to the task-irrelevant side, as evidenced by greater ipsilateral than contralateral early ERPs (P1 peaks) in Experiment 2, unlike in Experiment 1. Deploying attention even briefly to the task-irrelevant side of the screen might have made it more difficult to bring attention back to the correct side for remembering the memory array. Thus, participants who were slow in bringing attention back to the irrelevant side were “forced” to have to remember both sides of the visual display, with a maximum of eight items to remember. The slowdown (but eventual success) in bringing attention back to the task-relevant side of space can explain both the high behavioral performance and the meaningful pattern of behavioral results (which might be less sensitive to quick covert shifts of attention), and the confusing ERP results (which are more sensitive to small changes in timing of behavior; e.g. Besson, Kutas & Van Petten, 1992; Rugg & Coles, 1995).

However, several aspects of the design of the study and the obtained results suggest that the underlying reasons for the unexpected results in Experiment 2 were not based on IOR, at least in the strict sense of the term. First, there was no slowdown in responses in Experiment 2 relative to Experiment 1, which is a hallmark manifestation of IOR-based behaviors. If anything, RTs were numerically faster in the second experiment. However, IOR-like shifts of attention very early in the trial (during arrow presentation) might not have affected the less sensitive behavioral markers of performance, while strongly affecting the more sensitive ERPs. Second, IOR paradigms typically involve peripheral cues, and thus, it is unclear whether the central arrow would elicit IOR in Experiment 2. Central cues have also been used to elicit IOR behavior (i.e. the slowdown in response to a longer cue), but only in situations where eye movements were allowed (Klein, 2000). However, as follow-up eye movement analyses demonstrate, the Experiment 2 pattern of results remains even when the cleanest sample, in terms of objectively measured eye movements, is obtained. Thus, eye movements were unlikely to cause IOR behaviors in Experiment 2.

More broadly, it is possible that the more relaxed pace of Experiment 2 (created by the longer duration of the arrow) eliminated WM-mediated attentional differences that enabled high spans to both attend more to the arrow and to filter highly in the fast-paced Experiment 1, such that WM-mediated attention was less necessary for processing information in Experiment 2. However, both high and low spans' early attention to arrow (measured by the P1/N1 amplitudes to arrow) decreased in the second experiment, suggesting that another factor contributed to producing unexpected results in Experiment 2. Thus, a combination of the IOR-like attention wandering to the irrelevant side during the arrow presentation, and the less vigilant processing of the arrow (which might be WM-dependent) could have produced the unexpected and difficult to interpret Experiment 2 results, including low CDA amplitudes without much differentiation across trial types, for both high and low WM participants, which in turn produced difficult-to-interpret ERP indices of WM and filtering.

These results point to the fragility of this otherwise seemingly robust paradigm (Brignani, Bortoletto, Miniussi, & Maioli, 2010; Emrich, Al-Aidroos, Pratt, & Ferber, 2009; Fukuda, Awh, & Vogel, 2010; Fukuda & Vogel, 2009; Ikkai, McCollough, & Vogel, 2010; Wang, Most, & Hoffman, 2009). Very slight changes in the setup of the trial, which do not even involve changes to the memory-related aspects of the trial affected the results so drastically as to eliminate this paradigm's usefulness in capturing ERP-based measures of WM and filtering. The fragility of this paradigm has now been demonstrated by others. Sander, Werkle-Bergner, & Lindenberger (2010) found that changing the duration of the memory array from 100 to 500 ms eliminated the existence of the CDA component altogether. Murray, Kuo, Stokes, & Nobre (2009) found that the CDA disappeared if an additional cue was presented after the memory array, regardless of whether this cue was neutral, or whether it retroactively indicated which subset of items were task-relevant. Thus, future work should both strive to design experiments as closely as possible to the existing work in order to obtain the effects of interest, but should also explore and define the limits of this paradigm. Understanding why paradigms are often fragile is useful both practically in terms of maximizing the probability of obtaining interesting and interpretable results from the experiment, but additionally, and perhaps more importantly, for deepening the understanding and delineating the boundaries of conditions that elicit the psychological processes in question. The fragility of this paradigm (Murray et al., 2009; Sander et al., 2010) also indicates a strong need to examine the relationship between WM and filtering in different filtering contexts, which may be not be quite as sensitive to small changes in task setup.

Nonetheless, the RT-based results from Experiment 2 are encouraging and somewhat consistent with the predictions, possibly because RT-based measures might be less sensitive (Luck, Woodman, & Vogel, 2000), and thus less affected by small changes in the timing of the task. The RT-based findings show a positive relationship between WM and filtering in the Orthogonal condition, and no relationship in the Correlated condition, as predicted. However, there are at least two reasons why these RT-based results should be interpreted with caution. First, it is the low spans who appear to be changing strategies across the two conditions, not the high spans, which is contrary to the idea that high WM can support dynamic allocation of filtering resources, based on task demands. Second, results from the Correlated condition are difficult to interpret, given that low filtering (i.e. processing of the distractors) can either slow down processing (as in the Orthogonal condition), or instead can speed up processing (as is typically found in the congruent trials of the Flanker task). The formula used to calculate filtering efficiency assumes that greater distractor processing will slow down RTs (since it assumes low filtering equates RTs for 4Items and Distractors trial types); nevertheless, it is not clear whether this assumption is justified, given the potential for the congruent distractors speedup.

Despite these shortcomings, Experiment 2 results provide additional evidence to the growing body of literature suggesting that WM plays a pivotal role a number of processes seemingly unrelated to WM. In Experiment 1, high WM participants showed enhanced P1/N1 amplitudes, when ERPs were time-locked to the arrow onset, suggesting more enhanced early perceptual processing of the arrow and better spatial attention to the arrow. These ERP components have been localized to the extrastriate cortical areas in the occipital cortex, far away from the prefrontal regions thought to support WM and other executive functions. Nevertheless, these results are consistent with existing findings of high WM being associated with larger early perceptual ERPs, such as the N1 component (Brumback, Low, Gratton, & Fabiani, 2004), which were interpreted to indicate stronger attentional modulation in high WM participants. In addition, high WM participants were more likely to maintain attention to the task-relevant side of the screen after the memory array presentation, as suggested by the N2pc results, which have been source localized also to the extrastriate cortical areas, and perhaps to the posterior parietal regions (Hopf, Luck, Girelli, Hagner, Mangun, Scheich, & Heinze, 2000). Thus, WM predicts very early perceptual and attentional processes that are not commonly thought about in terms of higher-order executive functions (Hillyard et al., 1998).

In addition, preliminary analyses of ERPs time-locked to the test-array (see Figure 1A) indicated that high WM was also associated with greater amplitudes for both the N1 component and also for what appears to constitute the parietal old/new effect (Rugg & Curran, 2007). These results suggest that WM may also be associated with providing top-down control for successful memory retrieval, a finding consistent with Elward & Wilding (2010), in addition to the attentional processes that help during encoding and maintenance. In fact, the view of WM presented here, in terms of the ability to bias processing in favor of maintaining task-relevant information across interference and delays, is perfectly consistent with these somewhat surprising findings. Flexible allocation of top-down control can provide substantial support for early attention to a very brief arrow in Experiment 1 and for maintaining attention in the relevant side of space in Experiment 2. This view of WM has also been proposed to account for much of variance in tasks conventionally thought to tap nothing much beyond simple processing speed (Cepeda, Blackwell, Munakata, in prep.).

EXPERIMENT 3: EXPLORATION OF DYNAMIC FILTEIRNG IN CHILDREN AND ADULTS

Experiment 3 was designed to extend the investigation of the relationship between WM and filtering task-irrelevant information by exploring whether the shift in filtering strategy, predicted among high WM participants, can occur *dynamically* in the course of performing a task. The filtering task used in Experiments 1 and 2 (based on Vogel et al., 2005) did not allow for such testing, because the demand for filtering was either consistently high or consistently low throughout the entire task. Moreover, Experiment 2 demonstrated the fragility of the previously used paradigm (based on Vogel et al., 2005), thus highlighting the need to investigate the relationship between WM and filtering in other filtering contexts. Therefore, Experiment 3 introduced two additional paradigms in which the relationship between WM and filtering was investigated within the course of performing the task, and within participants. In addition, this relationship was explored developmentally, in six-year-old children. Exploring this relationship in development is important, because as discussed in the Introduction, developmental explorations may help to better identify the dynamics supporting the underlying mechanisms for higher-level cognitive functions, as highlighted by possible dissociations in children' and adult's performance.

Two new filtering paradigms were utilized in this experiment: the *Flanker* task, developed first by Eriksen and Ericksen (1974) and modified more recently (Fan, McCandliss, Sommer, Raz, & Posner, 2002; Rueda, Fan, McCandliss, Halparin, Gruber, Lercani & Posner, 2004), and the modified version of the *Garner* speeded classification paradigm (based on the recent adaptation by Baudouin, Durand, & Gallay, 2008; Garner, 1974).

In the Flanker paradigm, participants need to respond to the direction (left or right) of central arrow (or a picture of a fish, in the child-appropriate version). The central item may be flanked on both sides by other items (arrows or fish), which may point either in the same direction as the central item (congruent trials), or in the opposite direction (incongruent trials). The central item may not be flanked at all on some trials, or may be flanked by items that contain no direction information, such as horizontal lines (neutral trials). The slowdown on the incongruent trials, relative to neutral trials, is referred to as the incongruency cost. The speedup on the congruent trials, relative to neutral trials, is referred to as the congruency benefit.

There are large individual differences in both children and adult abilities to ignore the irrelevant flanking information (Gonzalez, Fluentes, Carranza, & Estevez, 2001; Fosella et al., 2001; Rueda et al, 2005). Moreover, these individual differences appear to be related to differences in executive functions, which share much variance with WM, and thus make this task an appropriate candidate for exploring the relationship between WM and filtering task-irrelevant information. For example, 4-6 year old children with the homozygous long allele of the dopamine transporter type 1 (DAT1) gene, linked to superior performance on a number of executive functions (e.g. Bertolino et al., 2006), showed a smaller incongruency cost (calculated there as the difference between incongruent and congruent RTs) on the Flanker task (Rueda et al., 2005). A similar result was found when exploring the relationship between DAT1 alleles and Flanker task performance in adults (Fossella et al., 2002), thus providing further support to the suitability of the Flanker task in exploring the relationship between WM and filtering task-irrelevant (flanking) information.

In the Garner task, participants need to make speeded judgments regarding features of one dimension (e.g. emotion: happy or sad), while ignoring features of another dimension (e.g.

face identity: person A or person B). In the *baseline* block, the irrelevant dimension does not vary. In the *correlated* condition, the relevant and the irrelevant features vary consistently (e.g. Person A happy; Person B sad), thus lowering the demand for filtering task-irrelevant (identity) information, as it is completely redundant with the task-relevant information (emotion). In contrast, in the *orthogonal* condition, the two dimensions vary inconsistently (e.g. Person A happy or sad; Person B happy or sad), thus increasing the need to filter task-irrelevant information in order to maximize accuracy and speed of responses on the task-relevant information. The incongruency cost on the Garner task is measured in terms of slowdown on the Orthogonal block, relative to the Baseline block, and the congruency benefit is measured in terms of the speedup on the Correlated block, relative to Baseline.

Both of these tasks should provide information about the specific filtering profiles used by participants. Low static filtering can be manifested by a large incongruency cost and a large congruency speedup, which would result from taking in all (or most) of the available information. High static filtering can be manifested by a relatively small incongruency cost and a relatively small congruency benefit, since attending to only the task-relevant information should minimize any effects of task-irrelevant features. Finally, the dynamic filtering profile can be inferred from a relatively small incongruency cost (produced by high filtering in the incongruent, or orthogonal situations), and a relatively large congruency benefit (produced by low filtering in the congruent, or correlated situations). Directly comparing the magnitudes of congruency and incongruency effects might be problematic, given possible differences in difficulty levels of filtering strongly vs. weakly; thus, the relative strengths of the relationships between incongruency and congruency costs will be examined.

Finally, the Flanker task should provide an additional test of whether filtering can be modulated dynamically, based on the changing task demands. It is well established that in paradigms where incongruent and congruent trials are interleaved, there are significant carryover effects from one trial to the next (Egner, 2007; Gratton, Coles & Donchin, 1992; Jha & Kiyonaga, 2010; Nieuwenhui, Stins, Posthuma, Polderman, Boomsma, & de Geus, 2006). Specifically, incongruent trials are performed faster than when they follow incongruent trials (II trials) relative to when they follow congruent trials (CI trials), and conversely, congruent trials are performed faster when they follow congruent trials (CC) than when they follow incongruent trials (IC). This *conflict adaptation* effect (i.e. faster RTs on II trials relative to CI trials; also known as the Gratton effect) is typically explained in terms of carryover effects of upregulated cognitive control on the incongruent trials, based on the conflict signal sent from the anterior cingulate cortex (ACC) and implemented by the lateral prefrontal cortex (PFC) (Botvinick, Braver, Barch, Carter, & Cohen, 2001; Egner, 2007; but see description of other explanations in the Discussion section). The reduction in congruency effect following incongruent trials (i.e. slower RTs for IC trials relative to CC trials) is typically explained in terms of reduced facilitation from congruent distractors, because an increase in cognitive control (during incongruent trials) can produce stronger attentional biasing exclusively to the task-relevant features, thus reducing the congruency effect (Egner, 2007). If high WM can support dynamic, flexible allocation of cognitive control in order to give rise to either high or low filtering, high WM individuals might show a larger conflict adaptation effect, such that they can strongly recruit cognitive control on incongruent trials (and show strong benefits on the subsequent incongruent trials), and strong release from cognitive control on the congruent trials (and show strong benefits on the subsequent congruent trials). Low WM participants are thus expected to exhibit relatively little conflict adaptation. To the best of my knowledge, there has not been any published work examining the relationship between WM and conflict adaptation.

Methods

*Participant*s

Adults: Seventy-one right-handed University of Colorado undergraduate students (44 female) participated in this two-session experiment. EEG was recorded during the first session, and results from that session are described in the Experiment 2 chapter. The second session was administered 4-10 days after the first session and involved several behavioral measures, which are described and analyzed here, in the context of Experiment 3. Eight participants failed to return for the second session (i.e. Experiment 3); therefore, the final sample for adult participants in Experiment 3 includes 63 participants. Only 25 participants completed the perceptual priming task, because this task was decided to be added after the data collection has begun, and not all participants had time to complete another task before the end of the session.

Children: Forty-four six-year-old children (*M* = 6 years, 3 months; range 6.1—6.7 years; 23 female) participated in this experiment. Seven participants failed to come back for the second session, which included the Complex Span task, leaving 37 children in the final sample for that task.

Materials and Procedure

Tasks were designed to be maximally similar for adult and child participants to maximize the ability to compare results across the two age groups. Changes were made to accommodate children's inability to complete several long tasks in a row without losing focus and motivation. The order the tasks for the two groups were also somewhat similar. Adult tasks were ordered as follows: Garner task, Complex span task, Flanker task, Forward and Backward digit spans, Luck and Vogel task, and the Perceptual Priming task. Child tasks were ordered as follows: The first session included the Garner task, Luck & Vogel task, Flanker Task, Forward and Backward digit spans and the Perceptual Priming task. The second session for children included the Complex Span task. The Complex Span task was moved to the second session for children for the purpose of keeping each session under one hour long, given that children are unable to complete long sessions. The Luck & Vogel task was moved earlier in the session for children because the Complex span task was moved to the second session, and it was deemed best to avoid having two filtering tasks in a row; thus, the Luck and Vogel task was inserted between the two filtering measures. Instructions were purposefully vague with respect to filtering, in order to see if high spans spontaneously adopt a dynamic filtering strategy, based on task demands, in accords with the predictions for this experiments.

Garner Filtering task. This task was modeled after the task used in Baudouin et al (2008), which was adapted from the Garner speeded classification paradigm (Garner, 1974), in which participants needed to quickly respond to one dimension of a presented item while ignoring the second dimension. Female faces with two different emotions (happy or sad) and two different identities were centrally presented on the screen as in Baudouin et al. (2008). Participants were instructed to respond based on the emotion of the picture (i.e. press one button for a happy face and another button for a sad face). Nothing was said about the irrelevant identity dimension. The face identity was never the task-relevant dimension because Baudouin et al. (2008) found that effects were larger when the emotion dimension was task-relevant in adults and comparable in children.

Critically, there were three different conditions, similar to classic Garner paradigms (see Figure 26). In the *Baseline* condition, only the relevant dimension (emotion) was varied while the irrelevant dimension (identity) did not vary (i.e. there were two possible types of stimuli: Person A happy, Person A sad, or Person B happy, Person B sad; counterbalanced across participants). In the *Correlated* condition, both dimensions were varied, but in a confounding manner, such that only two types of stimuli were presented: Person A Happy and Person B Sad, or Person A sad and Person B happy; counterbalanced across participants). Thus, in the correlated condition, the irrelevant dimension (face identity) was 100% predictive of trial outcome, and therefore a wide attentional focus (i.e. low filtering) was task-advantageous. Finally, in the *Orthogonal* condition, the relevant and the irrelevant dimensions were fully crossed (Person A Happy; Person A Sad; Person B Happy; Person B Sad), such that attending to the irrelevant dimension (face identity) was disadvantageous, thus encouraging high filtering strategy.

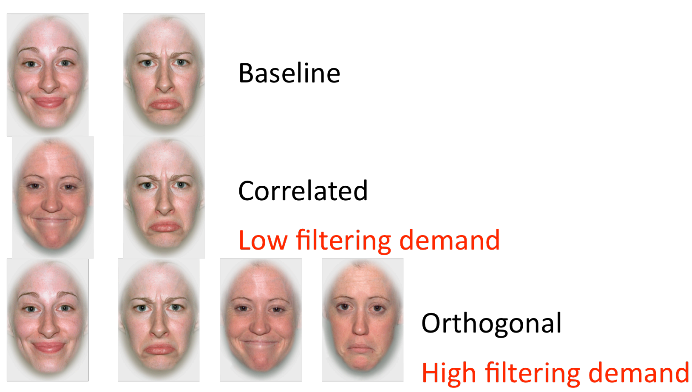


Figure 26. Examples of possible stimuli in the Baseline, Correlated, and Orthogonal blocks in the Garner task.

In order to decrease task-related variability and increase power in the individual differences approach, the order of blocks was fixed as: Baseline, Correlated, Orthogonal. The only difference between the adult and child version of the tasks was in terms of the number of trials per block: 40 trials per block in the child version and 60 trials per block in the adult version.

