Is Semantic Activation from Print Automatic?

An Investigation Using the Psychological Refractory Period and Task Set Paradigms

by

Darcy White

A thesis

presented to the University of Waterloo

in fulfillment of the

thesis requirement for the degree of

Doctor of Philosophy

in

Psychology

Waterloo, Ontario, Canada, 2017

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner	Dr. Naseem Al-Aidroos Associate Professor
Supervisor	Dr. Derek Besner Professor
Internal Member	Dr. Evan Risko Assistant Professor
Internal Member	Dr. Jonathan Fugelsang Professor
Internal-external Member	Emmanuelle Piérard Associate Professor

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

A portion of the literature review included in Chapter 1, and all of the research described in Chapter 2, has been published in *Acta Psychologica* (White & Besner, 2016). Changes have been made to improve the flow of the dissertation.

Abstract

The view that various visual word recognition processes are automatic in the sense that they are ballistic, intention free, unconscious, and capacity free, dominates the reading literature. Though results from multiple studies contradict the automatic perspective, its prevalence continues to this day. The present experiments address the automaticity of *semantic* activation from print by exploring whether it is (a) capacity limited, (b) requires intention, and (c) whether it is subject to performance optimization. First, I examine standard and semantic Stroop effects in the context of two Psychological Refractory Period Paradigm (PRP) experiments to address the issue of whether semantic activation from print is *capacity limited*. Included in these experiments is an SOA Proportion manipulation that has been proposed to encourage a strategic adoption of either parallel or serial processing (Miller, Ulrich, & Rolke, 2009). The results of the PRP experiments support the conclusion that semantic activation is, contrary to the wide spread view in the literature, capacity limited, and provide evidence that performance optimization plays a role in this paradigm, despite it not interacting with Congruency. Next, the same Stroop and SOA Proportion manipulations were employed in the context of the Task Set Paradigm to determine whether semantic activation from print requires *intention*, and whether performance optimization plays a role in this context. The results differed from those obtained in the PRP experiments; SOA Proportion modulated the SOA x Congruency interaction. To determine what drives the different results obtained using the PRP and Task Set paradigms, the final three experiments utilized a combination of these two paradigms. Combined, the results of these 7 experiments suggest that (1) semantic activation is capacity limited, (2) semantic activation requires intention, (3) strategic processes play a role in semantic activation, and (4) the overt response to the tone drives the capacity limitations observed in the present PRP experiments.

Acknowledgements

First, I would like to thank my supervisor, Derek Besner, for his continuous support and guidance, and for making sure that I know obscure sayings and which pens he thinks are best. He has always challenged me, which has helped me develop as an academic, a researcher, and a critical thinker. Though his approach is a bit in your face at times, I wouldn't want to change it. It has toughened me up and pushed me to work harder, and I know that it's always coming from a good place. As Derek himself sometimes says, "if you can deal with me, then you can deal with anyone."

I would also like to thank my parents for encouraging my interest in academics from a young age, and for instilling in me the motivation and independence that have helped me throughout my time as a graduate student.

Finally, I would like to thank Evan Risko and Jonathan Fugelsang, my internal committee members, for their valuable feedback on my dissertation.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

A central issue in cognitive psychology is the distinction between controlled and automatic processing, as is reflected in many cognitive psychology textbooks (Ashcraft & Klein, 2010; Eysenck & Keane, 2015; Galotti, Fernandes, Fugelsang, & Stolz, 2010; Goldstein, 2011; among others). Controlled processes (1) require various forms of attention, (2) require intention, (3) are capacity limited, (4) are subject to interference from other processes, and (5) are conscious, whereas automatic processes (1) do not require attention (2) do not require intention, (3) are capacity free, (4) cannot be interfered with by other processes, and (5) are unconscious (see Brown, Gore, & Carr, 2002; Logan, 1988; Posner & Snyder, 1975; Shiffrin & Schneider, 1977; among many others). It is commonly assumed that, over time, skills that are practiced become automatic in these aforementioned senses (LaBerge & Samuels, 1974; Posner & Snyder, 1975; Shiffrin & Schneider, 1977).

The belief that habits and well-practiced skills are automatic in many ways is by no means new (Bryan & Harter, 1899; Cattell, 1886; James, 1890; Thorne, 1955). This supposition has been applied to both motor skills and cognitive skills in a variety of cognitive research areas. Of particular importance for the present investigation is the pervasiveness of the automatic processing perspective in the reading literature. Specifically, my interest rests in the suggestion that semantic activation from print is automatic.

