

INDUCING PROACTIVE CONTROL USING A STROOP CUEING PARADIGM

by

Mariana Rachel Olsen

A thesis submitted in partial fulfillment
of the requirements for the degree

of

Master of Science

in

Psychological Science

MONTANA STATE UNIVERSITY
Bozeman, Montana

April 2014

©COPYRIGHT

by

Mariana Rachel Olsen

2014

All Rights Reserved

DEDICATION

This thesis is dedicated to the best parents anyone could ask for, Keith and Susan Nystuen. They always saw in me a better person than I'll likely ever be. I owe much of this wonderful life I lead to them and their unfaltering love, support, and patience.

I also dedicate this thesis to my aunt Donna and late uncle Dave. Whether it was buying me "Zoo Tycoon" and "Amazon Trail" for Christmas or taking me on behind-the-scenes tours at the Science Museum, they instilled in me a love for learning early on that still burns strong. Into adulthood, they were tremendous benefactors to my education. I owe much of my academic tenacity to them.

ACKNOWLEDGEMENTS

On the academic front, I'd like to extend my gratitude to Dr. Keith Hutchison, for his guidance and patience through data analysis and several drafts; Frank Marchak, for lending us his eye tracker and his expertise on troubleshooting and the many things that affect pupil dilation; Ian Handley, for his valuable feedback; Suzanna Powell, for copious amounts of moral support and occasional technical support; Samuel Sorensen, Maile All, Tess Edgar, Reema Najjar, Ahna Sundling, Katie Hart, and Taryn Dicotomis for their help with data collection; and the Attention and Memory Lab for their feedback and support.

I'd also like to thank those whose support was less direct to this thesis but no less important: Paul Zunkel, who has already done it all and commiserated with me through every step of the process; Brianna Olsen, who didn't take offense when I called her a low span right before forgetting to tell her when to exit off the freeway; Mary Hanvik, Peter Bernard, and my parents, for not letting their eyes glaze over as I explained my thesis to them; Grace and Artemis, for the occasional walks, nuzzles, and snuggle breaks that kept me sane; and finally, William Bernard, who has put up with me for the two-year journey of my master's degree and whose patience, love, and understanding have been nothing short of staggering.

TABLE OF CONTENTS

1. INTRODUCTION	1
Problems Interpreting List-based Evidence for Proactive Control	5
Utility of Predictive Cueing Paradigms	7
Current Research	9
Experiment 1	10
2. METHOD	12
Participants and Design	12
Stimuli	12
Automated Ospan Task	13
Stroop Cueing Task	15
Procedure	15
3. RESULTS	17
RT and Error Rates	17
Recall and Recognition	18
Working Memory Capacity	19
4. DISCUSSION	21
Experiment 2	22
5. METHOD	24
Participants and Design	24
Stimuli	24
Apparatus	25
Procedure	25
6. RESULTS	27
RTs and Error Rates	27
Recall	28
Pupil Dilation	29
Working Memory Capacity	29
Pupil Time course Analysis	32
7. DISCUSSION	36

TABLE OF CONTENTS CONTINUED

8. GENERAL DISCUSSION	39
REFERENCES.....	45

LIST OF TABLES

Table	Page
1. Mean RT and error rates following EASY and HARD cues in Experiment 1	17
2. Mean proportion of recall, hits, false alarms, and corrected recognition score for neutral words following EASY and HARD cues in Experiment 1	19
3. Mean reaction times and errors as a function of cue and WMC in Experiment 1	19
4. Mean proportion of recall, hits, false alarms, and corrected recognition score for neutral words following EASY and HARD cues as a function of WMC in Experiment 1	20
5. Mean reaction times and error rates following EASY and HARD cues in Experiment 2.....	27
6. Mean reaction times and error rates for stimuli following EASY and HARD cues by WMC group in Experiment 2.....	30
7. Mean Proportion of Recall as a Function of Cue and WMC in Experiment 2.....	31
8. Mean Change in Pupil Dilation as a Function of Cue, Accuracy, and WMC in Experiment 2.....	35

LIST OF FIGURES

Figure	Page
1. Mean pupil dilation following EASY and HARD cues by WMC	32
2. Mean pupil dilation across fixation delay following EASY and HARD cues	33
3. Mean pupil dilation over 1500 ms fixation interval by WMC group	34

ABSTRACT

Cognitive control refers to the relative ability to attend to relevant stimuli while ignoring irrelevant stimuli in service of a goal. Braver, Gray, and Burgess (2007) have proposed two complementary forms of control: proactive, which is preparatory in nature, and reactive, which is engaged after a stimulus or imperative event. The ability to use proactive control is often tested using the Stroop task; however, what is usually thought of as evidence for proactive control can be confounded with item-specific effects, sequential effects, and speed-accuracy trade-offs. To remedy this issue, the current study utilized a modified version of the Stroop task to examine the use of proactive control on a trial-by-trial basis. Two experiments tested participants' ability to flexibly engage proactive control in which participants were given 80% predictive EASY or HARD cues indicating whether an upcoming stimulus would be congruent or incongruent, respectively. I hypothesized that participants, especially those high in working memory capacity, would use the HARD cues to engage top-down control to suppress word-reading, leading to a) reduced Stroop interference following HARD cues, b) impaired recall and recognition for neutral words following HARD cues, and c) greater pupil dilation following HARD than EASY cues. In Experiment 1, participants showed reduced Stroop interference for stimuli following HARD relative to EASY cues. This effect was replicated in Experiment 2, with reduced Stroop interference in both reaction times and errors following HARD cues. However, neither recall or pupil dilation differed reliably as a function of cue or WMC. Together, these experiments demonstrate the utility of using a cueing procedure when examining proactive control in the Stroop task. Limitations and future directions in cueing research are discussed.

INTRODUCTION

The world around us is rich with sensory information. Whether we are walking through a crowded subway or a serene wilderness, our senses are flooded with different sights, sounds, smells, and tactile stimulations, all competing for our attention. Successful daily functioning requires that we attend to certain information while ignoring other information. In the field of psychology, this process is known as cognitive control.

Previous research has defined cognitive control as one's relative ability to focus on relevant stimuli while ignoring irrelevant stimuli in service of a goal (Miller & Cohen, 2001; Braver, Gray, & Burgess, 2007). Cognitive control is pervasive and has diverse real-world implications: Driving, studying, playing sports or games, and even many vocations require varying degrees of cognitive control. A decline in cognitive control can also be a reliable early predictor of the development of Alzheimer's disease (Hutchison, Balota, & Duceck, 2010).

Cognitive control is often described in terms of a dual-mechanism framework comprised of automatic and effortful components. In an early model, Norman and Shallice (1986) explained that successful execution of cognitive control requires overcoming a strong habitual response. To illustrate, they described two separate "thread" mechanisms: A bottom-up horizontal thread and a top-down vertical thread. The horizontal thread, *contention scheduling*, was automatic and thus required no component of attention to function. In this thread, an event or stimulus triggers multiple potential behaviors which may be congruent or conflicting. The competition between potential behaviors can be likened to a horse-race in which the strongest of the behaviors

will reach threshold first and, ultimately, be expressed. The vertical thread, referred to as the *Supervisory Attentional System (SAS)*, requires attention to bias the activation of potential behaviors through selective inhibition or facilitation. By removing the attentional component, behavior activation falls solely to contention scheduling, which may lead to sub-optimal behavior when the most habitual response is incorrect.

Arguing that most models at that time were too restrictive to account for the dynamic nature of cognitive control, Braver, Gray, and Burgess (2007) proposed a dual-mechanism account including proactive and reactive forms of control. Unlike other extant models, this framework was able to account for differences in control strategies that varied both between *and* within individuals across situations. As its name implies, *proactive control* is engaged before an imperative event. Similar to Norman and Shallice's (1986) SAS, proactive control is top-down in nature and requires the maintenance of a goal over time. It is an early selection/suppression mechanism by which an individual prepares to attend to (selection) or to ignore (suppression) certain information. Conversely, *reactive control* is engaged after the presentation of a stimulus or event and allows for just enough time to make a response. Reactive control is automatic and able to function without an attentional component. In sum, proactive control allows one to prepare for conflict before it arises, while reactive control requires one to resolve conflict after it is presented.

Thanks to the growing sophistication of neuroimaging techniques, the aforementioned processes have received an impressive amount of support. Researchers typically have participants perform a conflict task known to require top-down control and

collect imaging data to examine which areas of the brain are most active during successful and unsuccessful trials. The Stroop task is widely used in such research (Stroop, 1935). In the Stroop task, participants are presented with a series of color words (e.g., “blue”, “yellow”, “red”, and “green”) one at a time and are instructed to respond by naming the color in which the word is presented, not the word itself. Words appear in one of three conditions: Congruent (e.g., “blue” presented in blue), incongruent (e.g., “blue” presented in red), or neutral (e.g., “legal” presented in blue). Congruent stimuli present little conflict since reading the word and naming the color will both lead to a correct response. However, incongruent stimuli elicit conflicting responses, requiring one to override the pre-potent tendency of word-reading to respond correctly. The inherent conflict presented by incongruent stimuli leads to more errors and longer reaction times (RTs). This robust tendency is known as the “Stroop effect.”