Complex Span task. This task was adapted from Barrouillet, Gavens, Vergauwe, Gaillard, & Camos (2009) and was intended to measure verbal WM capacity. In this task, participants were required to remember increasingly longer strings of animal names, while concurrently naming the colors of smiley faces. Each animal appeared on the screen for 2000 ms, followed by a string of four smiley faces, which were on the screen for 1333 ms each, with a 667 delay between faces. Adult participants were asked to silently think of the name of the animal, while for child participants, the name of the animal was announced by the experimenter. Everyone needed report the color of the smiley face (red, blue, or yellow) out loud as they appeared. Participants were then prompted to report the entire string of animals in each trial by a question mark, centrally located on the screen. Animal strings started with only one animal per trial and increased in length until they reached the maximum of five animals for adults, and four animals for children. The original study terminated the task when participants failed to recall four series at a particular level ("truncated span measure" from Miyake & Friedman (2005); however, the current experiment opted for the more sensitive "proportion correct" measure (Friedman & Miyake, 2005), and thus administered all items for all the participants.

Flanker filtering task. The child version of the Flanker task was adapted from the child Attentional Network Task (ANT) (Rueda, Posner, Rothbart, & Davis-Stober, 2004; Rueda et al., 2004; Rueda, Rothbart, McCandliss, Saccomanno & Posner, 2005). In the ANT Flanker task, children were asked to respond as quickly as possible to the direction (left or right) of the centrally-located image on the screen, which could be presented by itself (neutral trials, Figure 27A), be flanked by congruent images (congruent trials, Figure 27B), or be flanked by incongruent images (incongruent trials, Figure 27C). In the child version of the Flanker task, the trials were blocked, such that the first and third blocks were incongruent and contained only incongruent and neutral trials. The second block was congruent and contained only congruent and neutral trials. The blocked design was utilized to maximize the possibility of some children dynamically changing the filtering strategies across incongruent blocks (on which high filtering was task-advantageous) and the congruent block (on which low filtering was task-advantageous). Each block contained 32 trials (16 neutral trials and 16 Flanker trials). The trial structure was identical to that in Rueda et al (2004): a variable 400-1600 ms fixation, followed by 150 ms of double cues[5], followed by another 450 ms long fixation, which was followed by the target items (neutral, incongruent, or congruent), which lasted no longer than 2500 ms to make the task reasonably fast-paced for young children. Auditory feedback was provided on every trial to maintain children's attention and interest in continue doing the task. Finally, there was a 1000 ms long inter-trial interval.

(A)  (B)  (C) 

[5] Only double cues were used (asterisks right above and below the central fixation cross) because Rueda et al. (2004) found the largest congruency effect in this condition. Congruency effects are not always present, and therefore, it was important to maximize the probability of obtaining the congruency effect, in order to test the prediction that high WM can support dynamic updating in filtering strategy

Figure 27. Stimuli used in the child version of the Flanker task on neutral (A), congruent (B), and incongruent trials (C).

In the adult version of the Flanker task, the stimuli and methods were identical to those used in Fan et al. (2002). There were also three types of trials: congruent (Figure 28A), incongruent (Figure 28B) and neutral (Figure 28C); which were interleaved. Participants completed 24 practice trials (on which they were given feedback), followed by three blocks of 96 trials each. Each trial started by a 400-1600 ms fixation cross, followed by a 100 ms cue, followed another 400 ms fixation cross, which was followed by the target (incongruent, congruent, or neutral), which was presented on screen for no longer than 1700 ms to make the task appropriately speeded. No feedback was provided in the adult version of the Flanker task on the experimental trials.

(A) (B) (C)

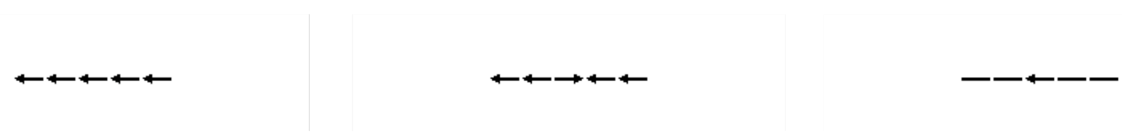


Figure 28. Stimuli used in the adult version of the Flanker task in congruent trials (A), incongruent trials (B), and neutral trials (C).

Forward and Backward digit spans. The digit span subscales from the WISC-R scale were used in order to maximize similarity between the child and the adult studies. The task was computerized; pairs of strings of digits of increasing length were presented both aurally and visually at the rate of one digit per second. At the end of the string, participants were cued to repeat strings of digits (in the forward and backward order, respectively) by a question mark centrally position on the screen. The task terminated after a participant failed to correctly report both strings of a given length. There was no difference in procedures for the child and the adult versions of these tasks.

Luck & Vogel task. A shortened version of the task used in Luck & Vogel (1997) was administered to both children and adults in order to provide an additional measure of visual working memory capacity. In this task, participants needed to make same/different judgments regarding the color of the squares that appeared on the screen in the memory and test arrays (see Figure 29). Squares were identical across memory and test arrays on 50% of trials, and one square changed colors on the remaining 50% of trials. The size of the stimuli and the visual angle was the same as in Luck & Vogel (1997). For adults, two blocks (set size 4 and set size 8) were administered, with 20 trials in each. WM capacity was estimated based on performance of the larger (8-item) set, as recommended in Luck & Vogel (1997). Participants viewed arrays of squares of different colors presented briefly during the memory array (100 ms) and after a brief delay (900 ms) were asked to make same/different judgments about the color of squares presented in the test array. For children, the task was modeled after Riggs et al. (2007), and included four blocks (set sizes 2, 3, and 4), with 12 trials for each set. The duration of the memory array was 500 ms for children, followed by a 900 ms delay interval, and ending with a test array that stayed on screen for 3000 ms, or until the child responded. The addition, the block with the set size of five items was added halfway through the experiment, in order to make the task slightly more difficult; however, because not all children received this manipulation, the WM capacity for children is estimated using performance on the set size of four items.

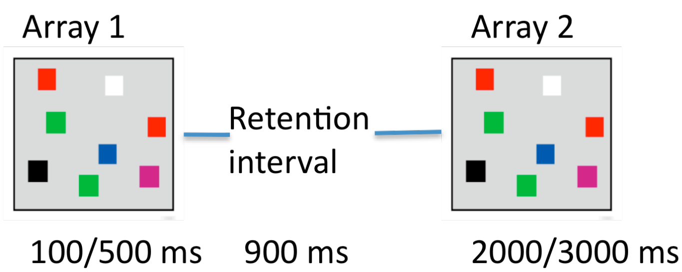


Figure 29. Example of a trial on the behavioral version of the Luck & Vogel task. The duration of Array 1 was 100 ms for adults and 500 ms for children. The maximum duration of Array 2 was 2000 ms for adults and 3000 ms for children. The maximum number of items to remember was eight for adults (as shown) and five for children.

Perceptual Priming task. Participants first saw a series of 10 complete (non-fragmented) line drawings of common objects (selected from Cycowicz, Friedman, Snodrass & Rothstein, 2000) on a computer screen, presented for two seconds each. Participants were instructed to watch the pictures carefully without saying anything (to minimize explicit processing/active rehearsal). Participants were then asked to guess the names of each picture. They were shown 10 new and 10 old pictures, fragmented as in Cycowicz et al. (2000). Pictures were presented in a random, but constant order, each starting with most fragmented and becoming progressively less fragmented, until the participant identified the picture, or until the full, non-fragmented picture was displayed. In the child version of the task, the experimenter asked whether the child knew the name of the picture and if so, asked the child to name the picture. The adult version was self-administered; i.e. adult were told to press “1” if they knew the name of the picture (and was then asked to write down its name on a scoring sheet) and to press “0” if they needed more information to guess the name of the picture. Trials answered incorrectly were removed from analyses. “Old” and “new” lists were counterbalanced, and constructed such that each list contained the same number of items from various categories, such as animals, household items, etc. The perceptual priming score was calculated as the difference between the level at which *old* and *new* pictures were recognized.

Data Trimming

Reaction time data trimming was identical to Experiment 1 and 2, with two exceptions. In the Garner task, I eliminated that exceeded *two* standard deviations for the mean of each block for each participant, to be consistent with procedures used in Baudouin et al., 2008. In addition, only minimal trimming was used for investigating the size of the conflict adaptation effect in the Flanker task. The order of the trials should be preserved as closely as possible when examining sequential effects; thus, the only trials that were removed were the incorrect trials and the trials that directly followed the incorrect trials, because it was not clear how to categorize the latter trials, in terms of whether they followed congruent, incongruent, or neutral trials.

Results

As elaborated below, Experiment 3 results provided additional evidence suggesting that high WM can support dynamic adjustment of top-down control, to allow either high or low filtering of task-irrelevant information, based on task demands, in both adults and six-year-old children. The filtering tasks used in this experiment enabled investigating whether filtering demand can be *dynamically* adjusted, within the course of performing a task. The addition of the developmental population enabled examination of the developmental origins of the relationship between WM and filtering.

The dynamic filtering account was tested by examining (1) changes in accuracy rates across different blocks and trial types, (2) changes in RTs across high and low filtering demand trial types (Flanker) and blocks (Garner), relative to baseline RTs, and (3) the size of sequential carryover effects following congruent vs. incongruent trials (Flanker task; Garner task was not suitable for this given its fully blocked design). The overall pattern of results on both Flanker and Garner tasks, with respect to filtering task-irrelevant information, is presented in Table 7. This table is color-coded to indicate which pieces of data were compatible with the dynamic filtering account (dark grey), which pieces were incompatible (white) with this account, and which pieces of data do not (or cannot) inform this account (light grey). Several pieces of data were consistent with the dynamic filtering account; dynamic changes were observed across both shorter time scales (e.g. sequential trial effects) and longer time scales (e.g. changes across blocks, and even across trials). However, challenges to the dynamic filtering account were also observed, particularly in terms of changes in magnitude of incongruency costs in the child Flanker task, and overall not very large effects (several were only marginal or almost marginal, if using the standard cutoff of p = .05). Moreover, as elaborated below, several pieces of data could be consistent with both static and dynamic filtering profiles, and therefore, neither strongly support, nor contradict the dynamic filtering account.

	Flanker task		**Garner task**	
	Kids	**Adults**	**Kids**	**Adults**
Accuracy changes across blocks	High WM: Neutral trials: Block 1 < Block 2 > Block 3 Low WM: constant acc	No relationship with WM	High WM: Base < Corr = Orth Low WM: Base = Corr = Orth	High WM: Base = Corr = Orth Low WM: Base = Corr < Orth
Incongruency cost - RT	High WM: smaller cost Low WM: larger cost	High WM: larger cost Low WM: larger cost	No relationship with WM for standard Incongruency cost; High WM showed smaller Orthogonal-Correlated difference	No relationship with WM
	High WM: no correlation between 2 Incongruent blocks Low WM: postive correlation	Analysis cannot be done because no intervening Correlated block		
Congruency benefit - RT	No relationship with WM	High WM: larger benefit Low WM: smaller benefit	No relationship with WM	No relationship with WM
Conflict adaptation	Effect present, but not modulated by WM	High WM: larger conflict adaptation effect	Cannot assess, design fully blocked	Cannot assess, design fully blocked

Table 7. Summary of Experiment 3 results. The dark grey color highlights data that are compatible with the dynamic filtering account. White regions show data that are incompatible with the dynamic filtering account. Light grey regions show data that neither support nor contradict this account.

In addition, the results from this experiment provided further support to the idea that WM supports many seemingly unrelated processes. Specifically, high WM was associated with faster performance across both Garner and Flanker tasks, in the adult participants. In children, there was no such WM-associated speedup, and even a trend in the Garner task for slower processing among high span children. In addition, in children there was a *negative* correlation between WM and performance on the perceptual priming task, which is thought to tap more posterior cortical regions, perhaps suggesting an early dissociation between prefrontal and posterior cortical regions.

Working Memory Tasks

In both children and adults, performance on the Complex Span task did not correlate with other working memory measures, possibly because this task may have relied more on retrieval-based than maintenance-based processes, because names of animals needed to be retrieved after naming strings of four colors; also, adult performed at ceiling, with a very limited range (the most frequent score was 96.7% correct). In children, the remaining WM measures (Forward and Backward digit spans, and the Luck & Vogel task) were correlated (all p's $< .15$; Figure 30) and aggregated, using the average of z-scores for each measure. In adults, the ERP K measure of WM was used for maximal comparability with Experiment 2, although the measures other than Complex span also mostly correlated (all p's $<.09$, except for lack of relationship between Forward span and Luck & Vogel task, $R = 0$, $p =$ n.s.; Figure 31), and the same patterns were obtained with an aggregate measure.

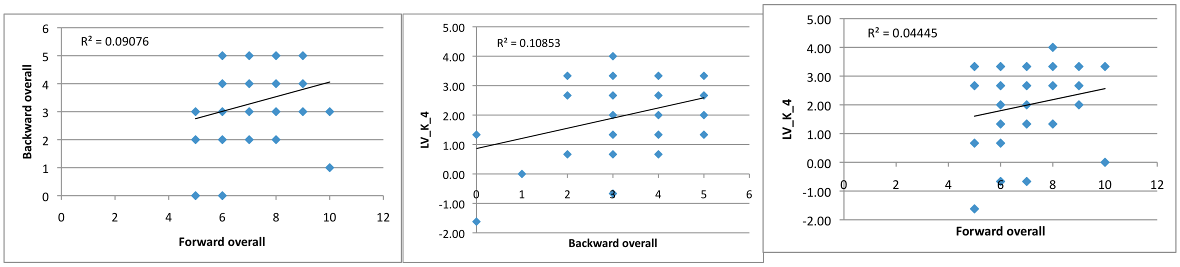


Figure 30. Correlations between performance on the Forward digit span, Backward digit span, and Luck & Vogel tasks for children.

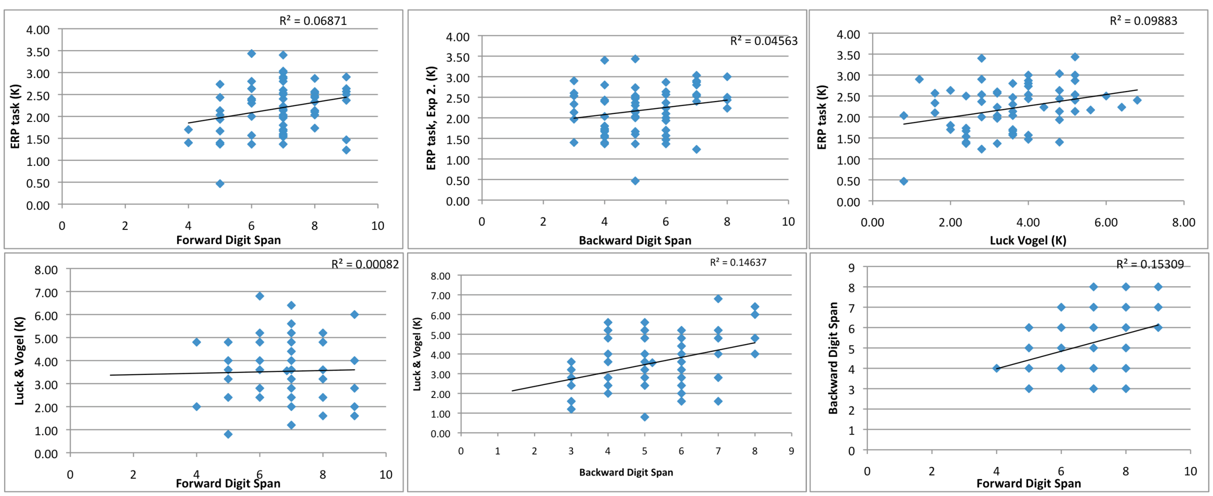


Figure 31. Correlations between performance on the Forward digit span, Backward digit span, Luck & Vogel, and the ERP-based WM tasks for adults.

Flanker Filtering Task

Children:

Accuracy:

As elaborated below, the accuracy-based results of the neutral trials in the children's Flanker task were consistent with the dynamic filtering strategy, such that high spans were dynamically adjusting their performance based on task context, whereas low spans were remarkably stable in their performance.

Children performed well, with the overall accuracy of 94%. Incongruent trials (Block 1 and Block 3) were performed less accurately (M = 92.1%) than congruent trials (Block 2) (M = 95.5%), $F(1,37) = 6.7$, $p = .014$. There was no effect of WM ($F(1,37) =.08$, $p = .77$), nor a WM by trial type interaction, $F(1,37) =2.1$, $p = .16$. Incongruent trials were performed equally accurately across the two incongruent blocks (M for Block 1 = 92.0% correct; M for Block 3 = 92.2% correct), $F(1,37) =.03$, $p = .86$. There was no effect of WM, nor an interaction with WM on incongruent trials accuracy, all p's > .4.

Neutral trials were affected by the context in which they were embedded (incongruent or congruent), but only for the high spans, $F(2,74) = 3.3$, $p = .044$ (see Figure 32). For low spans, there was no effect of block for the neutral trials accuracy, $F(2,38) = .25$, $p = .79$. In contrast, for the high spans, neutral trials embedded in the context of the incongruent trials were answered less accurately (M = 94.2% for Block 1; M = 91.6% for Block 3) than the neutral trials in congruent block (M = 97.2%), $F(1,18) = 8.3$, $p = .01$ for the quadratic contrast. Post-hoc LSD tests revealed that high spans became marginally more accurate on Block 2 (Congruent), relative to Block 1 (Incongruent), $p = .094$ and then significantly dropped their accuracy by Block 3 (Incongruent), $p = .013$. Performance on the neutral trials in the two incongruent blocks (Block 1 versus Block 3) did not differ, $p = .26$. This finding suggests that high spans might dynamically adjust their performance across blocks, such that they could take advantage of congruency in the context of the Congruent block and could be hurt by the context of incongruency in the Incongruent block, even on the neutral trials, whereas low spans' performance was consistent across blocks.