Visual Word Recognition

The dominant view in visual word recognition research is that semantic activation from print is automatic in a variety of senses (e.g., Augustinova & Ferrand, 2014; Brown, Gore, & Carr, 2002; Neely, 1977; Neely & Kahan, 2001; Posner & Snyder, 1975; among many others). Proponents of this automatic processing perspective posit that semantic activation occurs without intention, is ballistic (i.e., once initiated it cannot be stopped), occurs without conscious awareness, cannot be interfered with by other processes, and is capacity free (i.e., it does not require attentional resources).

Although the automatic processing view is widely held (Eysenck & Keane, 2015; Galotti, Fernandes, Fugelsang, & Stolz, 2010; MacLeod, 1991; Reisberg, 1997) there is a small literature whose results conflict with the conclusion that semantic activation from print is automatic (e.g., Baror & Bar, 2016; Besner, 2001; Besner & Reynolds, 2016; Besner, Risko, & Sklair, 2005; Besner & Stolz, 1999; Fagot & Pashler, 1992; Heyman, Van Rensbergen, Storms, Hutchison, & De Deyne, 2015; Labuschagne & Besner, 2015; Lachter, Forster, & Ruthruff, 2004; Lien, Ruthruff, Kouchi, & Lachter, 2010; Robidoux & Besner, 2015; Waetcher, Besner, & Stolz, 2011; among others). However, evidence that contradicts some of the main tenets of the automatic view has not shifted the dominant view: semantic activation in the context of visual word recognition is still purported to be automatic. Given the dominance of this perspective, it is surprising how little research has been conducted to address whether the key tenets of automaticity apply to the various stages of word processing.

The goal of my dissertation is to offer evidence that aims to further address this gap in the literature, as well as to discuss the potential role of performance optimization in the processes that underlie visual word recognition. Since the automatic criteria used by researchers is varied, and a process may be automatic in some ways and not others, it is difficult to assess all of the tenets of automaticity at once (e.g., Bargh, 1989). Instead, these criteria can be investigated individually to determine which characteristics of automaticity are true of a particular process (Moors & De Houwer, 2006). My focus in the present experiments is whether semantic activation is capacity free, occurs without intention, and whether it can be interfered with by other processes (specifically, performance optimization).

Stroop Tasks

All of the experiments discussed here make use of the Stroop task (which uses color words, e.g., *blue*; Stroop, 1935) as well as a variant in which semantic associates are employed (e.g., *sky* which is associated with the color blue; Dalrymple-Alford, 1972; Klein; 1964; Risko, Schmidt, & Besner, 2006). Example stimuli from both of these tasks are shown in Figure 1. Standardly, participants in the Stroop task are required to name the color a word is presented in while ignoring the color carrier word. For both Stroop variants, participants are slower to respond on incongruent trials (trials on which the color carrying word does not match the color that it is presented in, or is not associated with the color that it is presented in) than on neutral trials (trials on which the color carrying word is not a color word or color-associated word) or congruent trials (trials on which the color carrying word matches the color that it is presented in, or is associated with the color that it is presented in).