Studies using the Stroop task have consistently implicated the prefrontal cortex (PFC) in the engagement of top-down control (DePisapia & Braver, 2006; Miller & Cohen, 2001; MacDonald, Cohen, Stenger, & Carter, 2000; Vendrell et al., 1995). The PFC orchestrates attention in accordance with current goals by amplifying relevant information and/or dampening irrelevant information (Miller & Cohen, 2001). Support for this notion comes not only from neuroimaging studies, but from studies using patients with frontal lobe lesions. These frontal patients demonstrate marked difficulty in tasks known to require top-down control, such as the Stroop and the Wisconsin Card Sorting Task (Vendrell et al., 1995; Perrett, 1974).

Differences in cognitive control are not limited to clinical populations. As detailed by Braver et al. (2007), many factors can influence one's degree of cognitive control. Individual differences in cognitive control have been examined extensively using behavioral methods. One variable shown to affect cognitive control is *working memory capacity* (WMC), one's capacity for executive functioning (Kane & Engle, 2002). People with greater WMC perform exceptionally well on measures used to test top-down control (Engle, 2002; Hutchison, 2011) such as Stroop, anti-saccade, and dichotic-listening tasks. In these tasks, participants must selectively focus on one attribute while suppressing another and must maintain their goal to do this throughout the course of the task. For example, the anti-saccade task requires participants to look to the opposite direction of a flashing stimulus object in order to identify a briefly presented target letter. This requires them to overcome the habitual response of looking toward a flashing stimulus. Although performance does not differ between high- and low-WMC participants during pro-saccade trials (during which participants are instructed to look toward the flashing stimulus), performance does differ in anti-saccade trials with higher WMC predicting faster and more accurate responses (Kane, Bleckley, Conway, & Engle, 2001; Unsworth, Heitz, Shrock, & Engle, 2005).

More relevant to the current research are findings derived from studies using the Stroop task. Overall, participants with high WMC perform better on incongruent trials than do those with low WMC in terms of RTs and errors (Kane & Engle, 2003; Hutchison, 2011). These differences are enhanced when list-wide proportion-congruency (PC) is high (e.g., 75% of trials are congruent), and reduced or eliminated when list-wide

PC is low (e.g., 25% of trials are congruent). Researchers (Logan & Zbrodoff, 1979; Engle, 2002) propose that lists with low PC serve to repeatedly instill the goal of color-naming because the majority of trials are incongruent. Indeed, De Pisapia and Braver (2006) showed that when PC is low, there is sustained activation in the PFC, indicating prolonged engagement of top-down control. Within high-PC lists, external cues to name the color and not the word are scarce and participants must actively maintain the goal of color-naming, even though word-reading will often lead to a correct response. On the rare incongruent trials, goal-neglect leads to errors or longer RTs as participants overcome the conflicting word and color responses. According to Kane and Engle (2003), low-WMC participants are more likely to fall prey to such goal-neglect than their higher-WMC peers.

Problems Interpreting List-based Evidence of Proactive Control

Although such manipulations of PC seem to influence the level of top-down control, some research suggests that these list-wide effects may be confounded with other factors. For example, Jacoby and colleagues (Jacoby, Lindsay, & Hessels, 2003; Bugg, Jacoby, & Toth, 2008) argued that observed differences between high- and low-PC lists was due to the nature of the items contained in the list, not the list itself, because mostly-congruent/incongruent lists contain mostly-congruent/incongruent items. Specifically, certain items might be more (or less) likely to be congruent or incongruent than others. Rather than exerting list-wide control, participants could be responding to individual items on the basis of their possible congruence (item-specific proportion congruence) or

through incidental stimulus-response learning (contingency). In the former, the stimulus itself serves as an indicator of its possible congruence, leading the participant to exert or relax top-down control. For example, if the word “red” is seldom congruent, the word “red” will automatically signal the need for top-down control. Contingency learning occurs when participants learn word-color associations, making them quicker to respond to certain stimuli than others. For example, if “red” usually appears in green, participants will be quicker to respond “green” to this word compared to responses infrequently paired with the word “red.”

By orthogonally manipulating both list-wide and item-specific PC, Hutchison (2011) found evidence for both list-wide and item-specific levels of control. Further, the effects of both manipulations differed as a function of WMC. Replicating the pattern found by Kane and Engle (2003), Hutchison found that only participants low in WMC demonstrated list-based effects when item-specific effects were controlled, showing smaller Stroop effects in MI lists than MC lists. Further, item-specific effects were greater for low WMC participants. Specifically, these participants showed greater Stroop effects for MC words than MI words, demonstrating their impoverished ability to suppress the word dimension. Importantly, this study suggests that list-based effects can in fact reflect top-down control and are not always simply item-specific effects in disguise.

Sequential effects also pose a challenge to studies that manipulate list-wide PC. Specifically, there is evidence that people change their level of control on a trial-by-trial basis due to experienced conflict. This online adjustment is known as the Gratton effect

(Gratton, Coles, and Donchin, 1992), in which top-down control is enhanced following a conflicting trial. This effect has been found using Stroop, Simon, flanker, and priming tasks (Kerns et al., 2004; Sturmer, Leuthold, Soetens, Schroter & Sommer, 2002; Kunde, 2003) and presumably changes how information is processed by modifying the connections between the stimuli and goal-appropriate responding (Verguts & Notebaert, 2008). In the Stroop task, this effect manifests in reduced Stroop interference following an incongruent trial. Such trial-to-trial modulation of control makes list-based effects more difficult to interpret.

Utility of Predictive Cueing Paradigms

Clearly, it is difficult to disentangle proactive control from artifacts stemming from the stimuli themselves. Robust evidence for item-specific and sequential effects makes an argument for global proactive control tenuous at best. However, recent research by Bugg, Smallwood, and Lim (under review) tested an ingenious paradigm that allows for a more pure measure of proactive control. In four experiments, participants performed a modified version of the Stroop task in which each trial was preceded by an informative cue (“MATCHING” or “CONFLICTING”) or an uninformative cue (a string of Xs). In the first two experiments, informative cues were 100% valid. Also, different cue-stimulus intervals (CSIs) were either intermixed (Experiment 1; 500, 1000, 1500, and 2000 ms) or blocked (Experiment 2; 500, 1250, and 2000 ms) throughout the experiment. Results from both experiments provided evidence for proactive control, such that RTs were significantly faster following informative ($M = 707$ ms) than uninformative ($M =$

764 ms) cues. This benefit increased across CSI, especially for incongruent trials. Although no cueing benefit was observed for incongruent stimuli at the 500 or 1250 ms CSI, a cueing benefit was shown for both congruent and incongruent stimuli at the 2000 ms CSI. This finding suggests that cues are helpful in the engagement of proactive control, but it takes longer to prepare for overcoming an incongruent stimulus than to prepare for a congruent stimulus. However, Bugg et al. found no cueing benefit when cue validity was reduced to 75% (Experiment 3), but still found a cost in terms of longer RTs for incongruent stimuli following a “MATCHING” cue versus the neutral cue. Bugg et al. suggested that participants may use cues to engage proactive control only when such effort seems warranted. The higher the cue validity, the more likely a participant is to use it as an indicator of likely conflict. This logic is in line with that of Braver and colleagues (2007), who propose that engagement of proactive control is more likely when predictive, reliable cues are available.

However, it is possible that the cueing effects observed by Bugg and colleagues were not solely the result of using or not using proactive control. Since the cues were 100% valid, participants could be certain that any stimuli following a “MATCHING” cue would be congruent and that any stimuli following a “CONFLICTING” cue would be incongruent. Rather than exerting or relaxing proactive control, participants could have been switching to word-reading when they encountered a “MATCHING” cue, and back to the original task goal to name the color when they encountered a “CONFLICTING” cue. Slightly reducing the validity of the cues might remedy this issue, as will be discussed shortly.

Nonetheless, providing participants with reliable cues may be an effective way of inducing proactive control that is free of the confounds mentioned earlier in this paper. Specifically, researchers can provide pre-stimulus cues predicting upcoming congruence or incongruence on a trial-by-trial basis while maintaining item-specific congruence at 50%. Moreover, because the cues are randomized, specific incongruent or congruent cues are equally likely to follow previously congruent or incongruent trials. Thus, participants should be more likely to use proactive control to manage conflict, rather than relying on reactive strategies such as learning item-specific contingencies or color-word mapping (ISPC effects), and adjusting their level of control only after encountering an incongruent stimulus (sequential effects). However, the research by Bugg and colleagues does not speak to how individual differences might affect use of pre-stimulus cues. One such individual difference cited by Braver et al. (2007) was WMC. Based on previous research, I expect that participants with higher WMC are more adept at using cues to engage proactive control than are lower WMC participants.