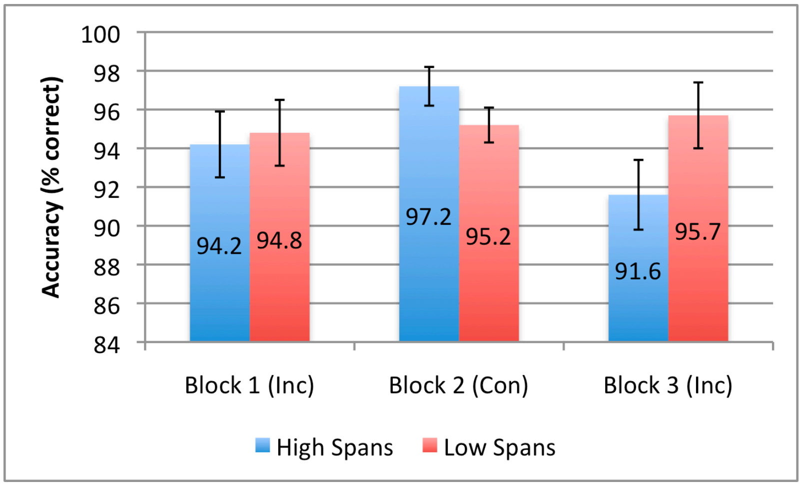


Figure 32. Accuracy on the neutral trials across the three blocks, for both high and low WM participants. Low spans showed a consistent profile across the three blocks, whereas high spans' neutral trial performance depended on the amount of incongruency: accuracy increased in the congruent, relative to the incongruent blocks.

Low WM participants were consistent not only at the group level (reported above), but also at the level of individual differences, such that they showed significant correlations between accuracy on neutral trials across all block comparisons, whereas high spans did not (with only a marginal correlation between block 1 and block 2 performance, and no significant correlations for any other comparisons) (see Table 8). These findings are particularly striking given that low WM is typically associated with worse consistency and reliability across measures (Hultsch, MacDonald, & Dixon, 2002; Long & Prat, 2002; Stuss, Murphy, Binns, & Alexander, 2003). Correlation for high and low WM participants were not significantly different from each other (all p's > .2) using the Fisher's Z transformation, but Fisher's Z is a notoriously conservative test (e.g. Berkson, 1978).

	Block 1 x Block 2	**Block 1 x Block 3**	**Block 2 x Block 3**
Low Spans (N = 21)	**R = .61**, β= .88 p = .003	**R = .58**, β = .61 p = .008	**R = .52**, β = .38 p = .02
High Spans (N = 20)	**R = .40**, β = 1.0 p = .083	R = .32, β = .30 p = .18	R = .14, β = .12 p = .54

Table 8. Low spans were more consistent that high spans in their performance on the Flanker task, as shown here by their significant relationships (bolded) between accuracy on the neutral trials across the three types of blocks.

Reaction Times:

As described in Table 7 and elaborated below, the RT-based results from the child Flanker task provided both support and challenges to the dynamic filtering account. The child Flanker task enabled analysis of the incongruency costs, congruency benefits, changes in incongruency costs after encountering the congruent block (Block 2), and the size of the conflict adaptation effect.

There was no relationship between WM and RTs (raw or log-transformed) on any of the trial types, even after the Standardized DfFit outlier removal procedure, all p's > .2.

The Incongruency cost was calculated as the difference between RTs on the Incongruent and Neutral trials in the Incongruent blocks (Blocks 1 and 3). The congruency benefit was calculated as the difference between RTs on the Neutral and the Congruent trials in the Congruent block (Block 2). There was no relationship between WM and the Congruency "benefit" (R = .04, p = .80, N = 41). However, there was a marginal negative relationship between WM and the Incongruency cost (R = -.27, p = .095, N = 39; see Table 9). These results suggest that low spans demonstrated the low filtering profile, because their incongruency cost was marginally larger than high spans'. However, from these analyses it is not clear whether high spans demonstrated high static filtering, such that they were filtering irrelevant information throughout; or, if they were instead using a more dynamic filtering strategy, such that they filtered strongly in the Incongruent blocks (resulting in the marginally smaller incongruency cost) and filtered weakly in the Congruent block (resulting in the comparable congruency benefit to that for the low spans, who are filtering weakly). The dynamic filtering account predicts no relationship between WM and the congruency cost (since both low filtering and dynamic filtering profiles are expected to produce relatively large congruency costs); however, caution should be used in interpreting this null effect as strong evidence in favor of the dynamic filtering account.

	Incongruent	**Congruent**	**Neutral**	**Incongruent – Neutral**	**Neutral – Congruent**
Full sample	R = .064, p = .69, N = 42	R = .07, p= .66, N = 42	R = .08, p = .61, N = 42	**R = -.29**, p = .07, N = 42	R = .04, p =.81, N = 42
After outlier removal	R = .078, p = .64, N = 38	R = .15, p = .36, N = 39	R = .15, p = .38, N = 39	**R = -.27**, p = .095, N = 39	R = .04, p = .80, N = 41

Table 9. Correlations between WM and trial types Log RTs on the child Flanker task, both before and after outlier detection process. Relationships with p-values < .1 are bolded. High spans show a marginally smaller incongruency cost.

To further elucidate the relationship between WM and filtering in the Flanker task, the relationship between the two Incongruent blocks (Block 1 and Block 3) was examined as a function of WM capacity. In the model where the Incongruency cost for Block 3 was predicted from the Incongruency cost from Block 1, WM, and the interaction between the two factors, there was a marginally significant interaction term, $t(1,41) = -1.8$, $p = .086$, such that the slopes for the high and low WM participants were marginally different (see Figure 33). For the high spans, there was no relationship between Incongruency costs in Block 1 and Block 3, $\beta = -.083$, $R = .08$, $N = 21$, $p = .74$. In contrast, for the low spans, the relationship was positive, $\beta = .85$, $R = .43$, $N = 21$, $p = .05$, such that the Incongruency cost in Block 1 strongly predicted the Incongruency cost in Block 3. The Fisher's Z-transform showed that these correlations were marginally different, $p =.10$. These findings suggest that the intervening Congruent block (Block 2) affected high spans' performance, but did not have any effect on that of the low spans. This interpretation is further corroborated by the finding that WM predicts incongruency costs almost marginally better in the first Incongruent block (Block 1: $R = .31$, $\beta = -.018$, $p = .045$, $N = 41$) than in the second Incongruent block (Block 3: $R = .054$, $\beta = -.005$, $p = .76$, $N = 36$), Fisher Z correlation comparison, $p = .12$. However, it is also challenged by the finding that the Incongruency cost did not go up more for high spans (who are thought to dynamically change their strategy from high filtering in Block 1 to low filtering in Block 3, after encountering the congruent Block 2) (M for Block 1 = .046; M for Block 3 = .064, $p = .27$ for the pair-wise comparison) than for the low spans (M = for Block 1 = .06; M for Block 3 = .084; $p = .08$ for the pairwise comparison), $F(1,40) = .12$, $p = .74$.

In summary, the basic block RTs results provide preliminary support to the idea that low WM could be associated with low filtering under certain task demands. The comparison of incongruency costs across the two Incongruent blocks suggest that only high spans may have adjusted their filtering strategy upon encountering the Congruent block, consistent with the idea that high WM can allow for dynamic adjustment of cognitive control, to support both high and low filtering of task-irrelevant information, based on task demands. However, the relatively weak (marginal) effects, along with an unsupported prediction regarding the change in magnitude of the Incongruency costs temper the strengths of these conclusions.

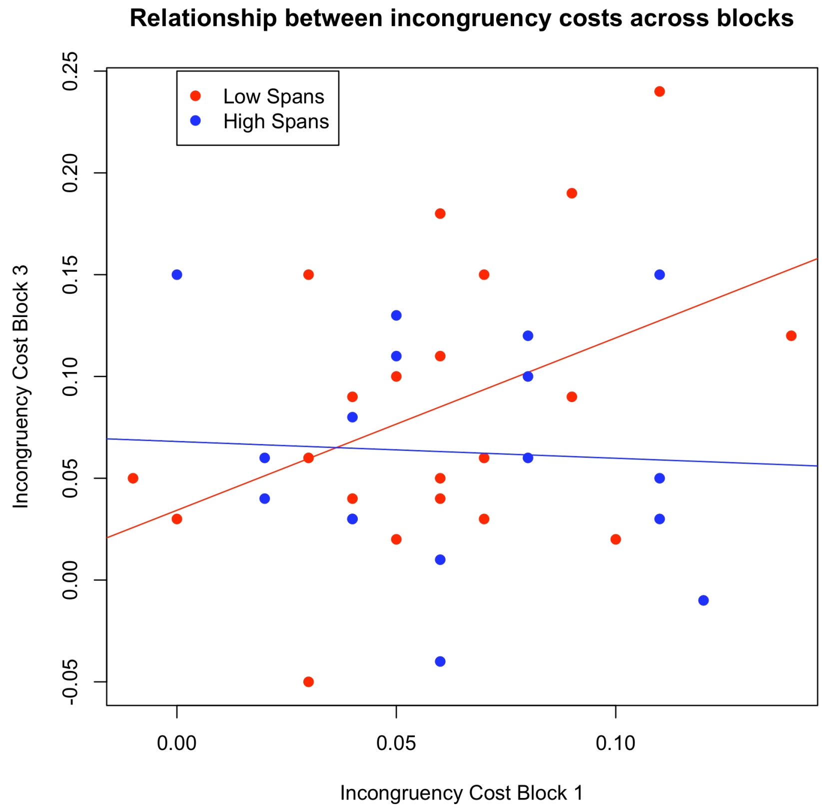


Figure 33. Low spans are consistent in the incongruency costs they show across Block 1 and Block 3, whereas High spans show no relationship, suggesting that high spans adjust their filtering strategy in Block 3 after encountering congruency (low filtering demand) in Block 2.

The child version of the Flanker task was not ideally suitable for testing for the conflict adaptation effect, given that the trials were blocked, such that the first and the third blocks contained only incongruent and neutral trials, and the second block contained only congruent and neutral trials. In addition, there were fewer trials in the child version of the Flanker task than in the adult version; thus, sequential effects might be less reliable and more difficult to observe. Nevertheless, trials were coded in terms of their sequential order. Of particular interest were incongruent trials that followed incongruent trials (II) and incongruent trials that followed neutral trials (NI). Despite the blocked design, the conflict adaptation effect was observed, such that II trials (*M* = 3.05, *SE* = .015) were faster than NI trials (*M* = 3.07, *SE* = .013), $F(1,40) = 4.7$, $p = .04$. This effect did not differ across the two blocks (Block 1 and 3), $F(1,40) = .002$, $p = .97$. This effect appeared to be related to the presence of conflict, rather than simply presence of flankers, because in the Congruent block, there was no difference in RTs for congruent trials following congruent trials (CC) (*M* = 3.01), congruent trials following neutral trials (NC) (*M* = 3.01), or neutral trials following congruent trials (CN) (*M* = 3.00), all *p*'s > .42. All of these trials took longer to respond to than neutral trials following neutral trials (NN) (*M* = 2.98), all *p*'s < .04, but this finding is consistent the finding that in children, congruent trials had a *cost* associated with them, relative to neutral trials. Critically, there were no interactions with WM capacity, all *p*'s > .23. Thus, the conflict adaptation effect, although present in children, was not modulated by WM. This could have occurred because children are less able to flexibly allocate cognitive control based on task demands, or because the blocked setup of the child Flanker paradigm was less suited for the subtle sequential effects.

Adults:

Accuracy:

Adults also performed well on their version of the Flanker task, with the average accuracy of 98%. Accuracy showed a marginal linear decrease as a function of block (*M* for Block 1 = 97.9%, *M* for Block 2 = 97.4%, and *M* for Block 3 = 97.4%), $F(1,61) = 3.4$, $p = .07$. There was no effect of WM and no interaction with WM, both *p*'s > .5. Thus, due to the lack of the relationship with WM, accuracy-based results from the adult Flanker task neither informed, nor contradicted the dynamic filtering account.

Reaction Times:

As described in Table 7 and elaborated below, the RT-based results from the adult Flanker task provided some support to the dynamic filtering account, in terms of the congruency benefit analysis and the conflict adaptation effect. The adult Flanker task did not allow for testing of the effects of intervening congruent block (as in the child version), because all trial types (incongruent, congruent, and neutral) were fully interleaved across the three blocks.

There was a negative relationship between WM and RTs on the individual trial types, such that high span adults were faster on all three types of trials (incongruent, congruent, neutral) (see Table 10).

	Incongruent	Congruent	Neutral	Incongruent – Neutral	Neutral – Congruent
Full sample	R = -.17, p = .18, N = 63	**R = -.24**, p= .05, N = 63	**R = -.21**, p = .099, N = 63	R = .085, p = .51, N = 63	**R = .24**, p =.055, N = 63
After outlier removal	**R = -.24**, p = .07, N = 58	**R = -.45**, p < .001, N = 58	**R = -.41**, p = .001, N = 60	**R = .32**, p = .013, N = 60	**R = .31**, p = .017, N = 59

Table 10. Correlations between WM and trial type Log RTs on the adult Flanker task, both before and after outlier detection process. Relationships with p-values < .1 are bolded. High spans are faster on all trial types, but show both a greater incongruency cost and a greater incongruency benefit.

As in the child version of this task, the incongruency cost was calculated as the difference between Incongruent and Neutral RTs. The congruency benefit was calculated as the difference between Neutral and Congruent RTs. In contrast to the child version of this task, however, there was a positive relationship between WM and both incongruency and congruency costs (see Table 2), both *p*’s < .02. This finding suggests that high WM was associated with both a *larger* incongruency cost and a larger congruency benefit. In fact, low spans showed a small but reliable congruency *cost* (M = -7.7 ms, SD 14.5), $t(28) = -2.9$, $p = .008$ when compared to the difference of 0 ms, suggesting that they were slowed down by all flankers, whereas high spans were able to extract and use the congruency information. These findings of the high incongruency cost and high congruency benefit suggest a *low* filtering profile for high span adults, relative to low spans. These results show that high WM is not always associated with a high filtering mode, as several previous studies (e.g. Vogel et al., 2005) have suggested.

Because all three trial types were interleaved with one another in the adult version of the Flanker task, this task was well suited to test for effects of WM on conflict adaptation. For this purpose, each trial was coded in terms of its position in the sequence of trials; of particular interest were incongruent trials that followed incongruent trials (II) and incongruent trials that followed congruent trials (CI). There was a strong conflict adaptation effect, such that II RTs (M = 594 ms or 2.76 when log-transformed) were shorter than CI RTs (M = 615 ms, or 2.78 when log-transformed), $F(1,61) = 13.4$, $p = .001$, or $F(1,61) = 15.7$, $p < .001$ when RTs are log-transformed. (Non-transformed RTs are shown to more clearly illustrate the scale for this effect).

If WM is associated with flexible, dynamic adjustment of filtering strategy, high WM participants might show a larger conflict adaptation effect, because they might be better able to up- and down-regulate cognitive control. Consistent with this idea, WM was associated with greater contrast between II and CI RTs. Specifically, when controlling for CI RTs, high WM was associated with *shorter* II RTs (so, a larger conflict adaptation effect), $t(53) = -2.5$, $p = .014$. Similarly, controlling for II RTs, high WM was associated with marginally *longer* CI RTs, $t(54) = 1.7$, $p =.092$, consistent with the idea that high WM predicts greater contrast between CI and II trials. Regressing WM on the difference between CI and II RTs yields a similar pattern, but did not reach significance: $R = .20$, $p = .12$ (Figure 34). After undergoing the Standardized DfFit procedure, the relationship weakened, but direction of the trend remained, $R = .14$, $p = .28$. However, the multiple regression approach is more powerful (Judd, McClelland, & Ryan, 2009) than a simple regression with difference scores, which are typically low in reliability (e.g. Lord, 1963). Thus, these results suggest that WM is positively associated with a larger conflict adaptation effect, consistent with the dynamic filtering account.

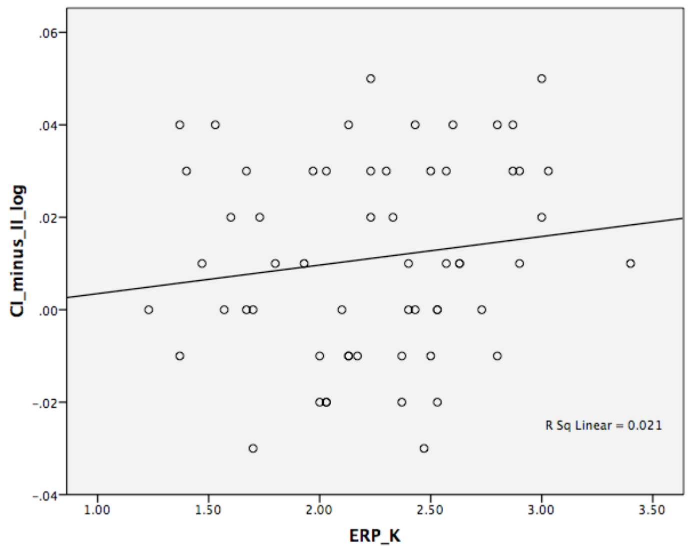


Figure 34. WM (K) is associated with a larger difference between CI and II trial RTs. The relationship is significant (p = .014) if the more powerful multiple regression method is used.

Garner Filtering Task

Children

The accuracy-based results in the child Garner task provided some support for the dynamic filtering account. The RT-based results neither supported nor contradicted the account, due to the overall lack of relationships with WM capacity. Conflict adaptation could not be investigated due to the fully blocked setup of the Garner task.

Accuracy:

Children performed well, with an average accuracy of 92% correct. There was no overall difference in accuracy as a function of block (Baseline, Orthogonal, or Correlated), $F(2,80) = 1.6$, $p = .21$. There was no effect of WM on accuracy, $F(1,40) = .39$, $p = .54$. However, there was a marginal interaction between block accuracy and WM, $F(2,80) = 2.4$, $p = .096$, such that for low spans, there was no difference in accuracy across the Baseline, Correlated, and Orthogonal blocks ($M = 92.7\%$, $M = 92.\%$, $M = 91.8\%$, respectively), $F(2,38) = .28$, $p = .76$. In contrast, for high spans, the effect of block was significant, $F(2,42) = 3.7$, $p = .03$, such that accuracy increased from the Baseline block ($M = 88.7\%$) to the Correlated block ($M = 92.5\%$), $p = .02$, and stayed constant in the Orthogonal block ($M = 91.5.\%$), $p = .37$. These findings are similar to those from the child Flanker task, where low spans were very consistent in their performance, while high spans shifted their behavior to become more accurate in the blocks with a low demand for filtering (Correlated block in the Garner task and Congruent block in the Flanker task). These data provide additional support to the dynamic filtering theory, given that low spans' accuracy is remarkable stable, whereas high spans' accuracy shifted as a function of filtering demand, across the Baseline and Correlated blocks.

Reaction Times:

Overall, RTs were comparable in the Baseline block ($M = 926$ ms, $SE = 38.7$) and the Correlated block ($M = 892$ ms, $SE = 28.5$), $p = .24$, and slower in the Orthogonal block ($M = 1082$ ms, $SE = 37.2$), both p's $< .001$. Thus, overall there was a significant incongruency cost, because Orthogonal RTs were longer than Baseline RTs, but no congruency benefit because Correlated RTs did not differ from Baseline RTs. Overall, RTs were very similar to the previous study (Baudouin et al., 2008), performed with slightly older children (6-8 year olds), although both the incongruency cost and the congruency benefit were found in the previous study (see Figure 35).