Figure 1. *The standard and semantic Stroop manipulations.*

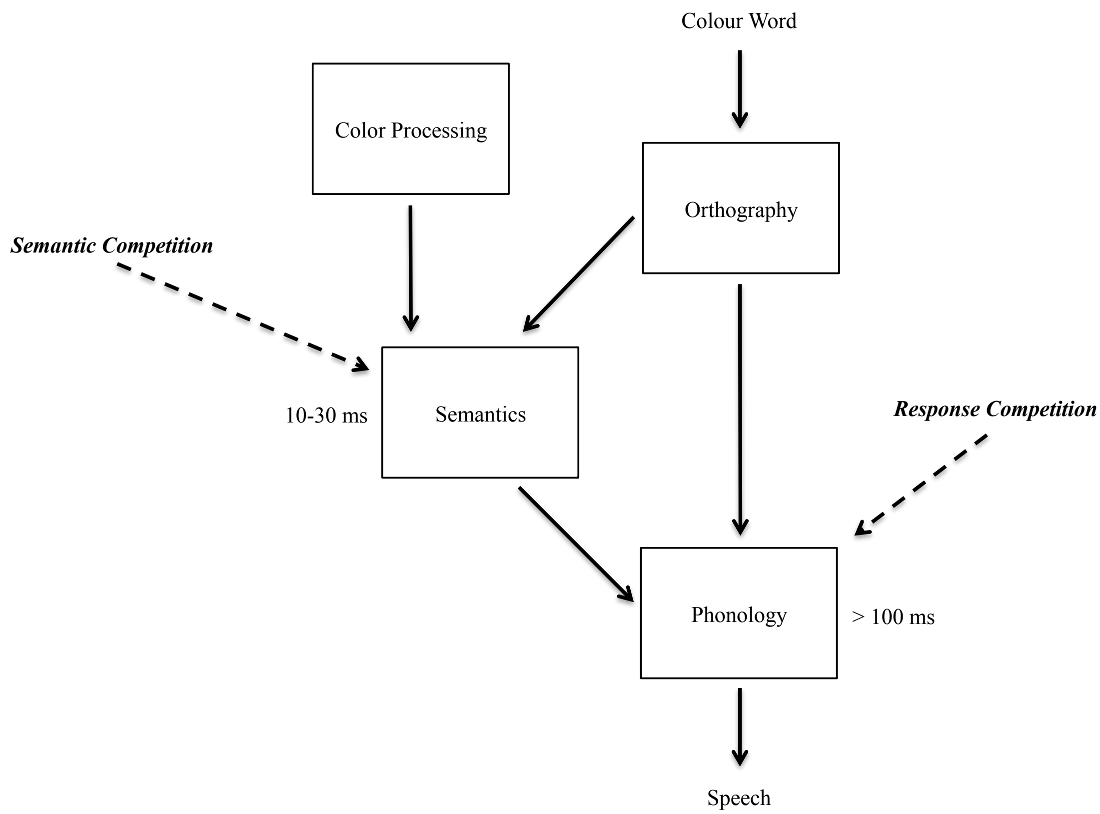
	Standard Stroop	**Semantic Stroop**
incongruent	blue	sky
neutral	table	table
congruent	blue	sky

These Stroop effects are often taken as strong evidence favoring automaticity (e.g., see MacLeod's 1991 review; see also textbooks such as Ashcraft & Klein, 2010; Eysenck & Keane, 2015; Galotti, Fernandes, Fugelsang, & Stolz, 2010; Goldstein, 2011). Indeed, Logan (1988, p. 511) asserts that "the major evidence for automatic processing comes from Stroop and priming studies, in which an irrelevant stimulus influences the processing of a relevant stimulus", and as Brown, Gore, and Carr (2002, p. 220) state "…the assumption of automated word recognition in the mature reader is the 'standard' or 'received' view in cognitive science, in part because of the impact exerted by results from the Stroop task." The argument that Stroop effects reflect automatic processing is based on the idea that participants should not process the irrelevant color word when it interferes with performance, yet they do so because they cannot prevent the automatic processing of the word.

The importance of employing both Stroop tasks results from issues associated with drawing inferences from examining the standard Stroop effect in isolation. The standard Stroop effect does not only reflect semantic level competition; there is also a large response competition component (Augustinova & Ferrand, 2014; Manwell, Roberts, & Besner, 2004). As is shown in Figure 2, the semantic component of the Stroop effect is often between 10 and 30 ms, whereas the response competition component is often 100 ms or greater. Given the much larger effect of response competition, as compared to the effect of semantic competition (at least when the response set includes the irrelevant color word), changes in the size of the semantic component of the standard Stroop effect might not be evident. If a particular manipulation affects *only* the semantic competition component, and this component is absorbed into the much larger response competition component, then there may not be significant changes in the size of the standard

Stroop effect. This makes it difficult to interpret the results obtained using this task as a reflection of processing at the semantic level.

Figure 2. *A depiction of the size of the semantic competition component of the Stroop effect as compared to the response competition component.*



It should also be noted that all of the present experiments use only incongruent and neutral items for both Stroop manipulations. In the context of the standard Stroop task, this avoids participants adopting a strategy in which they read the words to benefit performance on congruent trials (in which the color carrying word matches the color that it is presented in). Incongruent and neutral items are also used with semantic Stroop to maintain consistency across experiments.

Bayesian Analysis

Throughout my dissertation, I discuss whether there is evidence that two factors are additive, or whether they are under-additive. The importance of this distinction will be discussed in more detail in Chapter 2. For now, I would simply like to emphasize that the interpretation of results as additive or under-additive is central to the arguments being made here. For this reason I have included Bayesian analyses for the interactions of interest (Rouder, Speckman, Sun, & Morey, 2009) to supplement my repeated measures ANOVAs. The benefit of this type of analysis is that it provides a ratio of the strength of the evidence favoring the null, as compared to the strength of the evidence favoring the alternative hypothesis (something that cannot be done using a standard ANOVA). In other words, it will provide a ratio for the evidence in favor of the null hypothesis that there is no interaction (i.e., the factors are additive) versus the alternative that there is an interaction (i.e., the factors are under-additive).