Current Research

The aim of the current research is to extend the findings of Bugg and colleagues in two ways: By deciphering whether cues signaling conflict enhance performance by inducing suppression of the word dimension, and examining WMC as a possible parameter for cue usage within the Stroop task. To these ends, the following experiments used a modified Stroop paradigm similar to that used by Bugg and colleagues (2013). Participants saw a cue before each trial signaling that the upcoming stimulus would be

EASY (congruent) or HARD (incongruent). In contrast to Bugg et al., I used cues that were 80% valid rather than 100% valid. Using an 80% validity should allow for tightening or relaxing proactive control without changing the task from color naming to word-reading.

I predicted that participants would be able to adjust their degree of top-down control accordingly by strengthening or relaxing their goal of word-reading following HARD and EASY cues, respectively. Encountering a HARD cue before an incongruent trial should lead to increased proactive control which should attenuate conflict from the irrelevant word dimension, leading to a smaller Stroop effect (the difference in RT or errors between congruent and incongruent trials). Also, because HARD cues should trigger proactive control, and thus early suppression of the word dimension, participants should exhibit lower recall and recognition for neutral words that follow a HARD cue than for neutral words that follow an EASY cue. Finally, because high-WMC individuals are more likely to engage proactive control, these effects should be greater for high-WMC participants than low-WMC participants.

Experiment 1

The following experiment was conducted to establish the efficacy of the Stroop cueing paradigm. Specifically, I wanted to test the prediction that participants flexibly engage proactive control on a trial-to-trial basis, depending on the cue given. I predicted that the presentation of a HARD cue would lead to reduced Stroop interference on the following trial. If the engagement of proactive control is influenced by one's level of

WMC, this effect should be stronger for those with high WMC than those with low WMC.

I was also interested in how participants would manage the conflict of incongruent stimuli following a HARD cue. A recent study by Shipstead and Broadway (2013) demonstrated that participants with high WMC not only exhibited better performance during the Stroop task than lower WMC participants, but also had lower levels of later recognition for neutral words presented during the task. Shipstead and Broadway suggested that those with high WMC were more successful during the Stroop task because they were suppressing the word dimension as the stimulus appeared, making it less likely that they would recognize the word later. If proactive control leads to word-suppression, I expected participants to demonstrate reduced recall and recognition for neutral words presented after a HARD, but not an EASY, cue. Further, following Shipstead and Broadway, I expected low-WMC participants to have a lower ability to suppress word-reading and therefore demonstrate overall higher levels of recall and recognition than high-WMC participants.

METHOD

Participants and Design

Seventy-five undergraduate students enrolled in an introductory psychology class participated in the study for partial fulfillment of course credit. All participants were native speakers of English and had normal or corrected-to-normal color vision. I excluded data from 22 participants due to high error rates (> 20%) on the Stroop task ($N = 1$), high math errors (>15%) on the Ospan ($N = 7$), or missing data on the Ospan ($N = 4$), Stroop ($N = 7$), or recognition task ($N = 3$), yielding a final sample size of 53 participants.

Experiment 1 utilized a 2 (Cue: EASY/HARD) x 2 (Congruency: congruent/incongruent) x WMC mixed-subjects design, with Cue and Congruency manipulated within-subjects and WMC measured continuously between-subjects. To assess performance, I calculated mean RTs and error rates which served as my primary dependent measures. Recall and recognition of neutral items served as secondary dependent measures to examine the possibility of early suppression when participants expected conflict.

Stimuli

The Stroop cueing task used the four color words “red,” “blue,” “green,” and “yellow.” I also selected a total of 48 neutral words to examine the possibility of early suppression following HARD cues. Using the English Lexicon

Project website (Balota et al., 2007), I matched the neutral words to each of the four color words in dimensions of length, log-transformed HAL frequency, and orthographic neighborhood. Half of these words appeared at random during the Stroop task and the other half served as distracter items on the later recognition task. Of the 24 neutral words that appeared in the Stroop task, half were preceded by an EASY cue and half were preceded by a HARD cue. I created four lists to counterbalance the neutral words across the EASY and HARD cueing conditions and to allow them to serve equally often as a critical word (presented during the Stroop task) and as a distracter word (on the recognition task) across participants.

The cueing task consisted of 240 trials. Seventy-two of these served as critical trials, split evenly among congruent, incongruent, and neutral conditions, that I used in later analyses. The remaining 168 items were included as fillers to establish cue validity, ensuring that 80% of the trials following an EASY cue were congruent, and that 80% of the trials following a HARD cue were incongruent. Half of the filler items were always congruent and preceded by an EASY cue while the other half were always incongruent and preceded by a HARD cue. The overall proportion congruency (PC) of the list was .50.

Automated Ospan Task

The automated Ospan is a modified version of the traditional Ospan task (Unsworth, Heitz, Shrock, & Engle, 2005). Rather than the experimenter recording each participant's performance and proceeding to each new trial using the key pad, the

participant is able to respond and advance through trials using the computer mouse. The participants began by reading instructions on the screen, which informed them that the task involved solving arithmetic problems and memorizing letter sequences. Three sets of practice trials followed the instructions: Practice with solving the arithmetic problems, practice with letter memorization, and practice doing both.

Upon completion of the practice phase, participants moved on to the task itself. The experimenter instructed them to respond as quickly and accurately as possible with both the arithmetic and memorization portions of the task, and that they must keep their math accuracy at or above 85%. This score appeared in red font in the upper right corner of the screen so participants could monitor their performance. Each trial began with the presentation of an arithmetic problem in the center of the screen, such as “ $(2 \times 1) + 1$.” Following the problem was a blank screen and then the presentation of a number such as “4” with two buttons labeled “TRUE” and “FALSE” below it. If the presented number was the correct solution to the problem, the participants clicked “TRUE;” if it was an incorrect solution, they clicked “FALSE.” After providing a response, an upper-case letter appeared for participants to hold in memory. After 3 - 7 of these sequences, participants saw a screen with an array of 12 letters, some of which the participants had seen during the previous trials. The participants then clicked the boxes next to the letters they saw in the correct order. Points were only obtained if *all* the correct letters were selected *in the exact order* that they appeared. Otherwise, the score was zero. Scores on the Ospan reflect the total number of letters remembered in the correct order and range from 0 - 75.

Stroop Cueing Task

This task was similar to the classic Stroop task, with the critical addition of cueing. The experimenter told participants that each trial would be preceded by one of two cues, EASY or HARD, and that these cues would indicate the potential congruency of the upcoming stimulus. For stimuli following an EASY cue, there was an 80% chance that the word and color would match (congruent). For stimuli following a HARD cue, there was an 80% chance that the word and color would *not* match (incongruent). The experimenter strongly encouraged participants to use these cues to maximize their performance during the task.

Procedure

All participants provided informed consent and were tested individually. Upon arrival to the laboratory, I randomly assigned participants to one of four Stroop lists. Participants sat approximately 60cm from a computer monitor. A program developed for E-Prime software (Schneider, Eschman, & Zuccolotto, 2002) presented the stimuli and a microphone captured response latencies. The experimenter coded the responses using a keypad labeled with colored stickers corresponding to the four color words. Task order was counterbalanced among participants, such that half completed the Ospan first and half completed the Stroop task first.

At the beginning of the Stroop task, instructions appeared on the screen and were paraphrased by the experimenter. The experimenter informed participants that they would see a series of words presented one at a time and that the goal of the task was to

name the word's font color, not the word itself. The experimenter also told them that a predictive cue would precede each trial and that they should use the cue to maximize their performance. Participants responded to stimuli vocally using the microphone.

Participants completed 20 practice trials to familiarize themselves with the procedure. These trials were split equally between EASY and HARD cueing conditions and consisted of 9 congruent, 9 incongruent, and 2 neutral trials. The Stroop task itself consisted of 240 trials with self-paced rest breaks after every 50 trials. Participants first saw an EASY or HARD cue for 500ms, followed by a blank screen for 1500ms. The stimulus word then appeared and participants had 3000ms to make a response. The experimenter coded each response by pressing a key with a colored sticker corresponding to the participant's response.

Following the Stroop task, participants saw instructions to recall as many of the neutral words as they could within two minutes. The experimenter recorded and scored the responses until the two minute limit was up or the participant was unable to recall any more words. Finally, participants completed an old/new recognition task. They saw a series of 48 words, half of which had been presented during the Stroop task as neutral items. Participants responded "old" if they remembered seeing the word during the Stroop task and to respond "new" if they had not seen it. Upon completion of the four tasks, the experimenter debriefed, thanked, and dismissed the participant.

RESULTS

To correct for microphone errors, I omitted RTs faster than 50 ms and slower than 2000 ms from analysis, eliminating 0.5 % of trials. I considered only correct responses for RT analyses. For each participant, I calculated a separate mean and standard deviation for congruent, incongruent, and neutral trials in each cueing condition. A nonrecursive outlier removal procedure suggested by Van Selst and Jolicoeur (1994) removed 2.9% of correct RTs for participants. Arithmetic means based on individual participants' trimmed RTs and errors are presented in Table 1. Unless otherwise specified, all effects reported as significant are associated with a two-tailed p -value of .05 or less.