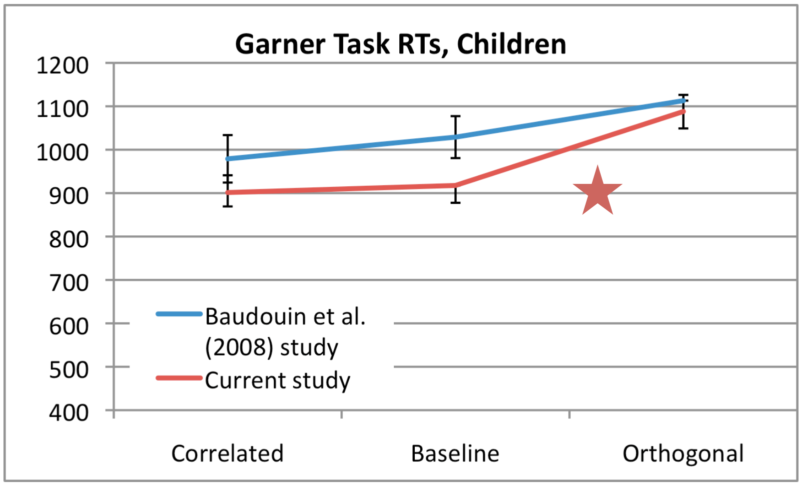


Figure 35. Group-level RTs for the child Garner task were comparable to those obtained in Baudouin et al. (2008), except Baseline RTs were significantly faster in the current study.

To examine the relationship between WM and potential filtering on the Garner task, regressions were conducted between WM capacity and both individual block RTs, along with differences across block RTs (see Table 11). High spans were faster than low spans on the Orthogonal RTs, R = -.32, p = .05, N = 37. However, there was no relationship between WM and Baseline RTs, R = -.07, p = .65, N = 39. There was a trend for high spans to be faster on the Correlated RTs, R = -.22, p = .2, N = 37. These results suggest that high span children may be particularly successful at filtering out task-irrelevant identity information on the Orthogonal block. However, when considering the relationship to Baseline performance, there was no relationship between WM and either the incongruency cost (p = .33) or the congruency benefit (p

=.54). Nonetheless, there was a significant negative relationship between WM and the difference between Orthogonal and Correlated RTs, $R = -.33$, $p = .04$, $N = 38$. However, this relationship is difficult to interpret: Orthogonal RTs are expected to decrease with WM (if high spans filter more, dynamically or statically), but it is not clear what should happen with Correlated RTs; they could increase with span if high spans continue to filter more, or not vary with WM (or even decrease with WM) if high spans dynamically update to filter less, like the low spans. Thus, the results from the child version of the Garner task are somewhat ambiguous with regard to the relationship between WM and filtering strategy.

	Baseline	**Correlated**	**Orthogonal**	**Orthogonal – Baseline**	**Baseline – Correlated**	**Orthogonal – Correlated**
Full sample	R = -.08, p = .60 N = 43	R = -.11, p = .50, N = 42	R = -.23, p = .14, p = 42	R = -.18, p = .26, N = 42	R = -.06, p = .72, N = 42	R = -.20, p = .20, N = 42
After outlier removal	R = -.07, p = .65, N = 39	R = -.22, p =.2, N = 37	**R = -.32**, p = .05, N = 37	R = -16, p = .33, N = 40	R = -.10, p = .54, N = 40	**R = -.33**, p = .04, N = 38

Table 11. Correlations between WM and trial type Raw RTs on the child Garner task, both before and after outlier detection process. Relationships with p-values < .1 are bolded. High spans were faster than low spans on the Orthogonal trial only, and have a smaller Orthogonal-Correlated difference.

Adults

Both the accuracy-based, and the RT-based results in the adult Garner task were ambiguous with respect to the dynamic filtering account, due to the overall lack of relationships with WM capacity. Conflict adaptation could not be investigated due to the fully blocked setup of the Garner task. However, the Garner task enabled direct comparison of performance across children and adults, due the identical task setup in the two age groups. The overall pattern of RTs across the three blocks was very similar for children and adults. However, in adults, high WM was associated with faster performance, whereas in children, the trend was for the opposite pattern.

Accuracy:

Adults performed well on the Garner task, with an overall accuracy of 97.5%. There was a significant effect of block, $F(2,124) = 5.6$, $p = .005$, such that accuracy was comparable in the Baseline block ($M = 97.5\%$, $SE = .003$) and the subsequent Correlated block ($M = 98.0\%$, $SE = .003$), $p = .14$, but decreased in the Orthogonal block ($M = 96.9\%$), both relative to the Baseline (marginally; $p = .06$) and relative to the Correlated block ($p = .003$). There was no effect of WM, $F(1,62) = .14$, $p = .71$. However, there was an almost marginal interaction between WM and block, $F(2,124) = 2.2$, $p = .11$, such that for high spans, there was no difference in accuracy across blocks ($M = 97.1\%$ for Baseline, $M = 97.8\%$ for Correlated, and $M = 97.2\%$ for Orthogonal), $F(2,62) = 1.2$, $p = .31$. In contrast, for the low spans, the effect of block was significant, $F(2,62) = 8.1$, $p = .001$, such that accuracy was constant in the Baseline ($M = 97.9\%$, $SE = .004$) and Correlated ($M = 98.2\%$, $SE = .005$) blocks, $p = .59$, but decreased in the Orthogonal block ($M = 96.6\%$, $SE = .005$), both p's $< .004$. This finding suggests that low spans adults were particularly hurt by situations where filtering task-irrelevant information was required for optimal performance (Orthogonal block). However, this finding is ambiguous with respect to the question of whether high WM can support dynamic filtering, based on task demands, since there was no relationship between accuracy and WM for high WM adults, possibly due to ceiling effects.

Reaction times:

There was a main effect of block, $F(2,124) = 41.6$, $p < .001$, such that RTs were marginally faster in the Correlated block (M = 497 ms, SE = 12.3) than in the Baseline block (M = 513 ms, SE = 15.8), p = .05, and significantly slower in the Orthogonal block (M = 584 ms, SE = 13.2) than in the preceding two blocks, both p's < .001. Thus, overall there were both a significant incongruency cost, and a significant congruency benefit. Overall, the pattern of RTs was very similar to the previous study, which also found incongruency costs and congruency benefits (Baudouin et al., 2008), but RTs were faster overall in the current study (see Figure 36). High spans were faster across all blocks (M = 504 ms, SE = 17.7) than low spans (M = 558 ms, SE = 17.7), $F(1,62) = 4.6$, p = .035. There was no interaction between WM and block, $F(2,124)$ = .82, p = .44, so both high and low spans were comparable in the extent of slowdown on the Orthogonal block and speedup on the Correlated block, despite high spans being faster overall.

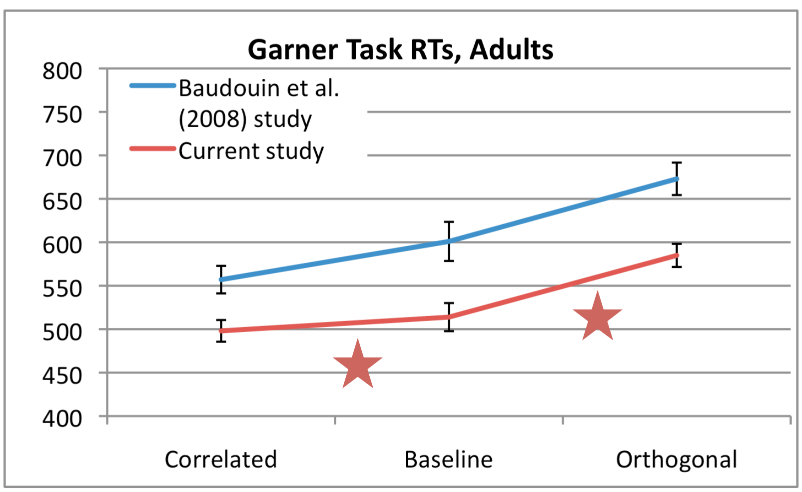


Figure 36. The pattern of group-level RTs for the adult Garner task was similar to that obtained in Baudouin et al. (2008); however, adults in the current study were faster overall.

To further examine the relationship between WM and potential filtering on the Garner task, regressions were conducted between WM capacity and both individual block RTs and differences across block RTs (see Table 12). Results showed that high WM was associated with faster performance across all three types of blocks (as in the adult Flanker task), but not related to either congruency benefits or incongruency costs, all *p*'s > .20. Adults' more proactive processing (Braver et al., 2001; Chatham, Frank, & Munakata, 2009) may have contributed to faster overall RTs for high spans on all three blocks, swamping possible differences across block RTs, and leading to difficulty in interpreting the relationship between WM and filtering in adults on the Garner task. However, it is unclear whey adults' proactive processing would not similarly obscure filtering effects in the Flanker task, where a similar WM-associated speedup was observed.

	Baseline	**Correlated**	**Orthogonal**	**Orthogonal – Baseline**	**Baseline – Correlated**	**Orthogonal – Correlated**
Full sample	**R = -.29**, p = .02 N = 64	**R = -.24**, p = .05, N = 64	R = -.20, p = .11, p = 64	R = -.16, p = .20, N = 64	R = -.19, p = .12, N = 64	R = .04, p = .75, N = 64
After outlier removal	**R = -.30**, p = .02, N = 60	**R = -.33**, p =.009, N = 61	**R = -.45**, p < .001, N = 59	R = 0, p = 1, N = 62	R = -.07, p = .59, N = 60	R = -.084, p = .51, N = 61

Table 12. Correlations between WM and trial type Raw RTs on the adult Garner task, both before and after outlier detection process. Relationships with p-values < .1 are bolded. High spans were faster than low spans on all trial types, but there was no effect of WM on the block difference RTs.

Comparing children and adults

The Garner task was best suited to compare performance as a function of development, because the setup of this task was identical for children and adults (with the exception of 40-trial blocks in children and 60-trial block in adults, to accommodate children's inability to complete very long tasks). Thus, RTs for each of the three types of block could be directly compared for children and adults, to investigate performance on this task as a function of age.

A 2 (population: children vs. adults) x 2 (WM: high vs. low) x 3 (block: Baseline, Correlated, Orthogonal) ANOVA reflected a main effect of block, $F(2, 202) = 65.8$, $p < .001$, such that RTs sped up from Baseline ($M = 730$ ms) to Correlated blocks ($M = 701$ ms), $p = .033$, and slowed down in the Orthogonal block ($M = 840$ ms), both p's < .001. This effect was tempered by an interaction with population age, $F(2,202) = 10.1$, $p < .001$, such that block differences were more pronounced in children than in adults (see Figure 37). The effect of block was not related to WM capacity, $F(2,202) =.90$, $p = .41$.

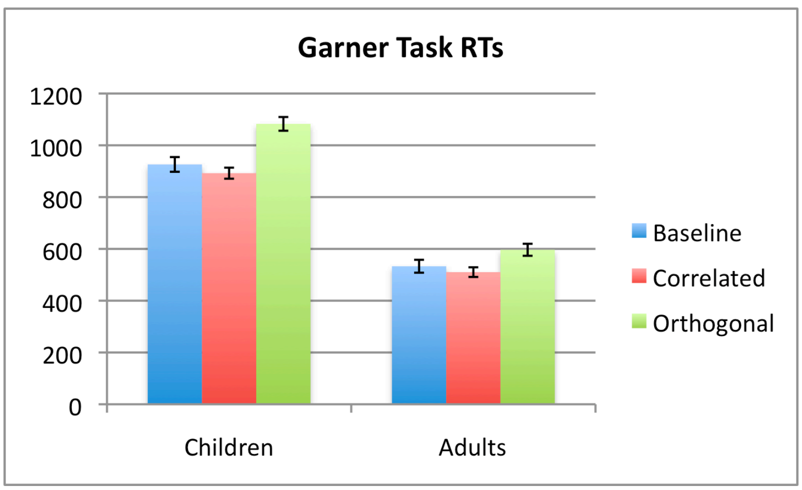


Figure 37. Block differences are more pronounced in children than in adults, even though overall patterns of RTs are comparable.

Children were slower ($M = 967$ ms) overall than adults ($M = 547$ ms), $F(1,101) =188.5$, $p < .001$, but this relationship was tempered by a cross-over interaction with WM, $F(1,101) = 8.3$, $p = .005$ (Figure 38), such that high spans adults were faster than low span adults, $F(1,61) = 10.6$, $p = .002$, whereas high span children were not significantly different from low span children, with the trend going in the opposite direction, with high span children being slower on this task $F(1,40) = 2.1$, $p = .16$.

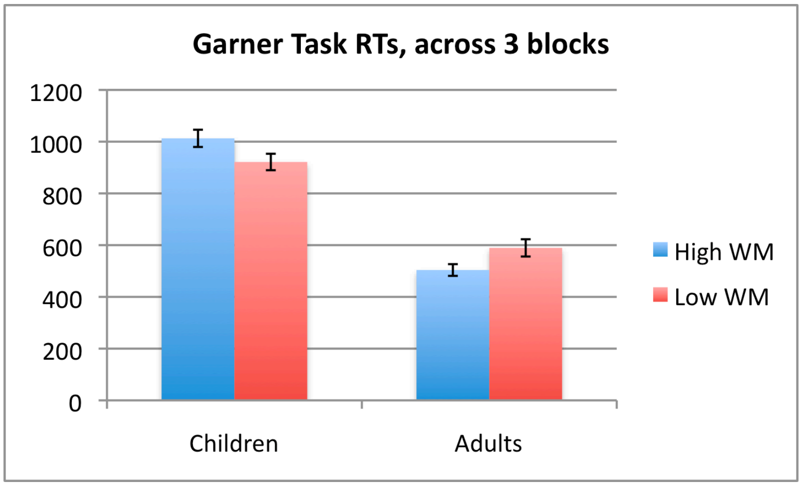


Figure 38. A cross-over interaction between WM and age on RTs in the Garner task: High WM adults are faster than low WM adults (p = .002), but there is a trend in the opposite direction for children (p = .16).

To continue testing the dynamic filtering account, Garner task RTs were examined as a function of each quartile (10 trials for children; 15 trials for adults) in each block, to see how the RTs change as a function of time (Figure 39). The addition of the quartile factor to the above ANOVA did not produce any significant changes to the effects described above. The effect of quartile was almost marginal, $F(3,606) = 2.0$, $p = .12$, such that RTs decreased during the middle two quartiles, across blocks. Critically, the quartile factor did not interact with any of the other factors (WM, population age, block, and their interactions), all p's > .26. These findings suggest that children and adults performed this task very similarly, by speeding up slightly in the Correlated block and slowing down in the Orthogonal block.

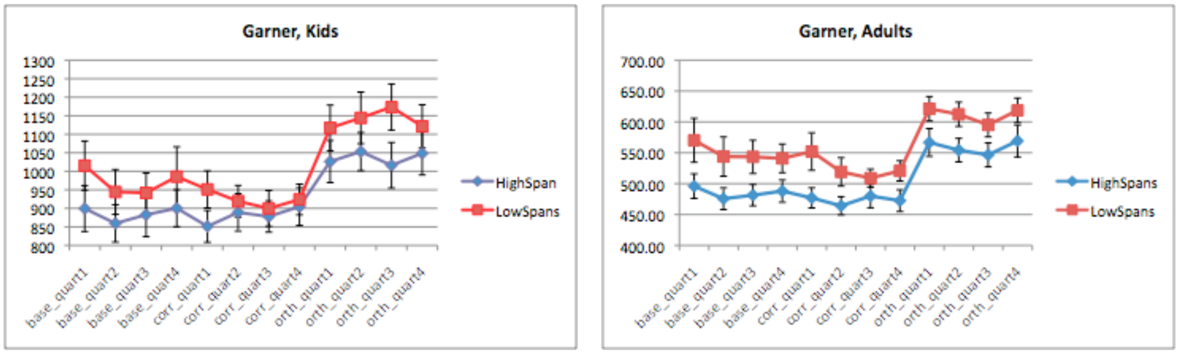


Figure 39. The pattern of Garner task RTs as a function of WM and quartile for each of the three blocks was remarkably similar across children and adults.

Perceptual Priming Task

For both children and adults, perceptual priming scores decreased in the course of the task, likely because perceptual priming rapidly decays with time (Maljkovic & Nakayama, 2000). Specifically, priming scores during the first half of the test phase (10 trials) were larger (M = 1.3, SE = .12 for kids; M = 1.2, SE = .10 for adults) than priming scores in the second half of the test phase (M = .53, SE = .14 for kids; M = .46, SE = .14 for adults), $F(1,41) = 16.1$, $p <$ .001 for kids, and $F(1,24) = 20.2$, $p < .001$ for adults. Thus, the perceptual priming scores from the first half the task were used in subsequent analyses.

In children, there was a marginal relationship between performance on the perceptual priming task and WM, such that perceptual priming performance was n*egatively* correlated with WM (Figure 40A), $R = .28$, $N = 42$, $p = .07$. In adults, there was no relationship between WM and perceptual priming performance (Figure 40), $R = .14$, $N = 24$, $p = .51$; but the sample size was limited, and these results should thus be interpreted with caution. Nevertheless, these findings suggest that high WM is not associated with improved performance on all cognitive tasks. At least early in development, high WM may impair performance on tasks that depend upon more stimulus-specific processing of detailed information, and there might be a dissociation, at the level of individual differences, toward more abstract, categorical processing supported by the PFC and the more stimulus-specific types of processing, supported by the posterior cortical regions (DeCaro, Thomas & Beilock, 2008; Kharitonova et al., in prep.).