Present Investigation

In three sets of experiments I investigate whether semantic activation from print is capacity limited, requires intention, and whether it is influenced by performance optimization. The two paradigms, which will be discussed in more detail in their respective chapters, used here are the Psychological Refractory Period Paradigm (which assesses capacity limitations) and the Task Set Paradigm (which assesses the role of intention). Both paradigms involve manipulating the stimulus onset asynchrony (SOA), so that two stimuli appear closer or further apart in time. The interpretation of the size of a manipulated factor as a function of decreasing SOA is what reveals whether a particular process is automatic in the context of these paradigms.

Broadly speaking, I found that semantic activation from print is indeed capacity limited, is affected by intention, and can be influenced by performance optimization. Additionally, results from experiments that combine the PRP and Task Set paradigms reveal what drives the capacity limitations observed in Chapter 2.

Chapter 2: Is semantic activation from print capacity free?

The first principle of automaticity that I will explore is that automatic processes are capacity free. This key idea of the automatic perspective is evident in early work by Logan (1985; 1997), among others.

> "Tasks that can be performed quickly, **effortlessly** [bold mine], and relatively autonomously are thought to be automatic and tasks that cannot are thought not to be automatic." – Logan, 1985 (p. 368)

> "Automatic processing is **effortless** [bold mine]. Non-automatic processing is effortful. In everyday life, the effortless[ness] of automatic processing is apparent first as a sense of ease and second as **the ability to do another task while performing an automatic one** [bold mine]." – Logan, 1997 (p. 125)

As was mentioned in the introduction, there is some evidence that is argued to contradict the automatic processing perspective. Some of this evidence comes from research using the Psychological Refractory Period (PRP) paradigm (Besner & Reynolds, 2016; Fagot & Pashler, 1992; Magen & Cohen, 2002; 2010). The PRP paradigm is useful in the present context because it can be used to determine whether a particular process is capacity demanding or not. I first briefly review this approach.

The PRP Paradigm

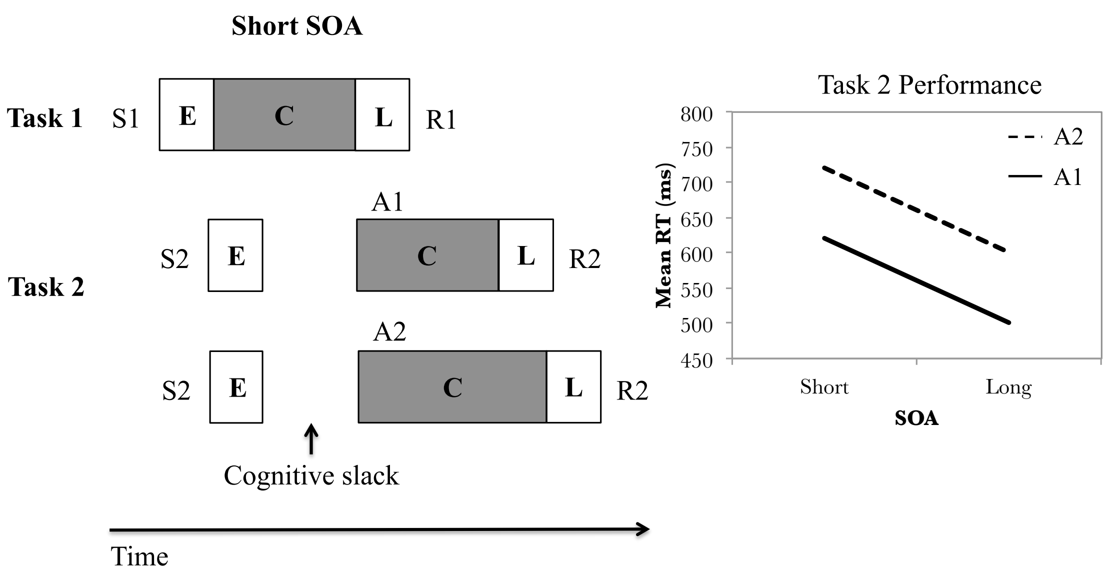
In a typical PRP experiment (see Pashler's 1994 review), participants respond to two stimuli presented sequentially, Stimulus 1 (S1) and Stimulus 2 (S2). Participants are told to respond to S1 before responding to S2. The stimulus onset asynchrony (SOA; the time between the onset of S1 and the onset of S2) is manipulated, as is a factor associated with Stimulus 2 processing.