Table 1: Mean RT and error rates following EASY and HARD cues in Experiment 1.

	Cueing			
	EASY		HARD	
	RT (ms)	Errors (%)	RT (ms)	Errors (%)
Incongruent	776	3.5	755	2.8
Congruent	624	0.5	638	0.5
<i>Stroop Effect</i>	<i>153*</i>	<i>3.0*</i>	<i>118*</i>	<i>2.3*</i>

Note: * = $p < .05$.

RTs and Error Rates

I first examined RTs using a 2 x 2 repeated-measures analyses of variance (ANOVA) with Cue (EASY vs. HARD) and Congruency (Incongruent vs. Congruent) as within-subjects variables. I found a significant Stroop effect, such that participants responded more quickly to congruent ($M = 631$ ms, $SE = 15$) than incongruent stimuli (M

= 766 ms, $SE = 17$), $F(1, 52) = 223.470$, $MSE = 4,324$, $\eta_p^2 = .811$. As predicted, there was a Cue x Congruency interaction, $F(1, 52) = 11.278$, $MSE = 1,450$, $\eta_p^2 = .178$.

Participants demonstrated smaller Stroop effects for stimuli following a HARD cue ($M = 118$ ms) than an EASY cue ($M = 153$ ms).

I next examined error rates using a 2 x 2 repeated-measures ANOVA with Cue and Congruency serving as within-subjects variables. The Stroop effect found in RTs was mirrored in the error data, $F(1, 52) = 7.716$, $MSE = 50$, $\eta_p^2 = .129$, such that participants committed more errors for incongruent ($M = 3.1\%$, $SE = .009$) than congruent stimuli ($M = 0.5\%$, $SE = .002$). No other effects or interactions approached significance ($ps > .333$).

Recall and Recognition

For each participant, I calculated a corrected recognition score for neutral words that followed EASY and HARD cues. This score was calculated by subtracting the number of false alarms (incorrectly classifying a new word as “old”) from the number of hits (correctly classifying a previously presented word as “old”). Recognition and recall scores are shown in Table 2. Mean corrected recognition and recall were analyzed using two paired-samples t-tests. Neither recognition nor recall differed as a function of Cue ($ts < 1$).

Table 2: Mean proportion of recall, hits, false alarms, and corrected recognition score for neutral words following EASY and HARD cues in Experiment 1.

	Proportion Recall	Rec. Hits	False Alarms	Corrected Rec.
EASY	0.04	0.55	0.23	0.32
HARD	0.04	0.55	0.23	0.32

Working Memory Capacity

Scores on the Ospan were negatively skewed and ranged from 0 – 71 ($M = 44$).

For illustrative purposes, I performed a quartile split, separating participants into high and low span groups based on their Ospan score (the top and bottom quarter of the distribution, respectively). Arithmetic means based on individual participants' trimmed RTs and errors are presented in Table 3.

Table 3: Mean reaction times and errors as a function of cue and WMC in Experiment 1.

	Cueing			
	<u>EASY</u>		<u>HARD</u>	
	RT (ms)	Errors (%)	RT (ms)	Errors (%)
High Span				
(N = 13)				
Incongruent	719	4.5	700	3.2
Congruent	571	0.6	592	0.0
<i>Stroop Effect</i>	147	2.9	108	3.2
Low Span				
(N = 13)				
Incongruent	857	3.2	834	1.3
Congruent	681	0.6	681	0.0
<i>Stroop Effect</i>	176	2.6	153	1.3

To assess the role of WMC on the ability to use proactive control, I analyzed RTs and error rates using two mixed-model ANOVAs with Cue and Congruency measured as

within-subjects variables and Ospan score measured as a continuous between-subjects variable. I also examined recall and recognition using two mixed-model ANOVAs with Cue as a within-subjects variable and Ospan score as a continuous variable. Mean RTs and error rates as a function of WMC, Cue, and Stroop are shown in Table 3, whereas recall and corrected recognition can be found in Table 4. In terms of RTs, a main effect of WMC was found, $F(1, 51) = 5.702$, $MSE = 42,642$, $\eta_p^2 = .128$, such that participants with higher Ospan scores responded more quickly overall than those with lower Ospan scores. Also, high spans tended to show less Stroop interference in RTs than low spans, but this Congruence x WMC interaction was only marginally significant, $F(1, 51) = 3.849$, $MSE = 4,099$, $p = .055$, $\eta_p^2 = .070$. There were no effects or interactions with WMC on error rates (all $ps > .25$). Further, WMC did not seem to affect overall recall as I found no main effect or interaction with Cue. However, pair-wise comparisons showed that whereas recall for low spans did not differ as a function of cue ($p = .654$, $\eta_p^2 = .004$), recall for high spans was significantly higher following an EASY cue than a HARD cue ($p = .048$, $\eta_p^2 = .076$). Corrected recognition was not affected by WMC or Cue ($ps > .237$).

Table 4: Mean proportion of recall, hits, false alarms, and corrected recognition score for neutral words following EASY and HARD cues as a function of WMC in Experiment 1.

	PR	Hits	FA	CR
High Span (N=13)				
EASY	0.05	0.58	0.28	0.30
HARD	0.02	0.56	0.28	0.28
Low Span (N=13)				
EASY	0.05	0.55	0.17	0.38
HARD	0.04	0.55	0.17	0.38

Note: PR = Proportion of recall, FA = False Alarm, CR= Corrected Recognition

DISCUSSION

My first experiment examined whether pre-stimulus cueing could induce proactive control and enhance the ability to overcome conflicting information. Specifically, I predicted smaller Stroop effects in RTs and errors following a HARD cue, and that this pattern would be especially true for participants with higher WMC. Following Shipstead and Broadway (2013), I further predicted that recall and recognition for neutral items would be poorest following a HARD cue due to early suppression of the word dimension, and poorer for high spans than low spans.

A significant Cue x Congruency interaction in RTs confirmed my first hypothesis. The data revealed smaller Stroop effects following a HARD cue than an EASY cue. This suggests that participants successfully used the cue to engage proactive control and thus overcome the conflicting information. Thus, using pre-stimulus cues may be a promising way of isolating and examining the use of proactive forms of control.

I also predicted that, due to greater executive control, high span participants would be more adept at utilizing the HARD cues than their low span peers. The absence of a Cue x Stroop x WMC interaction indicated that this was not the case. Whereas high spans tended to respond more quickly overall and tended to experience less Stroop interference, they did not seem to reap any added benefits from using HARD cues.

Finally, I had predicted that participants would demonstrate poorer recall and recognition for neutral items following a HARD cue than an EASY cue due to early suppression of word-reading. While the omnibus test showed no effects of or interactions with WMC, high spans showed significantly lower recall for neutral words that followed

a HARD cue than words that followed an EASY cue. Also, levels of corrected recognition for high and low spans numerically, but not significantly, replicated Shipstead and Broadway (2013). However, there appeared to be a floor effect in which overall recall and recognition were very low (0.9 and 0.3 words out of 24, respectively). This brings up a potential limitation of this experiment. The reader may recall that I used 24 neutral items presented once each during the task. It is entirely possible that such conditions would not allow participants to retain the items in memory regardless of whether they were engaging in word suppression.

In addition to the floor effects already mentioned, it is possible that not all participants used the cues to enhance their performance. Indeed, some participants later admitted to ignoring the cues during the task. If participants were not attending to the cues, they would not use them to engage proactive control. Thus, it might prove beneficial to use stronger encouragement to attend to the cues during the task.

Experiment 2

Along with sophisticated technologies like EEG and fMRI, researchers have turned to pupillometry as a physiological measure of cognitive activity (Kahneman & Peavler, 1969). Pupillometry involves the use of an eye-tracking apparatus to measure changes in pupil diameter. Extensive research has identified changes in pupil diameter as reliable indicators of many states including sexual preference, aversion, interest, mental effort, and anticipation (see Goldwater, 1972 for a review). Most relevant to the current experiment is mental effort. Research demonstrates that pupil dilation is greatest during

high cognitive load (Beatty, 1982; Peavler, 1974; Granholm, Asarnow, Sarkin, & Dykes, 1996).

The purpose of the following experiment was to further examine proactive control using pupillometry, as well as addressing two limitations of Experiment 1. First, recall for neutral words was quite low ($M = 0.9$ out of 24). Due to the possibility of floor effects, I reduced the number of neutral words in Experiment 2 from 24 to 8. Further, each of these neutral words was shown four times (once in each color), instead of just once, during the Stroop task. Second, some participants admitted to ignoring the cues during the Stroop task. Thus, I added a point value to each of the cues, such that EASY trials were worth 1 point and HARD trials were worth 10 points if the participant responded correctly. This provided participants with a more tangible incentive to attend to the cues.