(A) (B)

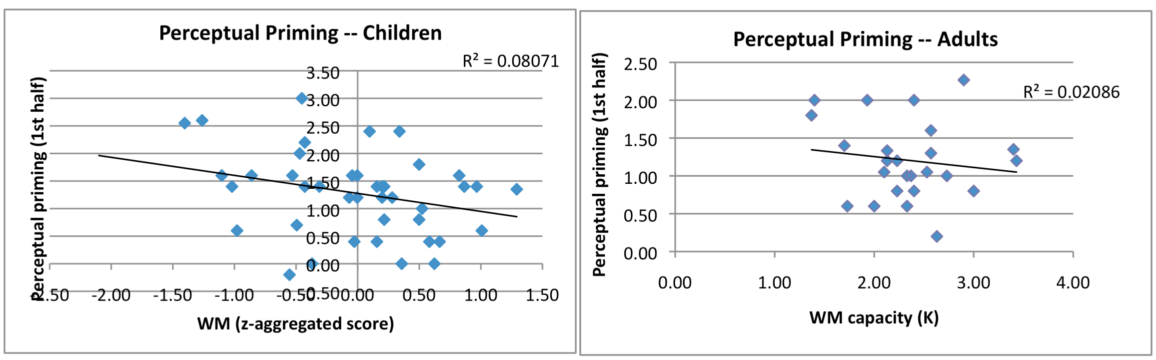


Figure 40. Perceptual priming scores are marginally negatively related to WM in kids (A) and not related to WM in adults (B).

Discussion

The purpose of Experiment 3 was to extend the investigation of the relationship between filtering and WM to new contexts (the Flanker task and the Garner task), which allow examining whether filtering can be adjusted *dynamically*, within the task, in both adults and six-year-old children. The Flanker task was ultimately more sensitive in eliciting meaningful individual differences in performance and the associated relationship with WM. Both the child and the adult versions of the Flanker task provided evidence for the dynamic filtering account, whereas data from the Garner task results were more ambiguous. Overall, results provided evidence that high WM can support dynamic adjustment of filtering strategy in both kids and adults, across both shorter time scales (across trials) and longer time scales (across blocks and tasks).

Flanker results

The child Flanker task provided three pieces of evidence suggesting that high WM can support dynamic shifts in filtering strategy, even among six-year-old children, whose working memory and filtering abilities are still underdeveloped (Bunge, Dudokovic, Thomason, Vaidya, & Gabrieli, 2002; Conklin, Luciana, Hooper & Yarger, 2007; Riggs et al., 2006). First, high WM was associated with a marginally smaller incongruency cost. Second, high WM participants were not consistent in the relationship between their incongruency costs across Block 1 and Block 3, whereas low WM participants were very consistent, both at the group level, and at the level of the individual differences, which is surprising from standard accounts, given that low WM is usually associated with more variability in behavior (Hultsch, MacDonald, & Dixon, 2002; Stuss, Murphy, Binns, & Alexander, 2003). Finally, high WM participants were less accurate on neutral trials in the context of incongruent blocks than in the context of the congruent block, whereas low WM subjects were again very consistent.

The first point (the incongruency cost finding) could imply either high static or dynamic filtering for the high spans. Both static high and dynamic filtering profiles predict a small incongruency cost; the dynamic filtering profile also predicts a large congruency benefit (due to task-advantageous low filtering); however, it is difficult to directly compare the magnitudes of these costs and benefits, given potential differences in underlying difficulty levels. However, the latter two points suggest that high spans were filtering *dynamically,* such that they were strongly influenced by the task context. First, high WM participants showed no relationship between incongruency costs in the first and the third Incongruent blocks, after encountering the Congruent block, suggesting a change in filtering strategy based on task context, whereas low WM participants' performance was consistent. Second, high WM subjects adjusted their accuracy on the neutral trials based on the overall block context, such that their neutral trials accuracy increased from the first (Incongruent) block to the second (Congruent) block and again decreased in the third (Incongruent block). These results suggest that high WM could support flexible and dynamic allocation of top-down control to support either high or low filtering, based on task demands, in children as young as six years old.

The adult Flanker task produced four main pieces that support the dynamic filtering account. First, high spans showed a greater incongruency cost than low spans. Second, high spans showed a greater congruency benefit than low spans. Third, low spans were actually slowed on congruent trials relative to neutral (while high spans were sped). Finally, high spans showed a larger conflict adaptation effect than low span participants, consistent with the idea that high WM can support flexible allocation of cognitive control in response to varying task demands.

The first two findings suggest that high spans were filtering *less* than low spans. The third finding suggests that low spans might be slowed in general by both types of flankers. Nonetheless, even with this general slowing for low spans, high WM participants still showed greater incongruency costs than low WM subjects. There are several possible mechanistic explanations for these obtained results. For example, filtering could reflect top-down support that limits the number of items that exceed threshold (such that low spans use more top down support to limit their window more), while local, competitive inhibition supports resolution of the target from those flankers (such that low spans have worse competitive inhibition so suffer more from flankers in general, even with their narrower window).

This account suggests that low spans apply a relatively low and fixed amount of filtering across different task contexts, while high spans are able to update their filtering strategy, based on task demands, as predicted. This idea is further supported by accuracy data from the child Flanker and Garner tasks, in which low spans showed remarkably stable performance across the three blocks, whereas high spans' accuracy changed significantly as a function of the block context (i.e. neutral trials accuracy in congruent vs. incongruent blocks in the in Flanker task, and the change in accuracy across Baseline and Correlated blocks in the Garner task).

Why was high WM associated with a smaller incongruency cost in children and a larger incongruency cost in adults? The adult Flanker task had interleaved trials, which perhaps led high spans to strategically filter less than low spans, consistent with dynamic updating (within the study, but not across trials). In contrast, the child Flanker tasks separated congruency and incongruency by blocks (such that the first and the third blocks contained only incongruent and neutral trials, and the second block contained only the congruent and neutral trials). The blocked setup might have encouraged shifting of the filtering strategy (which was successfully done by the high span participants)

There is a discrepancy between the current finding of high spans having a larger incongruency cost and a finding in the literature (Redick & Engle, 2006), which reports the opposite pattern of high spans showing a smaller incongruency cost. Despite tasks being identical (as in Fan et al., 2002), other factors may have influenced the opposite direction of the findings. In the current study, the Flanker task was administered after about approximately 40 minutes of other demanding tasks, whereas in the Redick & Engle study, the Flanker task was the first and only task administered. Thus, high spans could be fatigued in the current experiment and thus opt for the easier low filtering strategy, given that this strategy is particularly advantageous in the interleaved setup of the Flanker task.

Follow-up studies should continue examining the effects of blocked vs. interleaved trials in both children and adults; for example, by presenting adults with the blocked version of the Flanker task and the children with the interleaved version, in order to separate the effects of age from effects of task setup.

Garner Results

The results from the Garner task were more ambiguous with regard to informing the relationship between WM and filtering. The accuracy data from children can be interpreted to be consistent with the dynamic filtering account, given that high spans improved in accuracy between Baseline and Correlated blocks, potentially due to lowering the amount of filtering and taking advantage of the task-irrelevant but consistent identity information, whereas low spans performance was remarkably uniform across the three blocks. However, static filtering could also explain these results: accuracy on the Correlated block is expected in increase relative to the Baseline block if filtering is consistently imperfect, and some irrelevant (but useful in the context of the Correlated block) information is processed. Adult participants showed no WM modulation on Garner task accuracy for any of the blocks, thus also limiting the ability to inform the dynamic filtering account.

In terms of RTs, high spans children were faster than low span children only on the Orthogonal block of the Garner task, suggesting that high WM is associated with better ability to ignore the irrelevant identity information on the Orthogonal trials. However, the lack of the relationship between WM and the incongruency cost (i.e. the difference between Orthogonal and Baseline RTs) makes it difficult to precisely interpret these findings. High WM adults were faster on all types of blocks in the Garner task, with no effect of WM on the incongruency cost or the congruency benefit. This overall speedup of high span adults might have potentially masked any (small) differences in filtering. However, it is unclear why this masking would only be observed in the Garner, but not in the Flanker task, where there was also an overall effect of high WM being associated with faster RTs on all trial types, but also an effect of WM on both the incongruency cost and the congruency benefit.

There are several possibilities for why the Garner task might be less sensitive with regard to individual differences in filtering. One, the task was not technically speeded. Participants were asked to respond as fast as possible, but the faces remained on the screen until the participant made a response, a setup that was identical to that in the previous study, which used this paradigm in children and adults (Baudouin et al., 2008). Thus, the pace of this task was controlled by the participants, and might have therefore, decreased the need to filter. Future work should create a truly speeded version of this paradigm (similar to the Flanker task, where the images were presented on the screen for a maximum of 1700 ms in the adult version and 2500 ms in the child version) in attempt to increase the need to filter task-irrelevant information in the Orthogonal block and take advantage of the redundant information in the Correlated block.

In addition, the Garner task might have a less apparent need for filtering than the other measures, such as the Flanker task, or the Vogel et al. (2005) task used in Experiment 2. In the Orthogonal block of the Garner task, the irrelevant information (face identity) was orthogonal to task-relevant information (emotion), rather than conflicting (as it is in the Flanker task); thus, it was less critical to ignore these task-irrelevant features. Moreover, there might be less apparent benefit to not filtering in the Correlated block of the Garner task, since the task-irrelevant information was correlated rather than identical (as it was in the Flanker task), in the Correlated block. Thus, the setup of the Garner task might have made filtering less necessary and more difficult to do.

The Garner task, however, was suitable for comparing children's and adult's performance, due to identical task setup across the two age groups. Results revealed very similar patterns overall: a slight speedup in the Correlated block and a slowdown in the Orthogonal block, which were not mediated by age or WM. However, the overall speed was modulated by WM such that high WM adults performed faster than low WM adults; however, the trend went in the opposite direction for children: high WM was associated with a trend toward slower responding. The WM-associated speedup in adults is consistent with a number of similar findings reported in this dissertation, where WM was associated with faster performance (Flanker results), stronger attentional modulation, and perhaps stronger control over memory retrieval (Experiment 2 results). The trend toward a WM-associated slowdown in children is consistent with other findings reporting that in contrast to adults, children do not slow down sufficiently when they encounter more difficult problems (e.g. when needing to integrate across two separate dimensions in the Ravens' task) relative to easier problems (e.g. when needing to reason within a single dimension in the Raven's task) (Crone, Wendelken, van Leijenhorst, Honomichl, Christoff, & Bunge, 2009). Perhaps high WM children in the current experiment were the ones who allocated sufficient time to process the stimuli in the most task-appropriate manner, thus demonstrating slightly longer RTs. Future studies need to further examine this interesting possibility.

Conflict adaptation

The results from the conflict adaptation analyses in the adult Flanker task support the dynamic filtering hypothesis. High WM participants showed a larger conflict adaptation effect, consistent with the idea that high WM supports a more flexible allocation of cognitive control, based on the immediate task demands. The child version of the Flanker task was less well set up to examine the conflict adaptation idea, given its blocked design, which is less sensitive to sequential trial effects. Nevertheless, this analysis showed that children are also sensitive to sequential congruency effects. However, in children the conflict adaptation effect was not modulated by WM, unlike in adults. This could have occurred because children are less able to flexibly allocate cognitive control based on task demands, or instead, the lack of WM modulation could simply be due to the blocked setup of the child Flanker paradigm, which was less suited for finding subtle sequential effects. Future work will explore the relationship between WM and conflict adaptation in children in the interleaved design, similar to the one used in the adult version of the task in the current experiment.

The adult conflict adaptation results also help to inform the long-standing debate regarding the nature of the conflict adaptation effect (Egner, 2007; Mayr & Awh, & Laurey, 2003; Nieuwenhuis et al, 2006; Ullsperger, Bylsma, & Botvinick, 2005). The interpretation of this effect in terms of detecting conflict and up-regulating cognitive control on incongruent trials, leading to faster performance on subsequent incongruent trials (i.e. the *conflict monitoring* account) is only one of several proposed interpretations of the behavioral effect. Another prominent interpretation of this effect does not invoke cognitive control, and instead posits an episodic memory based explanation of the phenomenon (the *feature integration* account, sometimes also referred to as *repetition priming* account). Specifically, the feature integration model suggests that stimulus and response features that co-occur in time become integrated into a single episodic representation. A subsequent activation of a subset of features (e.g. the stimulus features) will automatically activate the rest of the features (e.g. the response-based features). Thus, complete overlap in all features (both stimulus-and response-based) and complete alternations in features (e.g. >>>>> followed by <<<<<, where no features repeat from one trial to the next) should result in fastest performance. In contrast, overcoming the previous binding of the stimulus and response features, as would be necessary in partial overlap situations, should elicit longer RTs. In the Flanker task, both II and CC trials have either complete overlap or complete changes in features, predicting fastest performance on these types of trials. In contrast, IC and CI trials produce partial overlap, predicting relatively longer RTs on these types of trials.

Several studies have set out to dissociate these possibilities. When controlling for feature integration (i.e. excluding trials with stimulus and response repetitions in a modified Flanker task with multiple possible values, such as the random selections of digits 1-9), the conflict adaptation effect was still observed (Ullsperger et al., 2005), suggesting that feature integration alone could not account for the effect. Similar findings were obtained in another study, where the Flanker task was modified to allow controlling for repetition priming by including six different stimulus features that were mapped to three different response choices: the conflict adaptation effect remained even after controlling for feature integration (Verbruggen, Notebaert, Liefooghe, & Vandierendonck, 2006). In contrast, another study did not find any conflict adaptation after controlling for feature integration (Mayr et al, 2003); however, in that experiment participants needed to switch between responding to vertical and horizontal stimuli (a manipulation intended to minimize feature overlap). In addition to minimizing feature overlap, this manipulation might have added a *task*-switching demand to this paradigm, and executive demands needed to switch between tasks (Monsell, 2003) might have concealed potential conflict adaptation effects (Enger, 2007; Ullsperger et al., 2005). Thus, the existing literature suggests that the conflict adaptation effect may still be present even when feature integration has been (correctly) controlled for.

The current results appear to lend additional support to the conflict monitoring account of the conflict adaptation effect. It is unlikely that the observed difference would result from overcoming repetition priming across stimuli with overlapping features, because the repetition priming account predicts effects that go in the opposite direction of the observed effects. Namely, if overcoming priming was related to WM, then high WM participants should be better able to overcome priming, thus showing a *smaller* conflict adaptation effect; however, high WM was associated with larger conflict adaptation. Of course, these two accounts are not mutually exclusive, and therefore, both can contribute to the observed effect (Egner, 2007). Nevertheless, the current findings lend additional support to the idea that the conflict adaptation effect is driven, at least in part, by contextual conflict monitoring, a prefrontal, top-down control ability related to WM (Botvinick et al., 2001; Chatham, Claus, Kim, Curran, Banich, & Munakata, in prep.).

However, additional analyses that included the neutral trials provided support for both explanations of the conflict adaptation effect. When there are only two values for each of the two dimensions, the complete vs. partial alternations are completely confounded with types of congruency (Enger, 2007), and thus the two accounts cannot be cleanly dissociated. The neutral trials can be used to attempt to reconcile the two possibilities. The neutral trials were similar in terms of conflict to the congruent trials (both have low conflict), but different from the incongruent trials (incongruent trials have high conflict). So if the conflict adaptation effect stems for the amount of conflict across different trials, then neutral trials should be treated similarly to the congruent trials and differently from the incongruent trials. However, neutral and congruent trials have very different features, and so if the conflict adaptation account stems from feature overlap, then congruent trials and neutral trials should be treated differently. Results support both of these possibilities (see Figure 41). On the one hand, neutral trials (red line) show complete overlap with congruent trials (blue line), consistent with the conflict monitoring account. On the other hand, incongruent trials that follow neutral trials (NI) behave very differently from incongruent trials that follow congruent trials (CI), $p = .004$, and very similarly to incongruent trials following incongruent trials (II), $p = .36$. Thus, RTs are longer when there is the most amount of partial feature overlap (CI trials), and shorter when there is either complete overlap or zero/minimal feature overlap (II trials and NI trials, respectively), consistent with the feature integration account.

Thus, it appears that both conflict monitoring and feature integration contribute to the conflict adaptation effect, which is consistent with previous conclusions (e.g. Egner, 2007). However, as described above, it appear that WM modulation of the conflict adaptation effect likely stems from the conflict monitoring, and not from the feature overlap aspect of this effect.

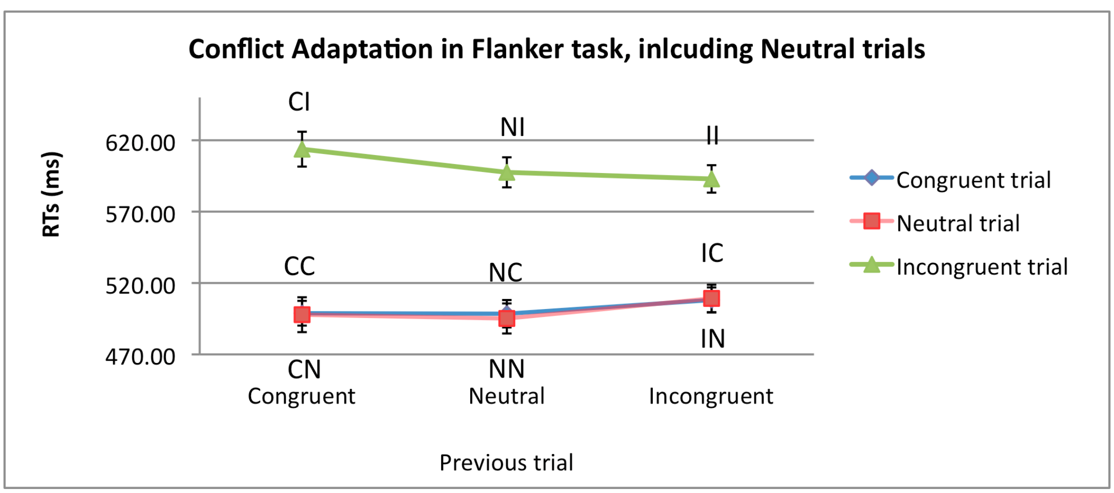


Figure 41. Neutral trials were treated both similarly to congruent trials (red line and blue lines are overlapping) and differently from congruent trials (CI RTs were longer than NI RTs), suggesting that both the amount of conflict and the amount of feature overlap contribute to conflict adaptation.

Ego depletion

Some of the results are consistent with *ego depletion* (e.g. Baumeister, Bratslavsky, Muraven, & Tice, 1998; Schmeichel, Vohs & Baumeister, 2003; Vohs & Heatherton, 2000), a notion that exerting cognitive control earlier in the session will deplete limited resources and leave less control available to use in subsequent tasks. Thus, assuming relatively comparable initial resources (within the same age of participants), those participants who were more successful on demanding executive tasks (i.e. high WM participants) might be more susceptible to the effects of ego depletion later in the session.