When the SOA between Tasks 1 and 2 is long, processing associated with S1 can finish before S2 is presented. This condition mimics single task experiments, and the effect of the manipulated factor should be the same size as when the task is being performed in isolation. In contrast, when the SOA is short, processing associated with S1 is still taking place when S2 is presented, which can create a processing bottleneck (capacity limitation) whereby processing of S2 must wait for some processing associated with S1 to finish.

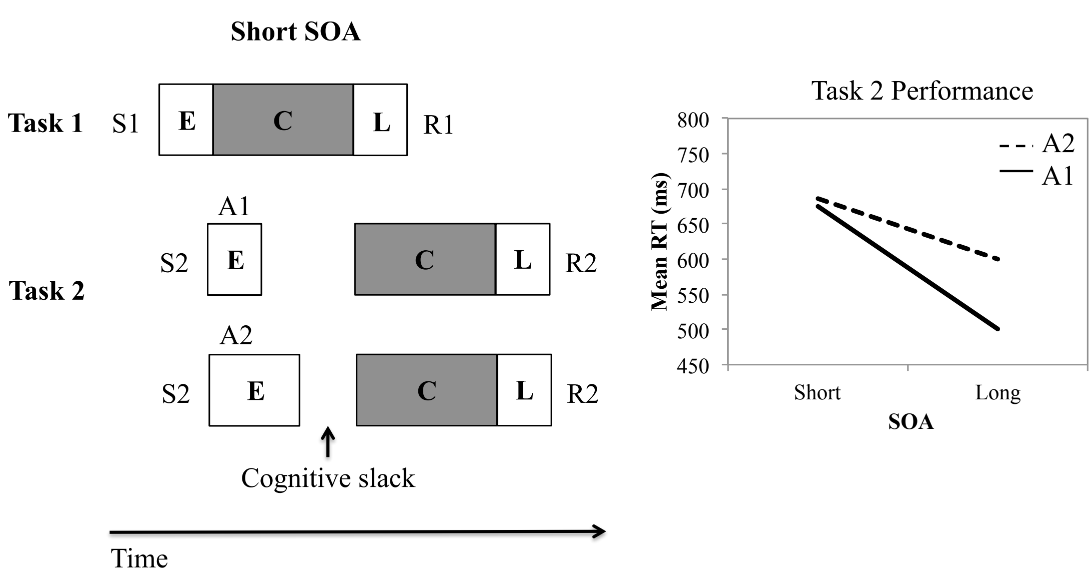
If the process associated with the manipulated factor in Task 2 is capacity limited, then we should see additivity of the manipulated factor and SOA on RT (i.e., the effect should be the same size in both the short and long SOA conditions). This pattern of results would imply that some processing of S2 was put on hold until processing of S1 was complete (it is standardly argued that such processing of S2 is *structurally* bottlenecked, resulting in serial processing as in Figure 3a; see Pashler, 1994). In contrast, if the manipulated factor indexes a process that is capacity free, then we should see under-additivity of our manipulated factor and decreasing SOA (i.e., the size of the effect should decrease as SOA decreases). This pattern of results would imply that processing associated with S2 was absorbed into the slack created by the processing of S1 (i.e., that some element(s) of the two stimuli were processed in parallel, see Figure 3b). That said, this is an oversimplification, because such under-additivity depends on prior processes also being capacity free. Thus, additivity of some factor and SOA may only mean that some prior process was bottlenecked. These constraints have not been widely recognized (see also Besner et al., 2009 for a case in which under-additivity implies both capacity free processing coupled with release from competition).

Figure 3. *A depiction of (a) additivity in PRP, and (b) under-additivity in PRP. "A" represents the manipulated factor, "E" represents early processes, "C" represents central processes, and "L" represents late processes.*

(a) If some processing associated with S2 is **capacity limited**:



(b) If processing associated with S2 is **capacity free**:



Investigations of the standard Stroop effect in the context of PRP have yielded additivity of Congruency and SOA, consistent with the inference that semantic activation from print is capacity limited (Fagot & Pashler, 1992; Magen & Cohen, 2002; 2010). However, there are two potential issues with this interpretation. First, it is possible that the observed additivity of Congruency and SOA actually results from a prior process being bottlenecked (i.e., feature identification, letter identification, or word identification). This is not the case, however, as prior stages have been shown to be capacity free in skilled readers (see Besner et al, 2009; O'Malley, Reynolds, Stolz, & Besner, 2008; Reynolds & Besner, 2006). Second, recall the issue with this interpretation that was addressed in the introduction; the standard Stroop effect consists of both semantic and response competition components. Thus, there could be capacity free processing of the semantic component of the standard Stroop effect that, having been absorbed into the much larger response competition effect, goes undetected.