For Experiment 2, I had four predictions: first, that participants would use HARD (but not EASY) cues to prepare for the upcoming stimulus by suppressing the word dimension; second, that this suppression following HARD cues would lead to reduced recall of neutral words; third, that pupil dilation would be greater following a HARD than an EASY cue due to increased cognitive effort; and fourth, that these patterns would be stronger for high spans than low spans.

METHOD

Participants and Design

Ninety-seven undergraduate students participated for partial fulfillment of course credit. I omitted data from 13 participants due to high math errors (>15%) on the Ospan ($N = 3$), missing Stroop data ($N = 4$), or technical difficulties with the eye-tracker ($N = 5$). Also, one participant was omitted as she did not meet the criteria of being a young adult (< 25 years old). I conducted analyses on a final sample of 84 participants.

Experiment 2 utilized a 2 (Cue) x 2 (Congruency) x WMC mixed-subjects design with Cue and Congruency manipulated within-subjects and WMC measured continuously between-subjects. In addition to mean RT, errors, and recall, I included mean pupil dilation as a dependent variable. I calculated mean pupil diameter for each participant during the interval between the cue and the stimulus by averaging the mean diameter for both eyes.

Stimuli

I used a list of 248 trials, divided into three blocks, for this experiment. Forty eight served as critical trials that were balanced among cue (EASY vs. HARD), congruency (congruent vs. incongruent), and color (red, blue, green, or yellow) dimensions. I included one hundred sixty-eight trials as fillers to ensure “easy” and “hard” cues were predictive 80% of the time and to yield a proportion congruency of .50. The remaining 32 trials were neutral trials, consisting of eight words matched to “red,”

“blue,” “green,” and “yellow” in dimensions of length, log-transformed HAL frequency, and orthographic neighborhood. Each word was presented four times, once in each color.

Apparatus

I collected pupil data during the Stroop task using a RED infrared eye tracker (SensoMotoric Instruments) running iView X software (Version 2.8), sampling at 120 Hz and at spatial resolution 0.03 RMS. Calibration of gaze direction was conducted for each participant at the beginning of the Stroop task. Blinks were corrected for using linear interpolation. To assess engagement of proactive control, I recorded pupil activity during the interval between cue and stimulus (1500ms).

Procedure

The procedure for Experiment 2 was identical to that of Experiment 1 with three exceptions. First, due to the age of the laboratory computers, I was unable to connect the S-R box used in Experiment 1. Thus, participants responded to stimuli using a key press rather than a microphone. I did not see this as a serious limitation, because previous researchers have found Stroop effects in both RTs and errors using both vocal and key press responses (Logan & Zbrodoff, 1998; see MacLeod, 1991, for a review). Stickers were labeled (“R”, “G”, “B”, and “Y”) and placed on four keys (1, 2, 3, and 5, respectively). The experimenter instructed participants to press the key that corresponded to the color of the stimulus. Second, the recognition task was omitted. Lastly, I included a point value to the right of each cue, so participants either saw “easy 01” or “hard 10”

preceding each trial. I intended these point values to increase participants' incentive to use the cues to maximize their performance. Based on previous research, I did not expect this change to unduly influence my results. Although incentive tends to increase dilation overall (Kahneman & Peavler, 1969; Peavler, 1974; Chiew & Braver, 2013), it does not in itself explain changes in dilation during cognitive tasks (Heitz, Shrock, Payne, & Engle, 2008). Thus, any difference in dilation following EASY and HARD cues is unlikely to be due to incentive stemming from the point values. The experimenter told participants to use the cues to get the highest score possible by the end of the task. Participants saw their current point total during each of the three self-paced rest-breaks.

RESULTS

To correct for abnormally long and short RTs, I omitted those faster than 200ms and slower than 1500ms from analysis, eliminating 6.1% of trials. As with Experiment 1, I considered only correct trials for RT analysis. For each participant, a separate mean and standard deviation was calculated for congruent, incongruent, and neutral trials in each cueing condition. Nonrecursive outlier removal (Van Selst & Jolicoeur, 1994) removed 1.6 % of correct RTs for participants. Arithmetic means based on individual participants' trimmed RTs and errors are presented in Table 5.

Table 5: Mean reaction times and error rates following EASY and HARD cues in Experiment 2.

	“Easy”		“Hard”	
	RT (ms)	Errors (%)	RT (ms)	Errors (%)
Incongruent	913	4.1	894	3.2
Congruent	766	0.6	774	1.4
<i>Stroop Effect</i>	<i>147***</i>	<i>3.5***</i>	<i>120***</i>	<i>1.8*</i>

Note: * = $p < .05$; ** = $p \leq .01$; *** = $p < .001$.

RTs and Error Rates

I first analyzed RTs using a 2 x 2 repeated-measures ANOVA with Cue (EASY/HARD) and Congruency (congruent/incongruent) serving as within-subject variables. A strong Stroop effect was shown in the form of a main effect of Congruency, $F(1, 83) = 175.838$, $MSE = 8,461$, $\eta_p^2 = .679$. Participants responded more quickly to congruent ($M = 770$ ms, $SE = 14$ ms) than incongruent stimuli ($M = 903$ ms, $SE = 15$ ms). Importantly, the Cue x Congruency interaction obtained in Experiment 1 was replicated

in Experiment 2, such that participants showed a smaller Stroop effect for stimuli following a HARD cue (120 ms) than an EASY cue (147 ms), $F(1, 83) = 6.439$, $MSE = 2,397$, $\eta_p^2 = .072$. Pair-wise comparisons showed that this pattern was driven by attenuated RTs for incongruent stimuli following HARD cues ($M = 894$ ms, $SE = 18$ ms), relative to EASY cues ($M = 913$ ms, $SE = 17$), $t(83) = 2.053$, $p = .043$. RTs to congruent stimuli did not differ between EASY ($M = 766$ ms, $SE = 14$ ms) and HARD ($M = 774$ ms, $SE = 14$ ms) cue conditions, $t(83) = -1.288$, $p = .201$.

Next, I analyzed errors using a 2 (Cue) x 2 (Congruency) repeated-measures ANOVA. As with RTs, I found a Stroop effect, such that participants committed more errors when responding to incongruent stimuli ($M = 3.7$, $SE = 0.6$) than congruent stimuli ($M = 1.0$, $SE = 0.2$), $F(1, 83) = 20.623$, $MSE = 30$, $\eta_p^2 = .199$. Mirroring the pattern shown in RTs was a marginal Cue x Congruence interaction in which participants showed a smaller Stroop effect in terms of errors for stimuli following HARD cues ($M = 1.8$, $SE = 0.7$) than EASY cues ($M = 3.7$, $SE = 0.9$), $F(1, 83) = 3.031$, $MSE = 20$, $\eta_p^2 = .035$.

Recall

To examine recall of the eight neutral words as a function of cue, I conducted a paired-samples t-test. Although participants tended to recall a higher proportion of words following an EASY cue ($M = 0.13$) than a HARD cue ($M = 0.11$), this trend did not reach significance, $t(83) = 1.382$, $p = .171$.

Pupil Dilation

To assess cognitive effort as a function of cue, separate pupil dilation means were calculated for each participant following EASY and HARD cues. I examined the difference in dilation following EASY and HARD cues using a paired-samples t-test. Although pupils were indeed more dilated during the fixation interval following an EASY cue ($M = 5.48$ mm, $SE = 0.09$ mm) than a HARD cue ($M = 5.46$ mm, $SE = 0.09$ mm), this trend did not reach significance, $t(83) = 1.593, p = .115$.

Working Memory Capacity

To examine the effect WMC might have on the ability to use predictive cues, I analyzed both RTs and errors using a 2 (Cue) x 2 (Congruence) mixed-model ANOVA with participants' Ospan score as a continuous between-subjects variable. Ospan scores were normally distributed and ranged from 0 – 75 ($M = 39, sd = 17$). As with Experiment 1, participants were split into high and low span groups based on upper and lower quartiles for illustrative purposes. Arithmetic means based on individual participants' trimmed RTs and errors are presented in Table 6.

Table 6: Mean reaction times and error rates for stimuli following EASY and HARD cues by WMC group in Experiment 2.

	Cue			
	EASY		HARD	
	RT (ms)	Errors (%)	RT (ms)	Errors (%)
High Span				
(N = 21)				
Incongruent	879	6.4	869	2.9
Congruent	759	1.4	775	1.2
<i>Stroop Effect</i>	121*	5.0*	94*	1.7
Low Span				
(N = 22)				
Incongruent	894	4.5	869	3.8
Congruent	758	0.4	760	1.1
<i>Stroop Effect</i>	136*	4.1*	108*	2.7°

Note: * = $p < .05$, ° = $p < .10$.