For instance, with regard to the accuracy-based results in the child Flanker task, it is possible that everyone starts out with fairly high accuracy in the first block (around 94%). In the congruent second block, high spans are able to switch to filtering less (based on task demand), and their neutral trials accuracy increases, given that low filtering makes the congruent task easier overall. Critically, in the third, incongruent block, high spans may recruit a mix of strategies: some participants may return to high filtering, whereas others may continue to use the low filtering strategy, resulting in lower neutral accuracy overall. In addition to this general "confusion" or mixture of filtering strategies, high spans may be more susceptible to ego depletion, since they have just engaged in the more executively demanding high filtering earlier in the task, and also potentially needed to switch between filtering strategies, which may have increased cognitive demands even further. Ego depletion may thus have contributed to lower neutral trial accuracy in the third block, and the smaller correlation between incongruent costs across first and third blocks, for high spans children only.

Further support for the ego depletion among high span children comes from several other recent experiments. First, children who switch between sorting rules (an ability thought to be supported by working memory, as shown in Blackwell, Cepeda, & Munakata, 2009) early in an hour-long experimental session, subsequently perform *worse* on a response inhibition task (an ability thought to also rely on PFC function (e.g. Aron, Robbins, & Poldrack, 2004) than children than who perseverated on the card-sorting task (Blackwell, 2010). In addition, piloting work for these experiments demonstrated that 6-year-old children showed a positive relationship between performance on working memory and visual priming tasks when the tasks are grouped closer together. However, as reported above, 6-year-olds show a *negative* relationship between WM and priming when the tasks are temporally separated, such that several other demanding tasks (which might have contributed to depleting executive resources) were administered. It is possible that children are more susceptible to ego depletion than adults because their more limited executive functions might be easier to deplete.

Future work needs to more closely examine this question by designing studies that parametrically manipulate the timing of the executive functions tasks within the session, to test for possible ego depletion effects, in both children and adults. Further, the assumption of comparable initial resources may not be justified; high WM individuals might possess more “resources”, and thus be less susceptible to ego depletion effects. Existing literature if vague with respect to identifying the nature of these resources; thus, future work needs to more carefully examine this issue.

WM-associated benefits

For the adult participants, WM predicted faster RTs on all trial types, across both Flanker and Garner tasks. This WM-associated speedup might be surprising from purely capacity-based perspectives on WM (e.g. Baddeley, 2003; Daneman & Carpenter, 1980). However, viewing WM more mechanistically, in terms of the ability to constrain top-down processing in the most task-efficient manner is consistent with these findings. High WM allows participants to actively maintain the task-relevant goals of responding as quickly as possible in both filtering tasks in this experiment. These findings are consistent with the Experiment 2 results, which suggested that high WM is associated with enhanced early perceptual and attentional processing, not typically associated with executive functions.

In children, the role of WM is less clear. One the one hand, children demonstrated no WM-associated speedup on the Flanker task and a trend in the direction of the slowdown on the Garner task. This suggests that WM plays a very different role in supporting behaviors early in development. It is possible that six-year-old children are not maintaining information in the fully proactive way, like adults do (Chatham et al., 2009). A more reactive strategy predicts the lack of WM-associated speedup that was observed in this experiment (due to high WM not being related to active maintenance, but rather, to retrieval-based processes). Such reactive processes might also be advantageous for slowing down responses overall, consistent with the observed trend in the child Garner task, and with the interpretation that in children, high WM might be associated with the ability to sufficiently slow down on difficult trials (e.g. Crone et al., 2009).

Other pieces of evidence argue in favor of children at least partially using adult-like, proactive strategies. Six-year-olds children are likely just starting to use proactive strategies

(Chatham & Munakata, in prep.), and thus there could be a mix of proactive and reactive strategies across children, or even within children on different trials or different tasks. For example, if children processed all WM tasks reactively, the potentially more retrieval-based Complex Span task should have correlated with performance on other WM tasks. In addition, children demonstrated a negative relationship between WM tasks, thought to tap PFC regions and the perceptual priming task, thought to tap more posterior brain regions. This finding is interpreted in terms of a possible early dissociation between more abstract, categorical processes that are related to WM, and more graded, fine-tuned processes that are related to perceptual priming. It is not clear why retrieval-based WM would be negatively related to perceptual priming performance, where keen attention to details is required. These interpretations are also consistent with the idea that top-down executive control (and goal maintenance in particular) is especially critical for speeding up processing speed tasks early in development, when even relatively easy tasks might require executive control (Cepeda, Blackwell, & Munakata, in prep.). Thus, future studies need to investigate the nature of WM in six-year-old children to help better understand its relationship to related processes, such as filtering task-irrelevant information, or reasoning abstractly.

General Discussion

The purpose of this dissertation was to investigate the relationship between WM and filtering, across situations with both high and low demand for filtering task-irrelevant information, in both adults and children. Most existing work has focused on examining unidirectional relationships between WM and filtering, such that some studies argued for a positive relationship (e.g. Fukuda & Vogel, 2009; Kane et al., 2005; Long & Prat, 2002; Vogel et al, 2005), while others argued for a negative relationship (e.g. Just & Carpenter, 1992; Waring et al., 2009), without attempts to reconcile this seeming conflict in existing patterns. This dissertation attempted to reconcile this apparent conflict by positing that high WM can support *both* high and low filtering, based on what is currently the most efficient strategy for a given task. This dynamic adjustment could occur via modulation of top-down control to focus on either a smaller subset of available information (high filtering) or a larger subset, encompassing both relevant and potentially irrelevant pieces of information (low filtering).

Summary of Findings

Three experiments tested the theory that high WM can support dynamic adjustment of filtering strategy, to support both high and low filtering, based on what is currently most task-advantageous. The first experiment demonstrated that participants with high WM exhibited a high filtering profile when high filtering was most advantageous, and a low filtering profile when low filtering was more appropriate. In the high filtering task, participants needed to ignore distractors (blue items) while making same/different judgments regarding the spatial orientation of the targets (red items). Filtering the distractors lowered the WM demand and was thus task-advantageous. In contrast, in the low filtering demand task, participants needed to switch between attending to color of an item to its shape, and vice versa. In the overlapping blocks, the currently irrelevant feature (e.g. the *purple* color of a purple triangle in the shape trial) became relevant on the subsequent switch trial (e.g. a *purple* circle presented in the color trial). Thus, broadly representing both the currently relevant (triangle) and the currently irrelevant (purple) features of an item was task-advantageous, since it allowed for faster processing of the subsequent trial.

However, given that filtering demand was manipulated in such different paradigms in Experiment 1 (visual WM task versus a task-switching paradigm), it is impossible to limit the interpretation of the differences in the obtained results solely to the different demands for filtering. Thus, Experiments 2 and 3 manipulated filtering demand within single paradigms. Experiment 2 modified the previously high filtering demand task (from Experiment 1) to create both high and low filtering demand versions. In the high filtering demand version, the distractors changed *orthogonally* to the targets; thus, paying attention to the distractors was disadvantageous, and therefore, distractors should have been filtered. In contrast, in the low filtering demand version, the distractors changed *consistently* with the targets, such that paying attention to either targets or distractors would produce the correct answer; thus, there was not much reason to filter distractors. Unfortunately, a small manipulation in the setup of Experiment 2 (i.e. extending the duration of the arrow that indicated which side of the screen participants needed to attend to for that trial) contributed to producing ERP results that were difficult to interpret. In Experiment 2 participants paid more attention to the task-irrelevant side of the screen early in the trial, which was most likely the result of the increased arrow duration.

Participants continued to attend less to the relevant side of the screen even after the onset of the memory array, relative to the first experiment; this continued attention to the irrelevant side likely also contributed to producing the confusing ERP results. These results included low ERP filtering estimates overall, with no modulation of filtering by WM, despite comparable behavioral performance. Despite these shortcomings, the results of this experiment were informative in highlighting the sensitivity of attention to subtle changes in experimental setup and the fragility of the seemingly robust effects.

Experiment 3 expanded the scope of the main question of this dissertation by testing whether the relationship between WM and filtering could dynamically change within task, in both adults and six-year-old children. The two filtering tasks that were used were a variant of the Flanker task (Eriksen & Eriksen, 1974; Rueda et al., 2004; 2005), and the Garner task (Baudouin et al., 2008; Garner, 1974). In the Flanker task participants needed to respond to the direction of a central item (left or right), while ignoring the flanking items, which could be either congruent (i.e. go in the same direction as the central item) or incongruent (i.e. go in the opposite direction of the central item). In the Garner task, participants needed to make judgments on values of one dimension (e.g. face emotion: happy or sad), while ignoring the values of another dimension (e.g. face identity: person A or person B); the irrelevant (identity) information could be invariant, consistent, or orthogonal to the relevant (emotion) information. Data from the Flanker task were mostly consistent with the dynamic filtering account, for both children and adults, although there were several critical caveats that are described in the Discussion section for Experiment 3 and the Limitations and Future Directions section below. Data from the Garner task were more difficult to interpret, perhaps due to the task's self-administered, unspeeded pace and its blocked design with orthogonal, instead of directly conflicting information. Both of these factors might have reduced the need to filter-task irrelevant information. Thus, the results of this dissertation provide preliminary support to the idea that high WM can be associated with both high and low amounts of filtering of task-irrelevant information, but future work is needed to more definitively test this theory.

Theoretical Implications

Although the results of this dissertation are not definitively conclusive, they extend the previous work by providing evidence that high WM can support *dynamic,* within-task adjustment of filtering strategies. This ability to dynamically adjust filtering strategies appears to be at least partially in place by six years of age. These results, although still tentative and awaiting future investigation, nevertheless can inform our understanding of how the ability to maintain information in working memory across delays and interference and the ability to adjust the focus on attention are related.

Relationship between WM and filtering

The link between WM and the ability to efficiently deploy attention is well established in the literature (Awh, Vogel, & Oh, 2006; Bleckley, Durso, Crutchfield, Engle, & Khanna, 2003; Kane & Engle, 2003; Lavie & de Fockert, 2005; Yi, Woodman, Widders, Marois, & Chun, 2004). Most of the experiments, however, have focused on the positive association between WM and filtering. For example, participants with high WM are less susceptible to various forms of long-term memory interference, such as the fan-effect and the retroactive and proactive interference (reviewed in Kane & Engle, 2002; 2003). High WM participants are also better at focusing exclusively on the target in the antisaccade task (Kane, Bleckley, Conway & Engle,

2001) and at naming the color of the ink, instead of reading the conflicting color word in the Stroop task (Kane & Engle, 2003; Long & Prat, 2002).

WM modulation of the Stoop effect is particularly apparent in conditions where most trials are congruent (such that reading the word will produce the correct answer on most trials), and the goal (to name the ink color) needs to be actively maintained for optimal performance on the incongruent trials (Kane & Engle, 2003). Thus, high WM participants might be better at actively maintaining the currently relevant task goals. However, in all of these tasks, the predominant task goal was to narrowly focus on the task-relevant and ignore the irrelevant information. Thus, we do not know from these experiments whether high WM unidirectionally supports high filtering, or whether it supports the most efficient allocation of attention, to dynamically support either high or low filtering, and we have simply not yet seen many tests of the situations where low filtering was required. In fact, many of the general intelligence tests have narrowly focused on precisely the situations that require high filtering (e.g. Hatch & Garnder, 1986). Results from this dissertation support the latter view, suggesting that WM can support the most *efficient* filtering strategy for a given situation and can thus support both high and low filtering. This view is consistent with the conceptualization that PFC-supported cognitive control supports "flexible regulation of behavior in the pursuit of internal goals" (Egner, 2007). In this dissertation such flexibility is thought to stem from up- and down-regulation of top-down control signals from the PFC regions to the posterior cortical regions, to allow either less or more information to be processed.

Flexibility takes a long time to develop ontogenetically (e.g. Blackwell et al., 2009; Zelazo, Frye, & Rapus, 1996); hence, from this perspective children ought to be able to control filtering less flexibly than adults, and thus show patterns that are less dynamic. These results are somewhat consistent with the existing data, since there are slightly more pieces of evidence supporting the dynamic filtering account in adults than in children. However, follow-up studies need to be conducted to more definitively compare the amount of dynamic adjustment of filtering across children and adults. Younger children should be even less flexible than the six-year-olds tested here, and therefore should be more impaired at dynamically adjusting filtering strategies. In addition, flexible allocation of cognitive control is impaired in many disorders, such as schizophrenia (e.g. Morice & Delahunty, 1996), so these populations should be particularly impaired at *dynamically* adjusting their filtering, despite not showing pronounced deficits in unidirectional filtering measures (e.g. Henik et al., 2002).

This dissertation has been advocating for the importance of flexible adjustment of filtering strategies to achieve optimal performance. However, too much flexibility might also be suboptimal. For example, variance related specifically to the ability to quickly shift between behaviors, after controlling for general executive functions (shifting-specific variance) is *negatively* related to IQ (Friedman et al., 2008), suggesting that too little stability can be detrimental to performance. In addition, PFC-supported flexibility might be harmful early in development because it impedes several types of learning, including learning conventions (Thompson-Schill, Ramscar & Chrysikou, 2009). Thus, flexibility is important for certain tasks (such as those requiring both high and low filtering), but might be suboptimal for others (e.g. IQ measures), and there might be important reasons why hypofrontality (and the associated lack of flexibility, along with perhaps less flexible adjustment of filtering strategies) develops slowly.

Broad effects of WM

Across the experiments, high WM incurred significant and often unexpected benefits on performance. For example, in adults (but not children), high WM was associated with faster responses even on baseline trials, on which there was arguably no obvious need to filter in either the Garner or Flanker tasks. WM was also associated with enhanced early perceptual and attentional processes in Experiment 2 ERP results. These results might be unexpected from the standard, capacity-based views of WM (e.g. Baddeley, 2003; Daneman & Carpenter, 1980), because maintaining more or fewer items in a static memory store should not necessarily be related to the scope of attention. However, these results are consistent with the conceptualization of WM in terms of the underlying mechanisms, such as active maintenance, updating, and top-down biasing (e.g. O'Reilly & Frank, 2006; Miller & Cohen, 2001), which should allow for flexible allocation of top-down control to enhance and speed up processing of information deemed to be relevant.

The finding of WM incurring benefits even on baseline trials is consistent with many recent investigations, which found that across a number of RT-based tasks, including the Flanker task used in this dissertation, bilingual participants were faster than their monolingual peers not only on conflict-related measures (e.g. the incongruency cost), but critically, also on baseline measures (e.g. neutral trials) that had no conflict to overcome (Bialystok & DePape, 2009; Bialystok et al., 2004; Bialystok, Craik & Ryan, 2006; Costa, Hernandez & Sebastian-Galles, 2008). The authors often leave these findings uninterpreted; however, these findings are very consistent with the presented framework and findings. Lifelong practice in interpreting and speaking in two languages may increase one's ability to maintain two or more meanings of the words bilinguals are exposed to, and to proactively monitor their environment. This maintenance practice may train WM, which in turn may help to bias processing in a way that encourages goal maintenance and helps to speed up processing even on simple baseline trials, in addition to the more complex trials where conflict needs to be overcome (Kharitonova & Miyake, in prep.).

Unpublished data from my own experiments are consistent with this interpretation: bilingual children were comparable to monolinguals in all measures other than the digit span measure, in which bilinguals outperformed monolinguals. (Advantages on baseline trials were only observed on RT-based, and not accuracy-based tasks (Bialystok & DePapa, 2009; Bialystok et al., 2004; Bialystok, Craik & Ryan, 2006; Costa, Hernandez & Sebastian-Galles, 2008), likely because RTs are more sensitive than accuracy-based data, and so it is not surprising that there we no advantages on the rest of our accuracy-based measures). Other studies have not found the bilingual advantage on the digit span and other purely capacity based measures of WM; thus, future work needs to focus on examining this issue further.

Critically, high WM is not a panacea for all cognitive tasks. On the perceptual priming task, where implicitly processing perceptual features that define the representations of an object (such as line elements of a drawing) forms the critical aspect of the task (Wiggs & Martin, 1998), children showed a marginal negative relationship with WM, while adults showed no relationship between WM and perceptual priming (although the limited sample size in adult participants precluded definitive interpretations). These results are consistent with existing work that shows high WM can be detrimental to processes that are more graded and implicit, such as implicit categorization tasks (DeCaro, Thomas, & Beilock, 2008). In addition, participants with COMT alleles associated with higher WM function (e.g. Durstewitz & Seamans, 2008) perform worse when they receive prior instructions that are inaccurate for the current task than participants with COMT alleles associated with lower WM function (Doll, Hutchinson, & Frank, under review). These findings are interpreted to stem from high WM participants' extreme reliance on explicit instructions, a finding consistent with prior research (Engle et al., 1991) and with the characterization of the PFC as supporting *explicit*, abstract types of reasoning (e.g. Badre, Kayser, & D'Esposito, 2010), in addition to supporting WM functions. High WM participants are also more susceptible to pressure-induced failure on demanding mathematics problems, possibly because high WM makes them more prone to using more explicit, categorical strategies, which become impaired under high pressure (Beilock & Carr, 2005). Finally, there are important developmental advantages for the slow development of the PFC, and the associated functions, such as WM (Thompson-Schill et al., 2009).

All of these results point to a potential dissociation between prefrontally-supported abstract and active representations (Kharitonova, Chien, Colunga, & Munakata, 2009) that enable flexible allocation of WM, and posterior cortical stimulus-specific representations that underlie more precise, perceptual-based, and perseverative behaviors. Future work needs to further explore this potential dissociation in the types of representations supported by prefrontal and posterior neural regions. The presence of such dissociation needs to be more definitively established, preferably in experimental (instead of purely individual differences) paradigms. Moreover, it is important to test whether the extent of the dissociation varies across development. For example, it is possible that children might show a larger tradeoff between using more prefrontal and more posterior regions (and relying on the associated mechanisms), by being "stuck" in one mode or the other, because of less well integrated neural regions (e.g. Fair et al., 2009) and low overall flexibility (e.g. Zelazo et al., 1996).

Limitations and Future Directions

There are some important limitations to this work, but also a number of promising avenues for future directions. Each experiment had its own set of limitations. Experiment 1 manipulated filtering demand across two very different paradigms; thus, factors other than the demand for filtering may have contributed to the obtained results. A major problem with

Experiment 2 involved the change in the design that increased the duration of the arrow, which indicated which side of the screen the participants needed to attend for that trial. This manipulation had been intended to decrease eye movements; instead, the rate of eye movements did not change, but the manipulation had inadvertent effects on attention, such that participants attended more to the task-irrelevant side of the screen early in the trial, which likely affected their filtering later in the trial, and produced results that were difficult to interpret. Future work needs to test the dynamic filtering account in the version of the task with the shorter arrow duration. Nevertheless, current results point to the fragility of this paradigm, and the high sensitivity of early attention to small changes in experimental setup. Future studies should also utilize an eye-tracker, which would detect eye movements and would allow excluding individual trials where eye movements were committed, instead of excluding participants with many eye movements. Additionally, applying Independent Components Analysis (ICA) might be fruitful in removing trials with eye movement artifacts (Jung, Makeg, Westerfield, Townsend, Courchesne, & Sejnowski, 2000). Given that participants likely treated the task very differently in Experiment 2, as indexed by very early changes in laterality of attention, it is less promising to apply this technique to the existing dataset.