Besner and Reynolds (2016) therefore conducted an experiment that included both standard and semantic Stroop items in the context of the PRP paradigm, so as to provide a clear answer to the question of whether semantic activation per se is capacity free or not. Their data replicated the observation that the standard Stroop effect is additive with SOA. Moreover, they also found the same pattern of results with semantic Stroop. These results suggest that semantic activation from print is indeed capacity limited in the sense of it being structurally bottlenecked.

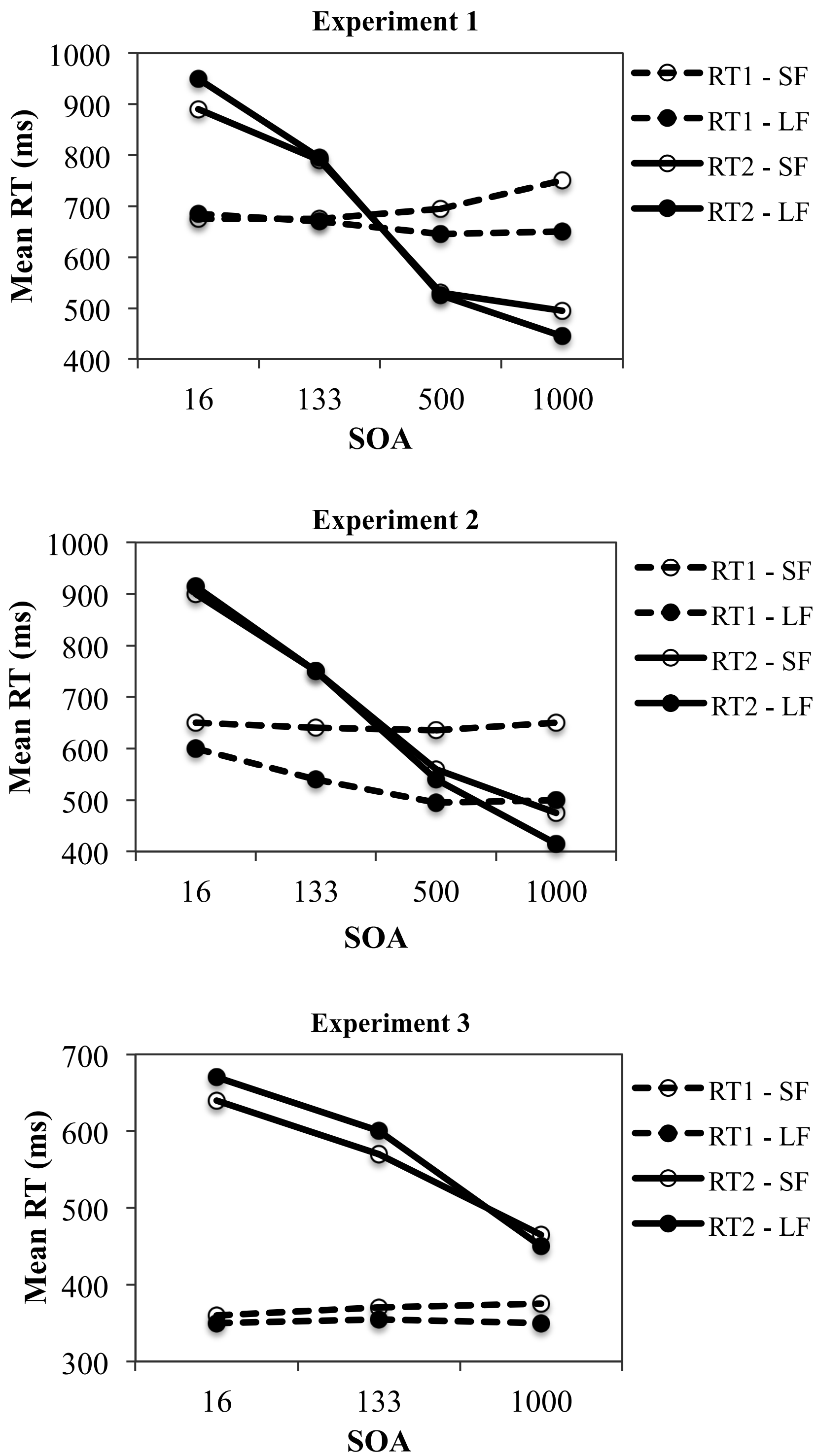
Performance Optimization

A different theoretical approach is seen in work by Miller, Ulrich, and Rolke (2009), who proposed that reports of additivity in the context of PRP do not force an interpretation couched in terms of capacity limitations. Instead, additivity might reflect performance optimization.