Analysis of RTs and errors using the full range of Ospan revealed no effects or interactions with WMC (all $ps > .250$). In terms of errors, high span participants tended to have higher error rates ($M = 2.9$, $SE = 0.6$) than low span participants ($M = 2.4$, $SE = 0.6$), but this main effect of WMC was not significant, $F(1, 82) = 1.341$, $p = .250$, $\eta_p^2 = .016$. While Cue did not interact with WMC, I found that high spans, and not low spans, showed a marginal effect of Cue. Their error rate was 3.9% ($SE = 0.8$) for stimuli following EASY cues, but dropped to 2.0% ($SE = 0.8$) following HARD cues. Pair-wise comparisons showed that this difference was only marginally significant ($p = .080$, $\eta_p^2 = .037$). Low spans had similar error rates for EASY ($M = 2.4$, $SE = 0.8$) and HARD ($M = 2.5$, $SE = 0.8$) cueing conditions ($p = .975$, $\eta_p^2 = .000$). Further, high spans showed a Cue x Congruency interaction in errors, such that these individuals demonstrated a significant

Stroop effect following an EASY cue ($p = .005$, $\eta_p^2 = .094$), but not a HARD cue ($p = .266$, $\eta_p^2 = .015$).

To examine the effect of WMC on recall and pupil dilation, a mixed-model ANOVA was conducted with Cue as a within-subjects variable and Ospan score as a continuous variable. There was no effect of WMC on recall, nor did WMC interact with Cue. Recall by cue and WMC are displayed in Table 7. The ANOVA showed no main effects of or interactions with WMC. Pair-wise comparisons showed that, whereas pupil dilations for low spans changed little as a function of cue (EASY: 5.65 mm, HARD: 5.64 mm), high spans showed marginally greater dilation following EASY cues ($M = 5.44$ mm, $SE = .174$ mm) than HARD cues ($M = 5.40$ mm, $SE = .172$ mm), $p = .062$, $\eta_p^2 = .042$. Pupil dilations by cue and WMC group are displayed in Figure 1.

Table 7: Mean Proportion of Recall as a function of Cue and WMC in Experiment 2.

	Cue	
	EASY	HARD
High Span	0.12	0.11
Low Span	0.11	0.10

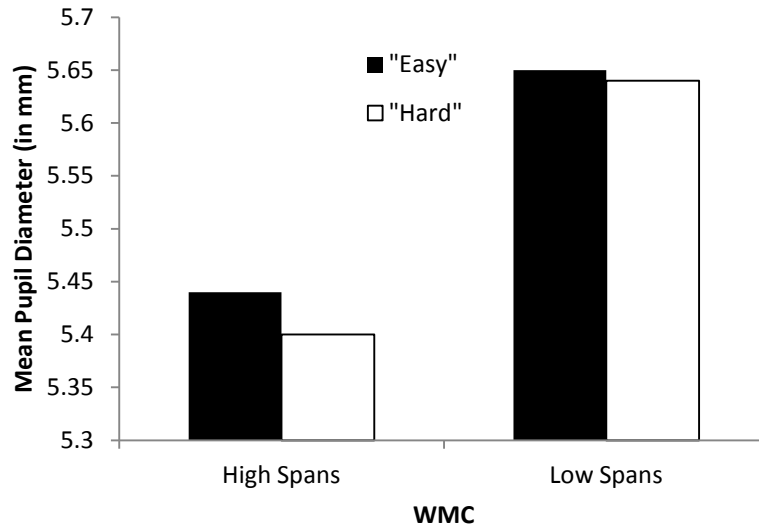


Figure 1: Mean pupil dilation following EASY and HARD cues by WMC. High spans are participants at the top quarter of the Ospan distribution and low spans are participants at the bottom quarter.

Pupil Dilation Time Course Analysis

In addition to previous analyses, I wanted to examine how cognitive effort may have changed over the 1500ms post-cue fixation. To this end, I calculated separate mean pupil dilations for each participant during six intervals (bins): 0, 300, 600, 900, 1200, and 1500ms after cue presentation. For each participant, I averaged these dilations across all 248 trials. To analyze potential cue differences in dilation across bins, I conducted a mixed-model ANOVA with Bin (0/300/600/900/1200/1500) and Cue (EASY/HARD) measured within-subjects and WMC (high span/low span) measured between-subjects.

A main effect of Bin indicated that dilation increased substantially over the course of the delay, $F(1, 112) = 140.115$, $MSE = 500$, $\eta_p^2 = .628$. Mean pupil diameter increased from 5.32 mm at baseline to 5.58 mm immediately before presentation of the

stimulus. This gradual increase from cue to stimulus suggests that participants were mentally preparing to make a response. However, the absence of a Bin x Cue interaction indicated that increase in dilation over the delay was not dependent on Cue ($F < 1$). The pattern of increasing dilation over the delay following EASY and HARD cues is shown in Figure 2.

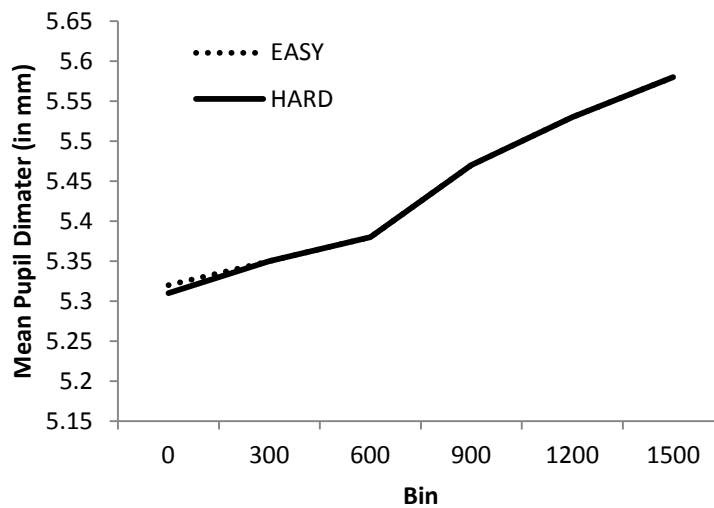


Figure 2: Mean pupil dilation across fixation delay following EASY and HARD cues.

I next examined the effect of WMC on dilation over the delay. No main effect or interaction with WMC was found. Low spans tended to show greater increases in dilation (.32 mm) over the fixation delay than high spans (.24 mm), but this Bin x WMC interaction did not reach significance, $F(3, 109) = 1.800$, $MSE = 490$, $p = .157$, $\eta_p^2 = .043$. Patterns of dilation over the fixation by WMC group and Cue are shown in Figure 3.

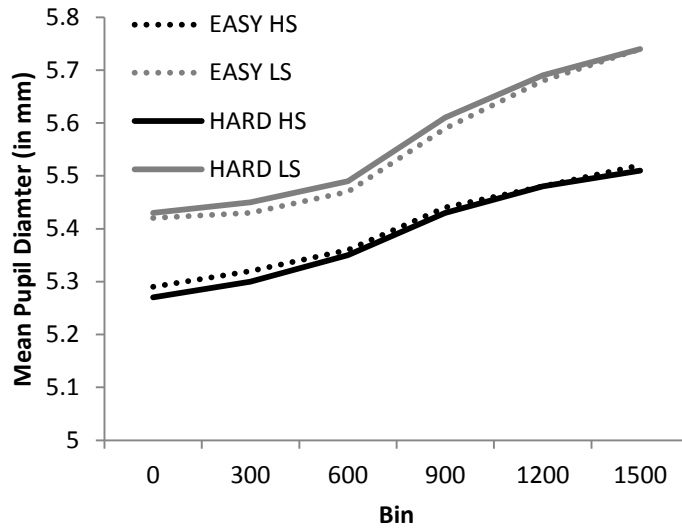


Figure 3: Mean pupil dilation over 1500 ms fixation interval by WMC group. HS = high span, LS = low span.

Finally, I wanted to see if change in pupil dilation over the fixation interval was predictive of subsequent accuracy. I calculated difference scores for each participant by subtracting baseline dilation from dilation in the 1500 ms bin across both cueing conditions for accurate and inaccurate trials. To examine the effect of dilation increase on trial accuracy, I used a 2 (Cue: EASY/HARD) x 2 (Accuracy: accurate/inaccurate) repeated-measures ANOVA. I found only a marginal main effect of accuracy, with participants showing a greater increase in dilation preceding an accurate trial ($M = .27$ mm, $SE = 0.04$ mm) than an inaccurate trial ($M = .21$ mm, $SE = 0.03$ mm), $F(1, 56) = 3.822$, $MSE = 1,910$, $p = .056$, $\eta_p^2 = .064$.

Next, I examined whether WMC affected change in dilation preceding accurate and inaccurate trials using a 2 (Cue) x 2 (Accuracy) x 2 (WMC group: high span/low span) mixed-model ANOVA with Cue and Accuracy measured within-subjects and

WMC measured between-subjects. I found only a marginal Accuracy x WMC interaction, $F(1, 29) = 2.132$, $MSE = 750$, $p = .155$, $\eta_p^2 = .068$. Pair-wise comparisons showed that there was an effect of Accuracy for high spans, who showed a marginally greater increase in dilation over the fixation interval preceding an accurate response ($M = 0.30$ mm, $SE = 0.05$ mm) than an inaccurate response ($M = 0.16$ mm, $SE = 0.10$ mm), $p = .066$, $\eta_p^2 = .112$. Dilation increase for low spans did not differ as a function of subsequent accuracy ($p = .945$, $\eta_p^2 = .000$). Change in dilation over the fixation interval by Cue, Accuracy, and WMC are illustrated in Table 8.