In Experiment 3, a major caveat stems from the way in which the child and the adult versions were set up. The tasks were designed to be as similar as possible for the child and adult participants, while also building closely on existing work with these populations. As a result, all trial types (incongruent, congruent, and neutral) in the adult Flanker task were interleaved, whereas trials were blocked in the child version, such that the first and the third block only contained incongruent and neutral trials, while the second block only contained the congruent and neutral trials. Each setup has its own advantageous, as each is most appropriate for some analyses, but not others. For example, the interleaved version of the adult Flanker task enabled a stronger comparison of the conflict adaptation idea, because the sequential effects were more apparent when the congruent and the incongruent trials were intermixed. In contrast, the blocked version of the child Flanker task was more appropriate for examining the dynamic filtering theory on a longer time scale: the middle congruent block might have detrimental effects on filtering in the last (incongruent) block, if one were able to dynamically update their filtering strategy, especially if such updating is difficult to implement. Thus, the next step will involve testing adults with the blocked version and the children with the interleaved version of the Flanker task, to test which effects are due to development and which are due to differences in task setup.

More broadly, next steps should focus on identifying the time-course of the dynamic adjustment (i.e. figuring out whether the changes occur across trials, across blocks, or even across tasks), and whether this time-course is dependent on the developmental stage of participants. Specifically, WM could affect filtering strategy in a very immediate sense (as reflected in conflict adaptation), or it could show long-lasting effects, spanning several blocks, or even tasks (as reflected by potential ego depletion effects). The underdeveloped WM in children may be sufficiently strong to support short-term changes in filtering strategies (e.g. in conflict adaptation), but may not be strong enough to support long-term benefits; if this were so, then children should be more susceptible to ego depletion effects. These ideas need to be examined in future work.

Finally, theoretically, it is not yet clear whether the amount of filtering is based on the *strength* of top-down control (such that high filtering is supported by strong control, and low filtering is supported by weak control), or rather based on the *breadth* of attentional scope (such that high filtering is supported by narrow attentional scope, and low filtering is supported by broad attentional scope). These possibilities do not have to be mutually exclusive. For example, high WM might support both stronger control and narrower focus (e.g. high WM participants might strongly maintain "attend to red items only", while low WM participants might weakly maintain "attend to everything you see"). The strength of the maintained signal likely constitutes at least part of the explanation, since the strength of the recurrent connectivity, which supports WM (e.g. Morton & Munakata, 2002) likely also helps to more strongly maintain only the task-relevant information. However, breadth of the attentional scope might also be important, because just weakly maintaining "red items" might not help to encode blue items, unless the blue items are particularly salient. It might be the case that breadth and strength are inherently related, such that decreasing strength automatically increases breadth, leading both to irrelevant items leaking in when cognitive control is weakened, and to relevant items being maintained and attended to less strongly. Computational models constitute an ideal medium for testing these possibilities, as they allow for each of these underlying mechanisms to be directly manipulated to observe subsequent effects on performance (e.g. Munakata, 2004), and thus should be used in future work to directly investigate these possibilities.

General Conclusions

The purpose of this dissertation was to explore the relationship between WM and filtering task-irrelevant information. The view of WM in terms of the underlying mechanisms, such as the ability to exert top-down control for processing and maintaining the task-relevant information, instead of a storage capacity to be filled, predicts that these processes should be reflected in performance in dynamic, rather than static ways. So, instead of high WM always supporting high filtering, the view presented here predicts that high WM can also be related to low filtering, in situations where low filtering is the more advantageous strategy. High filtering could result from upregulating cognitive control and actively maintaining task goals, leading to processing only the task-relevant information. Low filtering could result from loosening control over posterior areas, to allow for a larger amount of information to be processed. The dynamic filtering account was tested in three different experiments, with both adults and six-year-old children. Results overall were consistent with the predictions and showed that dynamic filtering could occur at both short and long time scales, in both children and adults. Several important caveats described to be addressed before making definitive conclusions; nevertheless, the current results provide an important advance in understanding the role of WM in dynamically adjusting filtering strategies based on task demands.

References

Aron, A., Robbins, T., & Poldrack, R. (2004). Inhibition and the right inferior frontal cortex. *Trends in cognitive sciences , 8* (4), 170-177.

Awh, E., Vogel, E., & Oh, S. (2006). Interactions between attention and working memory. *Neuroscience , 139* (1), 201-208.

Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience , 4* (10), 829-839.

Badre, D., Kayser, A., & D'Esposito, M. (2010). Frontal Cortex and the Discovery of Abstract Action Rules. *Neuron , 66* (2), 315-326.

Barrouillet, P., Gavens, N., Vergauwe, E., Gaillard, V., & Camos, V. (2009). Working memory span development: A time-based resource-sharing model account. *Developmental psychology , 45* (2), 477-490.

Baudouin, J.-Y., Durand, K., & Gallay, M. (2008). Selective attention to facial identity and emotion in children. *Visual Cognition , 16* (7), 933-952.

Beilock, S., & Carr, T. (2005). When high-powered people fail. *Psychological Science, 16 (2)* , 101–105.

Bellis, T., Nicol, T., & Kraus, N. (2000). Aging Affects Hemispheric Asymmetry in the Neural Representation of Speech Sounds. *Journal of Neuroscience , 20* (2), 791.

Berkson, J. (1978). In dispraise of the exact test: Do the marginal totals of the 2X2 table contain relevant information respecting the table proportions? *Journal of Statistical Planning and Inference , 2* (1), 27-42.

Bertolino, A., Blasi, G., Latorre, V., Rubino, V., Rampino, A., Sinibaldi, L., et al. (2006). Additive effects of genetic variation in dopamine regulating genes on working memory cortical activity in human brain. *Journal of Neuroscience , 26* (15), 3918-3922.

Besson, M., Kutas, M., & Petten, C. (1992). An event-related potential (ERP) analysis of semantic congruity and repetition effects in sentences. *Journal of Cognitive Neuroscience , 4* (2), 132-149.

Bialystok, E., & DePape, A. (2009). Musical expertise, bilingualism, and executive functioning. *Journal of Experimental Psychology: Human Perception and Performance , 35* (2), 565-574.

Bialystok, E., Craik, F., & Ryan, J. (2006). Executive Control in a Modified Antisaccade Task: Effects of Aging and Bilingualism . *Journal of Experimental Psychology: Learning, Memory, and Cognition , 32* (6), 1341-1354.

Bialystok, E., Craik, F., Klein, R., & Viswanathan, M. (2004). Bilingualism, Aging, and Cognitive Control: Evidence From the Simon Task. *Psychology and Aging , 19* (2), 290-303.

Blackwell, K. (2010). Mechanisms of cognitive control: Contributions from working memory and inhibition to task switching. *Ph.D. Dissertation*, Univeristy of Colorado.

Blackwell, K., Cepeda, N., & Munakata, Y. (2009). When simple things are meaningful: Working memory strength predicts children's cognitive flexibility. *Journal of Experimental Child Psychology , 103* (2), 241-249.

Bleckley, M., Durso, F., Crutchfield, J., Engle, R., & Khanna, M. (2003). Individual differences in working memory capacity predict visual attention allocation. *Psychonomic Bulletin & Review , 10* (4), 884-889.

Botvinick, M., Braver, T., Barch, D., Carter, C., & Cohen, J. (2001). Conflict Monitoring and Cognitive Control. *Psychological Review , 108* (3), 624-652.

Braver, T., Barch, D., Keys, B., Carter, C., Cohen, J., Kaye, J., et al. (2001). Context Processing in Older Adults: Evidence for a Theory Relating Cognitive Control to Neurobiology in Healthy Aging. *Journal of Experimental Psychology: General , 30* (4), 746-763.

Brignani, D., Bortoletto, M., Miniussi, C., & Maioli, C. (2010). The when and where of spatial storage in memory-guided saccades. *NeuroImage , 52*, 1611-1620.

Brumback, C., Low, K., Gratton, G., & Fabiani, M. (2004). Sensory ERPs predict differences in working memory span and fluid intelligence. *Neuroreport , 15* (2), 373.

Bunge, S., Dudukovic, N., Thomason, M., Vaidya, C., & Gabrieli, J. (2002). Immature frontal lobe contributions to cognitive control in children evidence from fMRI. *Neuron , 33* (2), 301-311.

Buschman, T., & Miller, E. (2007). Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Science , 315*, 1860-1862.

Cepeda, N., Blackwell, K., & Munakata, Y. (n.d.). Speed isn't everything: Complex processing speed measures mask individual differences and developmental changes in executive control.

Chao, L., & Knight, R. (1995). Human prefrontal lesions increase distractibility to irrelevant sensory inputs. *Neuroreport , 6* (12), 1605-1610.

Chatham, C. H., Claus, E. D., Kim, A., Curran, T., Banich, M. T., & Munakata, Y. (under review). Inhibition is out of control.

Chatham, C. H. & Munakata, Y. (in prep). From reactive to proactive: A shift in the dynamics of cognitive control in children between 5 and 6 years old.

Chatham, C., Frank, M., & Munakata, Y. (2009). Pupillometric and behavioral markers of a developmental shift in the temporal dynamics of cognitive control. *Proceedings of the National Academy of Sciences , 106* (14), 5529-5333.

Colflesh, G., & Conway, A. (2007). Individual differences in working memory capacity and divided attention in dichotic listening. *Psychonomic Bulletin & Review , 14* (4), 699-703.

Conklin, H., Luciana, M., Hooper, C., & Yarger, R. (2007). Working memory performance in typically developing children and adolescents: Behavioral evidence of protracted frontal lobe development. *Developmental neuropsychology , 31* (1), 103-128.

Costa, A., Hernandez, M., & Sebastiangalles, N. (2008). Bilingualism aids conflict resolution: Evidence from the ANT task. *Cognition , 106* (1), 59-86.

Cowan, N. (2001). The magical number 4 in short-term memory: a reconsideration of mental storage capacity. *The Behavioral and brain sciences , 24* (1), 87-185.

Cowan, N. (2005). Understanding intelligence: A summary and an adjustable-attention hypothesis. In O. Wilhelm & R.W. Engle (Eds.), *Handbook of understanding and measuring intelligence* (pp. 469-488). London: Sage.

Crone, E., Wendelken, C., van Leijenhorst, L., Honomichl, R., Christoff, K., & Bunge, S. (2009). Neurocognitive development of relational reasoning. *Developmental Science , 12* (1), 55-66.

Cycowicz, Y., Friedman, D., Snodgrass, J., & Rothstein, M. (2000). A developmental trajectory in implicit memory is revealed by picture fragment completion. *Memory , 8* (1), 19-35.

Daneman, M., & Carpenter, P. (1980). Individual differences in working memory and reading. *Journal of verbal learning and verbal behavior , 19* (4), 450-466.

de Fockert, J., Rees, G., Frith, C., & Lavie, N. (2001). The role of working memory in visual selective attention. *Science , 291*, 1803-1806.

DeCaro, M., Thomas, R., & Beilock, S. (2008). Individual differences in category learning: Sometimes less working memory capacity is better than more. *Cognition , 107* (1), 284-294.

Diamond, A., & Kirkham, N. (2005). Not quite as grown-up as we like to think. *Psychological Science , 16* (4), 291-297.

Doll, B., Hutchison, K., & Frank, M. (under review). Dopaminergic genes predict individual differences in susceptibility to confirmation bias.

Drew, T., McCollough, A., Horowitz, T., & Vogel, E. (2009). Attentional enhancement during multiple-object tracking. *Psychonomic Bulletin & Review , 16* (2), 411-417.

Durstewitz, D., & Seamans, J. (2008). The dual-state theory of prefrontal cortex dopamine function with relevance to catechol-o-methyltransferase genotypes and schizophrenia. *Biological Psychiatry , 64* (9), 739-749.

Egner, T. (2007). Congruency sequence effects and cognitive control. *Cognitive, Affective, & Behavioral Neuroscience , 7* (4), 380-390.

Eimer, M. (1996). The N2pc component as an indicator of attentional selectivity. *Electroencephalography and Clinical Neurophysiology , 99* (3), 225-234.

Elsabbagh, M., & Johnson, M. (2010). Getting answers from babies about autism. *Trends in cognitive sciences , 14* (2), 81-87.

Elward, R., & Wilding, E. (2010). Working memory capacity is related to variations in the magnitude of an electrophysiological marker of recollection. *Brain Research , 1342*, 55-62.

Emrich, S., Al-Aidroos, N., Pratt, J., & Ferber, S. (2009). Visual Search Elicits the Electrophysiological Marker of Visual Working Memory. PLoS One, *4* (11), 1-9.

Engle, R., Carullo, J., & Collins, K. (1991). Individual differences in working memory for comprehension and following directions. *Journal of Educational Research , 84* (5), 253-263.

Eriksen, B., & Eriksen, C. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics , 16* (1), 143-149.

Fair, D., Cohen, A., Power, J., Dosenbach, N., Church, J., Miezin, F., et al. (2009). Functional Brain Networks Develop from a "Local to Distributed" Organization. *PLoS Computational Biology , 5* (5), 1-14.

Fan, J., McCandliss, B., Sommer, T., & Raz, A. (2002). Testing the efficiency and independence of attentional networks. *Journal of Cognitive Neuroscience , 14* (3), 340-347.

Friedman, N., & Miyake, A. (2005). Comparison of four scoring methods for the reading span test. *Behavior Research Methods , 37* (4), 581-590.

Friedman, N., Miyake, A., Young, S., DeFries, J., Corley, R., & Hewitt, J. (2008). Individual differences in executive functions are almost entirely genetic in origin. *Journal of experimental psychology. General , 137* (2), 201-225.

Fukuda, K., & Vogel, E. (2009). Human Variation in Overriding Attentional Capture. *Journal of Neuroscience , 29* (27), 8726-8733.

Fukuda, K., Awh, E., & Vogel, E. (2010). Discrete capacity limits in visual working memory. *Current opinion in neurobiology , 20*, 177-182.

Garner, W. R. (1974). *The processing of information and structure.* Potomac, MD: Lawrence Erlbaum Associates, Inc.

Geisser, S., & Greenhouse, S. (1958). An extension of Box's results on the use of the F distribution in multivariate analysis. *The Annals of Mathematical Statistics , 29* (3), 885-891.

González, C., Fuentes, L., Carranza, J., & Estévez, A. (2001). Temperament and attention in the self-regulation of 7-year-old children. *Personality and Individual Differences , 30* (6), 931-946.

Gratton, G., Coles, M., & Donchin, E. (1992). Optimizing the use of information: Strategic control of activation of responses. *Journal of Experimental Psychology: General , 121* (4), 480-506.

Green, A., Fugelsang, J., Kraemer, D., Shamosh, N., & Dunbar, K. (2006). Frontopolar cortex mediates abstract integration in analogy. *Brain Research , 1096*, 125-137.

Hatch, T., & Gardner, H. (1986). From testing intelligence to assessing competences: A pluralistic view of intellect. *Roeper Review , 8* (3), 147-150.

Henik, A., Carter, C., Salo, R., Chaderjian, M., Kraft, L., Nordahl, T., et al. (2002). Attentional control and word inhibition in schizophrenia. *Psychiatry Res , 110* (2), 137-149.

Hillyard, S., & Anllo-Vento, L. (1998). Event-related brain potentials in the study of visual selective attention. *Proceedings Of The National Academy Of Sciences Of The United States Of America , 95* (3), 781-787.

Hillyard, S., Vogel, E., & Luck, S. (1998). Sensory gain control (amplifcation) as a mechanism of selective attention: electrophysiological and neuroimaging evidence. *Phil. Trans R. Soc. Lond. , 353*, 1257-1270.

Hopf, J., Luck, S., Girelli, M., Hagner, T., Mangun, G., Scheich, H., et al. (2000). Neural sources of focused attention in visual search. *Cerebral Cortex , 10* (12), 1233-1241.

Hultsch, D., MacDonald, S., & Dixon, R. (2002). Variability in reaction time performance of younger and older adults. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences , 57* (2), P101-P115.

Ikkai, A., McCollough, A., & Vogel, E. (2010). Contralateral delay activity provides a neural measure of the number of representations in visual working memory. *Journal of Neurophysiology , 103* (4), 1963-1968.

Jha, A., & Kiyonaga, A. (2010). Working-Memory-Triggered Dynamic Adjustments in Cognitive Control. *Journal of Experimental Psychology: Learning , 36* (4), 1036-1042.

Judd, C., McClelland, G., & Ryan, C. (n.d.). Data Analysis: A Model Comparison Approach.

Jung, T., Makeig, S., Westerfield, M., Townsend, J., Courchesne, E., & Sejnowski, T. (2000). Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects. *Clinical Neurophysiology , 111* (10), 1745-1758.

Just, M., & Carpenter, P. (1992). A Capacity Theory of Comprehension: Individual Differences in Working Memory. *Psychological Review , 99* (1), 122-149.

Kane, M., & Engle, R. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic Bulletin & Review , 9* (4), 637-671.

Kane, M., & Engle, R. (2003). Working-memory capacity and the control of attention: The contributions of goal neglect, response competition, and task set to stroop interference. *Journal of Experimental Psychology: General , 132* (1), 47-70.

Kane, M., Bleckley, M., & Conway, A. (2001). A Controlled-Attention View of Working-Memory Capacity. *Journal of Experimental Psychology , 130* (2), 169-183.

Kane, M., Brown, L., McVay, J., Silvia, P., Myin-Germeys, I., & Kwapil, T. (2007). For whom the mind wanders, and when. *Psychological Science , 18* (7), 614.

Karmiloff-Smith, A. (1998). Development itself is the key to understanding developmental disorders. *Trends in cognitive sciences , 2* (10), 389-398.

Karmiloff-Smith, A., Thomas, M., Annaz, D., Humphreys, K., Ewing, S., Brace, N., et al. (2004). Exploring the Williams syndrome face-processing debate: the importance of building developmental trajectories. *Journal of Child Psychology and Psychiatry , 45* (7), 1258-1274.

Kharitonova, M. & Miyake, A. (in prep). Bilingual advantage in executive functions: A Critical review.