According to Miller and colleagues, since participants in PRP experiments are told to respond to both tasks as quickly as possible, they aim to improve performance by reducing their total response time (TRT). TRT is the sum of the time taken to respond to both tasks (i.e., the time from the onset of S1 to the response to S1, added to the time from the onset of S2 to the response to S2). Further, Miller and colleagues argue that serial processing of Tasks 1 and 2 is almost always optimal in PRP experiments. Their account also posits that parallel processing of elements in Tasks 1 and 2 is only optimal when (i) the SOA is short, (ii) there are more short SOA trials than long SOA trials, and (iii) total processing time is not much longer than what would be required to process serially.

Miller and colleagues (2009) investigated these ideas by manipulating SOA Proportion in the context of PRP. SOA Proportion was blocked, so that some blocks had more short SOA trials than long SOA trials (parallel processing is claimed to be more efficient), and others had more long SOA trials than short SOA trials (serial processing is claimed to be more efficient). They found that there was a larger effect of SOA on RT2 in the Long Frequent (LF) condition (more long SOA trials than short SOA trials) than in the Short Frequent (SF) condition (more short SOA trials than long SOA trials). Additionally, RT1 was slower in the SF condition than the LF condition (see Figure 4).

Figure 4. *Evidence of performance optimization in PRP (a reproduction of Miller, Ulrich, & Rolke, 2009, Figures 3a – Experiment 1, 5a – Experiment 2, and 7a – Experiment 3).*



This pattern of results is what is expected if participants prepare for parallel processing when the short SOA is more likely, and for serial processing when the long SOA is more likely. These data are also argued to be inconsistent with a structural bottleneck account because that account predicts no effect of SOA proportion on RT1 or RT2.

Miller and colleagues' findings thus raise the possibility that additivity of factor effects in PRP may be the result of performance optimization rather than a structural bottleneck. However, it is important to note that Miller and colleagues (2009) did not follow up their study with an investigation of how the distribution of SOAs affects a *manipulated factor* in Task 2. The present studies therefore seek to determine whether prior demonstrations of additivity of Stroop Congruency and SOA in the context of the PRP paradigm reflect structural limitations, performance optimization, or both.

To anticipate the present results, Miller and colleagues SOA proportion effect replicates, but so does the additivity of Stroop congruency in both the standard and semantic form. It therefore appears that some processes are structurally bottlenecked, whereas others are subject to performance optimization. More specifically, SOA Proportion does not affect the processes underlying the Stroop effect.

Experiment 1: Standard Stroop in the Context of PRP

The two tasks in this experiment were tone identification (Task 1) and color naming (Task 2). SOA Proportion was manipulated across blocks to determine whether the size of the standard Stroop effect varies as a result of which SOA is more likely. One block had more short SOA trials than long SOA trials (Short Frequent; 80% Short), and the other had more long SOA trials than short SOA trials (Long Frequent; 80% Long). Following Miller and colleagues' (2009) logic, if previous reports of additivity of Congruency and SOA in the context of the PRP paradigm were the result of performance optimization, then the Long Frequent block (which is supposed to promote serial processing on short SOA trials) should yield additive effects of SOA and Congruency, and the Short Frequent block (which is supposed to promote parallel processing on short SOA trials) should yield an under-additive interaction between Congruency and decreasing SOA. In contrast, the structural bottleneck account predicts that additivity of Congruency and SOA should again be observed, despite an effect of SOA Proportion.

Method

Participants. Forty-eight undergraduate students from the University of Waterloo participated for course credit. Each participant was tested individually and had normal or corrected-to-normal vision, normal color vision, normal or corrected-to-normal hearing, and reported English as their first language.

Stimuli. The stimulus set for Task 1 consisted of a high tone (1500 Hz) and a low tone (500 Hz). The stimulus set for Task 2 consisted of the neutral words *keg*, *jail*, *table*, and *palace*, (from Manwell, Roberts, & Besner, 2004) and the color words *red*, *blue*, *green*, and *yellow*. The stimuli were presented in four of the E-Prime preset colors: red (RGB: 255, 0, 0), blue (RGB: 0, 0, 255), green, (RGB: 0, 128, 0), and yellow (RGB: 255, 255, 0). Items were presented in Courier New font size 18. All letters appeared in lower case and were uniformly colored for both neutral and color words. All of the color word trials were incongruent (e.g., *red* presented in blue) so as to remove any potential benefit from reading the word.