Table 8: Mean Change in Pupil Dilation as a function of Cue, Accuracy, and WMC in Experiment 2.

	Cue	
	EASY (SE)	HARD (SE)
High Span		
Accurate	0.31 (0.05)	0.30 (0.05)
Inaccurate	0.16 (0.11)	0.17 (0.09)
Low Span		
Accurate	0.30 (0.05)	0.30 (0.05)
Inaccurate	0.37 (0.10)	0.24 (0.08)

Note: Dilation is reported in mm.

DISCUSSION

The purpose of Experiment 2 was to replicate my earlier results and to include pupil dilation as a dependent measure to see if this index of cognitive effort would differ as a function of cue. I had four predictions: first, I predicted that participants would exert greater proactive control when presented with a HARD cue and that this would manifest in reduced RTs and error rates for stimuli following a HARD cue; second, I predicted that recall of neutral words would be lower following a HARD cue than an EASY cue; third, I predicted that pupil dilation would be greater following a HARD cue than an EASY cue, signaling the engagement of proactive control; and fourth, I predicted that these patterns would be stronger for high span than low span participants as they are believed to have greater capacity for executive control and should thus be better at using the HARD cues to engage proactive control.

Replication of the Cue x Congruence interaction provided support for my first prediction. Participants showed a smaller Stroop effect in both RTs and errors for stimuli following HARD cues than EASY cues. Further, the benefit of the HARD cue in RTs was seen for incongruent stimuli following a HARD cue, but not for congruent stimuli. This suggests that the presentation of a predictive HARD cue allows participants to prepare for, and subsequently minimize, conflicting information.

Although I reduced the number of neutral words, recall in Experiment 2 changed little from Experiment 1. Participants tended to recall fewer words following a HARD cue than an EASY cue, but this difference was not significant. Also, overall recall was still very low. When prompted for recall, many participants seemed embarrassed by how

few words they were able to recall (if any) and claimed that they had been paying too much attention to the font color to notice the words that had appeared.

In terms of pupil dilation, the data actually trended in the opposite direction of what I predicted. While not quite significant, participants tended to show greater dilation following an EASY cue than a HARD cue. Based on the reduced Stroop effect in terms of both RTs and errors following a HARD cue, one would expect to find evidence of increased top-down control in the form of greater pupil dilation. On the other hand, I did find that pupil dilation increased across the fixation interval, which suggests that participants were preparing for the upcoming trial. Also, greater increases in pupil dilation were marginally greater preceding an accurate response than an inaccurate response, indicating that greater cognitive effort as indexed by pupil dilation was indeed beneficial in that it led to more accurate responding.

Finally, WMC did not seem to play a major role in ability to use cues to modulate proactive control. I predicted that high spans should show smaller Stroop effects than low spans in the HARD cueing condition. The null Cue x Congruency x WMC interaction did not support this prediction. I also expected that HARD cues would lead to lower levels of recall and greater pupil dilation in high spans compared to low spans. Neither of these predictions were confirmed by our data. First, recall did not seem to be dependent on WMC as high and low spans showed similarly low levels of recall. Further, neither of the groups showed a difference in recall between EASY and HARD cueing conditions (in contrast with Experiment 1, in which high spans showed marginally greater recall for neutral words following an EASY cue than a HARD cue). Exploring

the influence of WMC on pupil dilation, I found evidence that ran counter to my prediction. Upon closer examination of the Cue x WMC interaction, I found that whereas low spans showed no reliable difference in dilation across cueing conditions, high spans showed marginally greater dilation following an EASY cue than a HARD cue. Finally, the relation between pupil dilation increase and accuracy depended on WMC. High spans showed greater increase in pupil dilation over the fixation interval preceding an accurate response than an inaccurate response, while dilation increase for low spans did not differ as a function of accuracy.

GENERAL DISCUSSION

Importantly, the replicated Cue x Congruency interaction demonstrated the existence of proactive control. In both experiments, I found smaller Stroop effects in RTs following a HARD cue than an EASY cue and, in Experiment 2, this pattern was also found in the error data. Together, these two experiments suggest that pre-stimulus cueing may be an effective way of examining engagement of proactive control that is free of previously cited confounds like sequential effects and ISPC effects. Further, these cues do not necessarily have to be 100% valid as indicated by Bugg and colleagues (in press). The cue validity used in this study (80%) seemed reliable enough for participants to justify engaging the extra effort following a HARD cue versus an EASY cue.

In Experiment 2, the benefit of pre-stimulus cueing appeared on incongruent trials that followed a HARD cue. RTs were significantly reduced compared to incongruent trials following an EASY cue, indicating that participants were using the HARD cue to prepare for upcoming conflict. To enhance the cueing effects found in this study, future research should consider increasing list-wide PC from .50. Goldfarb and Henik (2013) demonstrated that cueing benefits were greater when list-wide PC was high (80%). When list-wide PC is lower, participants' default strategy might be to prepare for conflict, in which case they would reap little added benefit from pre-stimulus cues. However, if list-wide PC is high, conflict is less likely and cues would be seen as more useful. Using the current cueing paradigm, I would expect to see reduced recall and recognition for neutral words following HARD cues, and significantly greater pupil dilation following a HARD cue than an EASY cue.

Also, I found little evidence that WMC influences the ability to use pre-stimulus cues. In their article on proactive and reactive control, Braver and colleagues (2007) suggested that proactive control might be more feasible for those with higher WMC since “individuals possessing greater cognitive resources will be those most willing and able to adopt a proactive mode” (pg. 88). However, both high and low spans appeared to benefit from the HARD cues, showing similar patterns of reduced Stroop effects in both RTs and errors. The only evidence I found for my prediction was in Experiment 2, wherein high spans had marginally smaller error rates following a HARD cue than an EASY cue. They also showed marginally smaller Stroop effects in errors following a HARD cue than an EASY cue, while error rates for low spans remained stable across cueing conditions.

One interpretation of my null WMC effects could be that the availability of predictive cues boosts the performance of low spans in a way similar to MI lists (Engle, 2002), blurring potential differences between those of high and low WMC. While high spans do not seem to experience any additive benefits of MI lists or HARD cues, such conditions may be adequate to provide a form of cognitive scaffolding for low spans. Specifically, the presence of external cues reduces their need to maintain the task goal over time. Including an uninformative cue in the current cueing paradigm might be a way to test this possibility. If HARD cues do, in fact, provide such scaffolding to low spans, we would expect to see low spans perform worse than high spans following an uninformative cue, but equivalent to high spans following informative EASY and HARD cues. Another possibility is that perhaps conflict was not frequent enough for WMC differences to be apparent. While Braver et al. proposed that individuals with high WMC

are more likely to use proactive control, they elaborated that such individuals might do so only when task demands were sufficient to warrant such use of cognitive resources.

Perhaps utilizing a higher list-wide PC (as mentioned earlier) would lead to better contrast in use of proactive control between high and low spans. Future research could explore this possibility by looking at cueing effects of high and low spans across high-PC and low-PC lists.

Looking at recall for neutral words across both experiments, I found no reliable evidence that a HARD cue would lead to greater suppression of word-reading than an EASY cue. Recall and recognition for both cue conditions were identical in Experiment 1. Experiment 2 showed that participants were somewhat more likely to recall neutral words following an EASY cue than a HARD cue, but this difference was not significant. As I have previously stated, overall recall was low across both experiments with the majority of participants recalling only one or two words. It is likely that the instructions to “name the color and not the word” were sufficient in getting participants to block out the word dimension to the best of their abilities. This may also explain why the findings in only one of my experiments (Experiment 1) was in line with Shipstead and Broadway (2013), with high spans showing reduced recall for neutral words following a HARD cue relative to an EASY cue.

My findings in terms of pupil data were particularly intriguing. Previous research (Granholm et al., 1996; Peavler, 1974) has shown that pupil dilation is a reliable indicator of cognitive effort, with dilation increasing with task difficulty and effort expenditure. For this reason, I predicted that pupil dilation would be greater following a HARD cue

than an EASY cue, because participants would be engaging proactive control to prepare for upcoming conflict. What I found in the data was actually in the opposite direction of my prediction: Pupil dilation was marginally greater following an EASY cue than a HARD cue. While unexpected, this finding might be explained by previous research on strategic control. Gratton, Coles, and Donchin (1992) proposed that in certain tasks, preparing for conflict becomes the default strategy. Participants may only change strategies when they are expecting information that is *not* conflicting. In the case of my study, this change in strategy would have occurred after participants encountered an EASY cue. This updating of one's strategy manifested in the data of Gratton et al. in a spike of the P300 wave, an ERP component often associated with strategy updating (Donchin & Coles, 1988a). It is not unreasonable to think that the greater dilation I found following an EASY cue than a HARD cue was due to updating one's cognitive strategy. If I were to use a higher PC (as in Goldfarb & Henik, 2013), I might expect greater dilation following a HARD cue than an EASY cue, as word reading would often lead to a correct response and HARD cues would signal the need to update one's strategy to word suppression and enhanced proactive control.