Kharitonova, M., Chien, S., Colunga, E., & Munakata, Y. (2009). More than a matter of getting'unstuck': flexible thinkers use more abstract representations than perseverators. *Developmental Science , 12* (4), 662-669.

Kharitonova, M., Hulings, J., & Munakata, Y. (in prep.). The Role of Representations in Executive Function: Investigating a Developmental Synergy Between Flexibility and Abstraction.

Klein, R. (2000). Inhibition of return. *Trends in cognitive sciences , 4* (4), 138-147.

Lavie, N. (2005). Distracted and confused?: Selective attention under load. *Trends in cognitive sciences , 9* (2), 75-82.

Lavie, N., & de Fockert, J. (2005). The role of working memory in attentional capture. *Psychonomic Bulletin & Review , 12* (4), 669-674.

Lavie, N., Hirst, A., de Fockert, J., & Viding, E. (2004). Load theory of selective attention and cognitive control. *Journal of Experimental Psychology: General , 133* (3), 339-354.

Long, D., & Prat, C. (2002). Working memory and Stroop interference: An individual differences investigation. *Memory and Cognition , 30* (2), 294-301.

Lord, F. M. (1963). Elementary models for measuring change. In C. W. Harris (Ed.), Problems in measuring change (pp. 21–38). Madison, WI: University of Wisconsin Press.

Luck, S., & Hillyard, S. (1994). Spatial Filtering During Visual Search: Evidence From Human Electrophysiology. *Journal of Experimental Psychology: Human Perception and Performance , 20* (5), 1000-1014.

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

Luck, S., Woodman, G., & Vogel, E. (2000). Event-related potential studies of attention. *Trends in cognitive sciences , 4* (11), 432-440.

MacLeod, C. M. (2007). The concept of inhibition in cognition. In: D.S. Gorfein and C.N. MacLeod, Editors, Inhibition in cognition, APA, Washington, pp. 3–24.

Maljkovic, V., & Nakayama, K. (2000). Priming of popout: III. A short-term implicit memory system beneficial for rapid target selection. *Visual Cognition , 7* (5), 571-595.

McCollough, A., Machizawa, M., & Vogel, E. (2007). Electrophysiological measures of maintaining representations in visual working memory. *Cortex , 43*, 77-94.

McDonald, J., Hickey, C., Green, J., & Whitman, J. (2009). Inhibition of return in the covert deployment of attention: Evidence from human electrophysiology. *Journal of Cognitive Neuroscience , 21* (4), 725-733.

McNab, F., & Klingberg, T. (2007). Prefrontal cortex and basal ganglia control access to working memory. *Nature Neuroscience , 11* (1), 103-107.

Miller, E., & Cohen, J. (2001). An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci., 24*, 167-202.

Monsell, S. (2003). Task switching. *Trends in cognitive sciences , 7* (3), 134-140.

Morice, R., & Delahunty, A. (1996). Frontal/executive impairments in schizophrenia. *Schizophrenia Bulletin , 22* (1), 125-137.

Morton, J., & Munakata, Y. (2002). Active versus latent representations: A neural network model of perseveration, dissociation, and decalage. *Developmental psychobiology , 40* (3), 255-265.

Munakata, Y. (2004). Computational cognitive neuroscience of early memory development. *Developmental Review , 24* (1), 133-153.

Murray, A., Kuo, B., Stokes, M., & Nobre, A. (2009). The Effect of Retro-Cueing on an ERP Marker of VSTM Maintenance. *precedings.nature.com* .

Nieuwenhuis, S., Stins, J., Posthuma, D., Polderman, T., Boomsma, D., DE, G., et al. (2006). Accounting for sequential trial effects in the flanker task: Conflict adaptation or associative priming? *Memory & cognition , 34* (6), 1260.

Olesen, P., Macoveanu, J., Tegner, J., & Klingberg, T. (2007). Brain activity related to working memory and distraction in children and adults. *Cerebral Cortex , 17* (5), 1047.

Ophir, E., Nass, C., & Wagner, A. (2009). Cognitive control in media multitaskers. *PNAS , 106* (37), 15583-15587.

O'Reilly, R., & Frank, M. (2006). Making working memory work: a computational model of learning in the prefrontal cortex and basal ganglia. *Neural Computation , 18* (2), 283-328.

Redick, T., & Engle, R. (2006). Working memory capacity and attention network test performance. *Applied Cognitive Psychology , 20* (5), 713-721.

Riggs, K., Mctaggart, J., Simpson, A., & Freeman, R. (2006). Changes in the capacity of visual working memory in 5- to 10-year-olds. *Journal of experimental child psychology , 95* (1), 18-26.

Rueda, M., Fan, J., McCandliss, B., Halparin, J., Gruber, D., Lercari, L., et al. (2004). Development of attentional networks in childhood. *Neuropsychologia , 42* (8), 1029-1040.

Rueda, M., Posner, M., Rothbart, M., & Davis-Stober, C. (2004). Development of the time course for processing conflict: an event-related potentials study with 4 year olds and adults. *BMC neuroscience , 5* (1), 39.

Rueda, M., Rothbart, M., McCandliss, B., Saccomanno, L., & Posner, M. (2005). Training, maturation, and genetic influences on the development of executive attention . *PNAS , 102* (41), 14931-14936.

Rugg, M.D. & Coles, M.G.H. (1995). Electrophysiology of mind: event-related brain potentials and cognition. Oxford: Oxford University Press.

Rugg, M., & Curran, T. (2007). Event-related potentials and recognition memory. *Trends in cognitive sciences , 11* (6), 251-257.

Samuel, A., & Kat, D. (2003). Inhibition of return: A graphical meta-analysis of its time course and an empirical test of its temporal and spatial properties. *Psychonomic Bulletin and Review , 10* (4), 897-906.

Sander, M., Werkle-Bergner, M., & Lindenberger, U. (2010). Age and individual differences in working memory capacity: Insights from ERPs. *CNS poster* .

Schacter, D., & Buckner, R. (1998). Priming and the brain. *Neuron , 20*, 185-195.

Schmeichel, B., Vohs, K., & Baumeister, R. (2003). Intellectual Performance and Ego Depletion: Role of the Self in Logical Reasoning and Other Information Processing. *Journal of Personality and Social Psychology , 85* (1), 33-46.

Srinivasan, R., Nunez, P., Tucker, D., Silberstein, R., & Cadusch, P. (1996). Spatial sampling and filtering of EEG with spline laplacians to estimate cortical potentials. *Brain Topography , 8* (4), 355-366.

Stuss, D., Murphy, K., Binns, M., & Alexander, M. (2003). Staying on the job: the frontal lobes control individual performance variability. *Brain , 126* (11), 2363-2380.

Thompson-Schill, S., Ramscar, M., & Chrysikou, E. (2009). Cognition Without Control. *Current Directions in Psychological Science , 18* (5), 259-263.

Ullsperger, M., Bylsma, L., & Botvinick, M. (2005). The conflict adaptation effect: it's not just priming. *Cognitive, Affective, & Behavioral Neuroscience , 5* (4), 467-472.

Verbruggen, F., Notebaert, W., Liefooghe, B., & Vandierendonck, A. (2006). Stimulus-and response-conflict-induced cognitive control in the flanker task. *Psychonomic Bulletin & Review , 13* (2), 328-333.

Vogel, E., & Machizawa, M. (2004). Neural activity predicts individual differences in visual working memory capacity. *Nature , 428* (6984), 748-51.

Vogel, E., McCollough, A., & Machizawa, M. (2005). Neural measures reveal individual differences in controlling access to working memory. *Nature , 438* (7067), 500-503.

Vohs, K., & Heatherton, T. (2000). Self-regulatory failure: A resource-depletion approach. *Psychological Science , 11* (3), 249-254.

Wang, L., Most, S., & Hoffman, J. (2010). Contralateral delay activity is sensitive to the spatial distribution of items in working memory: An ERP study. *Journal of Vision , 9* (8), 599-599.

Wang, R., & Spelke, E. (2002). Human spatial representation: Insights from animals. *Trends in cognitive sciences , 6* (9), 376-382.

Waring, J., Payne, J., Schacter, D., & Kensinger, E. (2010). Impact of individual differences upon emotion-induced memory trade-offs. *Cognition & Emotion , 24* (1), 150-167.

Wiggs, C., & Martin, A. (1998). Properties and mechanisms of perceptual priming. *Current opinion in neurobiology , 8* (2), 227-233.

Woodman, G., & Luck, S. (1999). Electrophysiological measurement of rapid shifts of attention during visual search. *Nature , 400* (6747), 867-868.

Yadon, C., Bugg, J., Kisley, M., & Davalos, D. (2009). P50 sensory gating is related to performance on select tasks of cognitive inhibition. *Cognitive, Affective, & Behavioral Neuroscience , 9* (4), 448-458.

Yi, D., Woodman, G., Widders, D., Marois, R., & Chun, M. (2004). Neural fate of ignored stimuli: dissociable effects of perceptual and working memory load. *Nature Neuroscience , 7* (9), 992-996.

Zacks, R., & Hasher, L. (n.d.). Directed Ignoring: Inhibitory regulation of working memory.

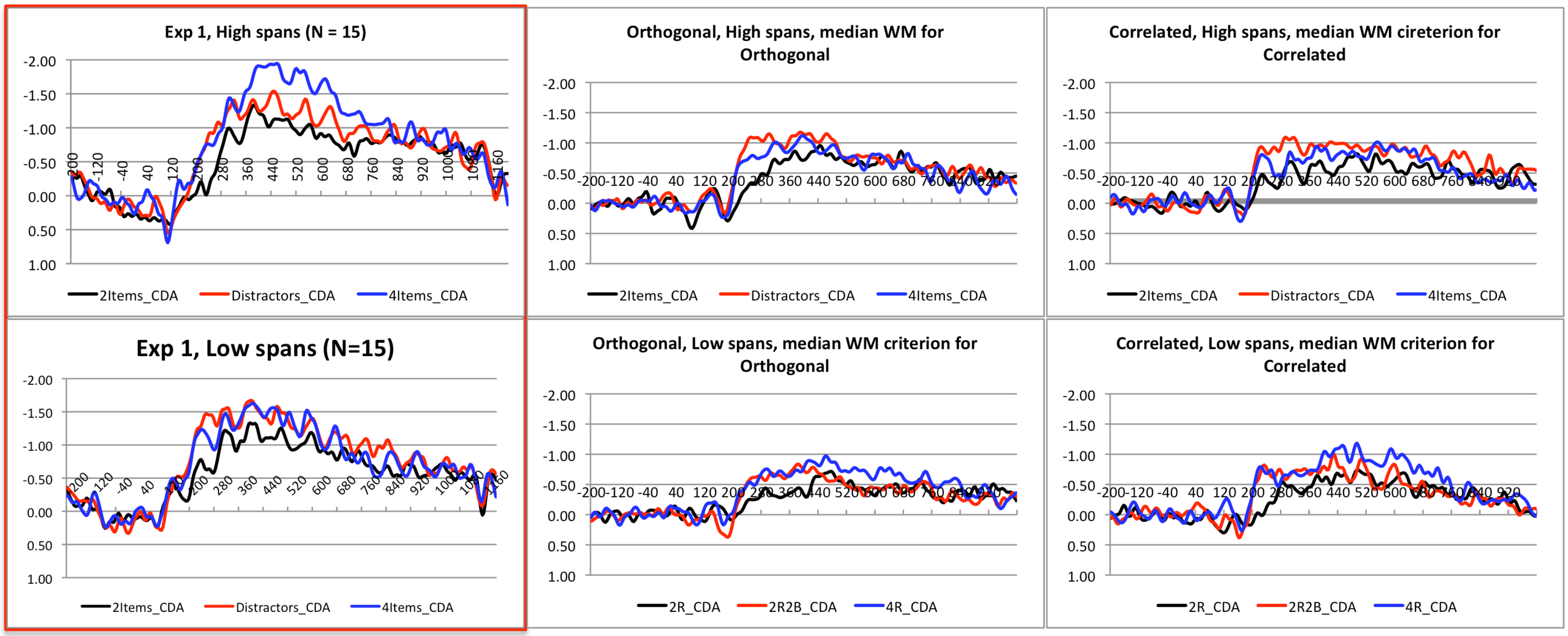
Zanto, T., & Gazzaley, A. (2009). Neural suppression of irrelevant information underlies optimal working memory performance. *Journal of Neuroscience , 29* (10), 3059.

Zelazo, P., Frye, D., & Rapus, T. (1996). An age-related dissociation between knowing rules and using them. *Cognitive Development , 11*, 37-63.

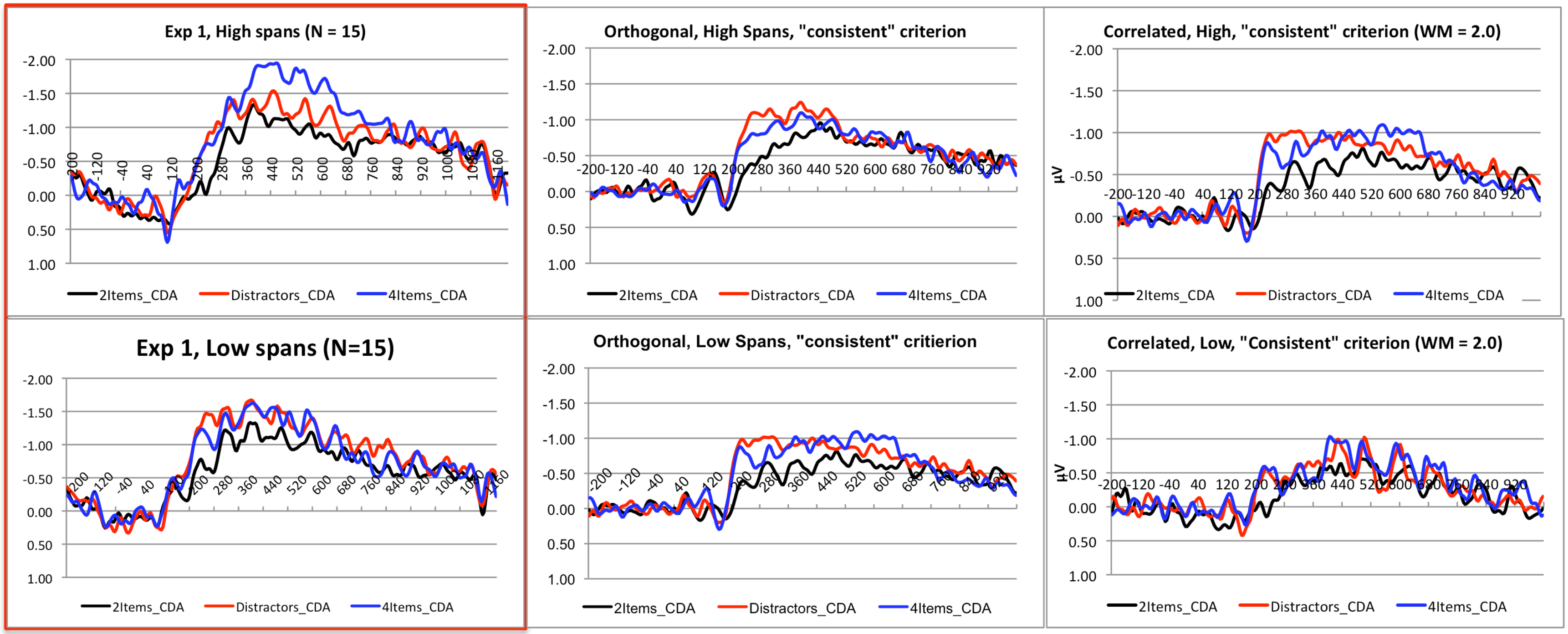
APPENDIX A

WM Span (High/Low basis)		Experiment 1	Experiment 2 – Orthogonal	Experiment 2 – Correlated	Implication for Experiment 2
Consistent cutoff for WM across all conditions, K = 2.0	High	CDA WM = -.64 CDA FE = .62 (N=15)	CDA WM = -.14 CDA FE = .31 (N=18)	CDA WM = -.30 CDA FE = 0.22 (N=17)	• No relationship between behavioral WM span and either CDA WM or CDA FE • CDA FE and CDA WM values are numerically smaller in Exp 2 for both high and low spans
	Low	CDA WM = -.29 CDA FE = 0.0 (N=15)	CDA WM = -.30 CDA FE = .22 (N=10)	CDA WM = -.19 CDA FE = .34 (N=17)	
Median split for each condition (K = 2.0 for Exp 1, K = 2.18 for Orthogonal, K = 2.38 for Correlated)	High	CDA WM = -.64 CDA FE = .62 (N=15)	CDA WM = -11 CDA FE = -.60 (N=14)	CDA WM = -.21 CDA FE = -.45 (N=12)	• Low spans filter more than high spans in Exp 2, unlike in Exp 1. • High spans' FE is negative in Exp 2 because CDA amplitudes for Distractors trials are largest (see graphs below). • Relationship between CDA WM and Behavioral WM reverses, such that high spans have less negative CDA WM values
	Low	CDA WM = -.29 CDA FE = 0.0 (N=15)	CDA WM = -.28 CDA FE = 0.62 (N=14)	CDA WM = -.33 CDA FE = 0.68 (N=12)	
High – Low (Z-combined based on forward span, backward span, behavioral Luck-Vogel task)	High	No data available	CDA WM = -.11 CDA FE = .13 (N=12)	CDA WM = -.33 CDA FE = .64 (N=10)	• No relationship between behavioral WM span and CDA WM • Correlated Low spans' FE is negative in Exp 2 because CDA amplitudes for Distractors trials are largest. • CDA FE is higher for high than low spans in Correlated, contrary to predictions
	Low		CDA WM = -.34 CDA FE = .34 (N=15)	CDA WM = -.24 CDA FE = -.23 (N=9)	
High – Low (Z-combined, based on forward/backward digit span)	High	No data available	CDA WM = -.13 CDA FE = -.29 (N=12)	CDA WM = -.35 CDA FE = .70 (N=10)	• No relationship between behavioral WM span and CDA WM • Correlated Low spans' FE is negative in Exp 2 because CDA amplitudes for Distractors trials are largest. • CDA FE is higher for high than low spans in Correlated, contrary to predictions
	Low		CDA WM = -.35 CDA FE = .52 (N=13)	CDA WM = -.22 CDA FE = -.41 (N=9)	

- High/low breakdown based on median WM values from each experiment/condition (Median split for each condition (K = 2.0 for Exp 1, K = 2.18 for Orthogonal, K = 2.38 for Correlated)



- High/low breakdown based on median WM value from Experiment 1, to be consistent across experiments (WM K = 2.0)



- High/Low breakdown based on median of Z-combined WM