While exploring the effect of WMC on pupil dilation, I found that overall, low spans showed greater pupil dilation than high spans (although this difference did not reach significance, $p = .348$). This makes sense, as low spans likely have to work harder to suppress conflict than high spans. In fact, my data seems consistent with that of Heitz and colleagues (2008). In their study, low spans showed greater pupil dilation than high spans while encoding letters in a reading span task, indicating greater effort expenditure.

Although high spans were exerting control, it likely took less effort for them to do so. Looking at the effect of cue on pupil dilation, low spans showed similar dilation following EASY and HARD cues, but high spans showed significantly greater dilation following an EASY cue than a HARD cue. It is unclear why this might be the case. Future research should examine individual differences in cognitive effort when participants are told to expect (or not expect) conflict.

Finally, examination of changes in pupil dilation over the fixation delay yielded some interesting findings. First, dilation increased across delay regardless of cue, suggesting that participants were preparing mentally for the upcoming stimulus. Also, dilation increased more during the fixation interval preceding an accurate response than an inaccurate response. Second, while high spans showed lower overall dilation, both high and low spans showed a similar pattern of increase. Interestingly, increase in dilation prior to stimulus onset was only predictive of accuracy for high spans. While both high and low spans appeared to be preparing for the upcoming stimulus, it seems that high spans are somehow more efficient in this preparation. Thus, pupil dilation might be an effective measure of proactive control for higher-WMC individuals than it is for lower-WMC individuals.

In sum, the current cueing paradigm seems like a promising way to examine use of proactive control in conflict tasks such as the Stroop. Future research should continue to examine how participants use these cues to prepare for and subsequently overcome conflict, the degree to which WMC and other individual difference variables play a role in how these cues are used, and how patterns of pupil dilation might predict accuracy and

speed. Findings from such research may shed light on how cueing could be used to benefit individuals known for an impoverished capacity for proactive control, including older adults and dementia patients. The null WMC effects in the current study indeed suggest that such cues may be widely beneficial across individuals.

REFERENCES CITED

- Balota, D. A., Yap, M. J., Cortese, M. J., Hutchison, K. A., Kessler, B., Loftis, B., Neely, J. H., Nelson, D. L., Simpson, G. B., & Treiman, R. (2007). The English Lexicon Project. *Behavior Research Methods*, *39*, 445-459.
- Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological Bulletin*, *91*(2), 276-292.
- Braver, T. S., Gray, J. R., & Burgess, G. C. (2007). Explaining the many varieties of working memory variation: Dual mechanisms of cognitive control. In A. Conway, C. Jarrold, M. J. Kane, A. Miyake, & J. N. Towse (Eds.), *Variation in working memory* (pp. 76-106). New York, NY: Oxford University Press.
- Bugg, J. M., Smallwood, A., & Lim, Y. B. (under review). The next trial will be conflicting!: The effects of explicitly pre-cueing congruency on cognitive control.
- DePisapia, N. & Braver, T. S. (2006). A model of dual control mechanisms through anterior cingulate and prefrontal cortex interactions. *Neurocomputing*, *69*, 1322-1326.
- Donchin, E. & Coles, M. G. H. (1988a). Is the P300 component a manifestation of context updating? *Behavioral and Brain Sciences*, *11*, 355-372.
- Engle, R. W. (2002). Working memory capacity as executive attention. *Current Directions in Psychological Science*, *11*(1), 19-24.
- Goldfarb, L. & Henik, A. (2013). The effect of a preceding cue on the conflict-solving mechanism. *Experimental Psychology*, *60*(5), 347-353.
- Goldwater, B. C. (1972). Psychological significance of pupillary movements. *Psychological Bulletin*, *77* (5), 340-355.
- Granholm, E., Asarnow, R. F., Sarkin, A. J., & Dykes, K. L. (1996). Pupillary responses index cognitive resource limitations. *Psychophysiology*, *33* (4), 457-461.
- Gratton, G., Coles, M. G. H., & Donchin, E. (1992). Optimizing the use of information: Strategic control of activation of responses. *Journal of Experimental Psychology: General*, *121*, 480-506.
- Heitz, R. P., Schrock, J. C., Payne, T. W., & Engle, R. W. (2008). Effects of incentive on working memory capacity: Behavioral and pupillometric data. *Psychophysiology*, *44*, 1-11.

- Hutchison, K. A. (2007). Attentional control and the relatedness proportion effect in semantic priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *33*, 645-662.
- Hutchison, K. A. (2011). The interactive effects of listwide control, item-based control, and working memory capacity on Stroop performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37* (4), 851-860.
- Kahneman, D. & Peavler, W. S. (1969). Incentive effects and pupillary changes in association learning. *Journal in Experimental Psychology*, *79* (2), 312-318.
- Kane, M. J. & Engle, R. W. (2003). Working-memory capacity and the control of attention: The contributions of goal-neglect, response competition, and task set to Stroop performance. *Journal of Experimental Psychology: General*, *132* (1), 47-70.
- Kane, M. J. & Engle, R. W. (2002). The role of prefrontal cortex in working-memory capacity, executive function, and general intelligence: An individual differences perspective. *Psychonomic Bulletin and Review*, *9*(4), 637-671.
- Kane, M. J., Bleckley, M. K., Conway, A. R. A., & Engle, R. W. (2001). A controlled attention view of WM capacity. *Journal of Experimental Psychology: General*, *130*, 169-183.
- Kerns, J. G., Cohen, J. D., MacDonald, A. W., Cho, R. Y., Stenger, V. A., & Carter, C. S. (2004, February 13). Anterior cingulate conflict monitoring and adjustments in control. *Science*, *303*, 1023-1026.
- Kunde, W. (2003). Sequential modulations of stimulus-response correspondence effects depend on awareness of response conflict. *Psychonomic Bulletin and Review*, *10*, 198-205.
- Logan, G. D. & Zbrodoff, N. J. (1979). When it helps to be misled: Facilitation effects of increasing the frequency of conflicting stimuli in a Stroop-like task. *Memory and Cognition*, *7*, 166-174.
- Logan, G. D. & Zbrodoff, N. J. (1998). Stroop-type interference: Congruity effects in color-naming with typewritten responses. *Journal of Experimental Psychology: Human Perception and Performance*, *24* (3), 978-992.
- MacDonald, A. W., Cohen, J. D., Stenger, V. A., & Carter, C. S. (2000). Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control. *Science*, *288*, 1835-1838.

- MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, *109*(2), 163-203.
- Miller, E. K. & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*, 167-202.
- Norman, D. A. & Shallice, T. (1986). Attention to attention: Willed and automatic control of behavior. In R. Davidson, R. G. Schwartz, & D. Shapiro (Eds.), *Consciousness and self-regulation: Advances in research and theory* (pp. 1-18). New York: Plenum Press.
- Peavler, W. S. (1974). Pupil size, information overload, and performance differences. *Psychophysiology*, *11* (5), 559-566.
- Perret, E. (1974). The left frontal lobe of man and the suppression of habitual responses in verbal categorical behavior. *Neuropsychologia*, *12*, 323-330.
- Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime reference guide*. Pittsburgh, Psychology Software Tools, Inc.
- Shipstead, Z. & Broadway, J. M. (2013). Individual differences in working memory capacity and the Stroop effect: Do high spans block the words? *Learning and Individual Differences*, *26*, 191-195.
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*, 643-662.
- Sturmer, B., Leuthold, H., Soetens, E., Schroter, H., & Sommer, W. (2002). Control over location-based response activation in the Simon task: Behavioral and electrophysiological evidence. *Journal of Experimental Psychology: Human Perception and Performance*, *28*, 1345-1363.
- Unsworth, N., Heitz, R. P., Shrock, J. C., & Engle, R. W. (2005). An automated version of the operation span task. *Behavior Research Methods*, *37* (3), 498-505.
- Van Selst, M. & Jolicoeur. (1994). A solution to the effect of sample size on outlier elimination. *The Quarterly Journal of Experimental Psychology*, *47A*, 631-650.
- Vendrell, P., Junque, C., Pujol, J., Jurado, M. A., Molet, J., & Grafman, J. (1995). The role of prefrontal regions in the Stroop task. *Neuropsychologia*, *33*, 341-352.
- Verguts, T. & Notebaert, W. (2008). Hebbian learning of cognitive control: Dealing with Specific and Nonspecific adaptation. *Psychological Review*, *115*(2), 518-525.