Design. Experiment 1 consisted of three within-subject factors, SOA (50 ms vs. 1500 ms); SOA Proportion (Long Frequent Condition – 80% long SOA, 20% short SOA vs. Short Frequent Condition – 20% long SOA, 80% short SOA), and Congruency (incongruent vs. neutral). SOA Proportion was blocked within subjects and counterbalanced for order across participants. Each block had 30 practice trials and 240 experimental trials. Both SOA and Congruency varied randomly from trial to trial.

Apparatus. Task 2 stimuli were displayed on a 22-inch LG Flatron W2242TQ color monitor (29.5 cm high x 47.5 cm wide). The display had a refresh rate of 60 Hz and a resolution of 1680 x 1050 pixels. The auditory stimuli were presented with a set of Logitech X-140 2.0 speakers. E-Prime 2.0 experimental software was used to present the stimuli and record data. The experiment was run on an Ultra Vault PC with an Intel® Core™2 Quad CPU @ 2.40 GHz processor. Participant responses were collected via an Altec Lansing microphone headset attached to a voice key assembly. RTs were measured to the nearest millisecond.

Procedure. Participants were seated approximately 70 cm away from a computer monitor. Each block started with 30 practice trials, followed by 240 experimental trials. Each trial began with the presentation of a fixation marker (+). The fixation marker remained on the screen for 500 ms. 500 ms after the offset of fixation, a 50 ms tone played over the speakers. Participants were instructed to indicate whether the tone was high or low in pitch with a key press. The two response keys were “x” and “n”. Stimulus-response key correspondence was counterbalanced across participants.

Either 50 or 1500 ms after the onset of the tone, a colored word appeared on the screen. Participants were told to name the font color of the word aloud and ignore the color carrier word. The stimulus remained on the screen until participants made a response. Following the participant’s response, the screen remained blank until the researcher coded the response as correct, incorrect, or a mistrial (i.e., the microphone triggered too early or too late). The fixation marker appeared 100 ms after the researcher’s response, indicating the beginning of the next trial. Participants were told to give priority to Task 1, in that they were to make a response to Task 1 before Task 2, and to perform both tasks quickly while maintaining a high degree of accuracy.

Results

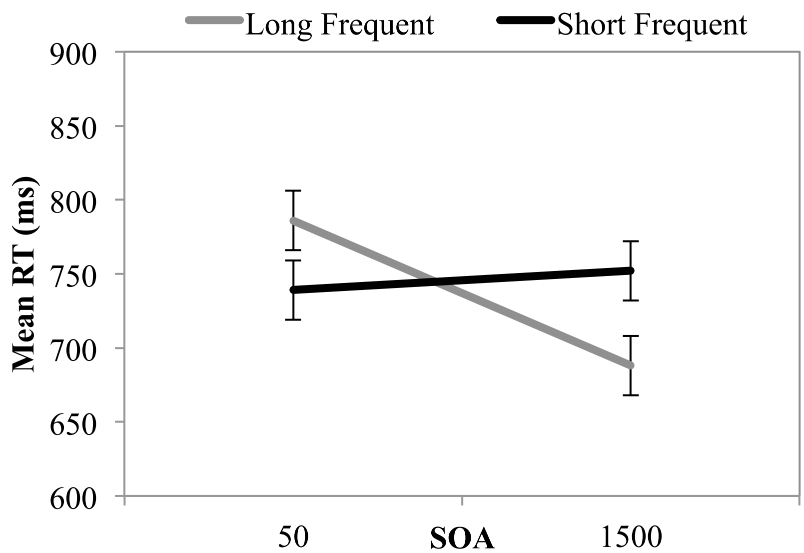
Prior to data analysis, one participant was removed due to a failure to follow instructions (e.g., talking during trials). 5.9% of trials were removed due to microphone errors (i.e., the voice key triggered too early or too late), and 7.8% of trials were removed due to an error being made on either Task 1 or Task 2. Following error removals, trials that had RTs of less than 150 ms or greater than 3000 ms were removed (0.6% of the trials with correct responses were removed using these cutoffs). Trials that were more than 3 standard deviations from the mean in any given condition for Task 1 or Task 2 were also removed, resulting in the removal of 2.2% of trials with correct responses. The same cutoff (3 standard deviations) was also used to determine whether participants were error or RT outliers. The outlier removal procedure was identical in all of the experiments discussed in my dissertation, and will therefore not be outlined again. There were 5 error outliers and 2 RT outliers for Task 1, and 1 RT outlier for Task 2, leaving data from 39 participants for further analysis.

Task 1: Tone Identification

Response Times

First, to investigate whether Miller and colleagues' (2009) results replicate, we looked at the SOA Proportion x SOA interaction (collapsed across Congruency). There was a marginally significant main effect of SOA, $F(1,38) = 4.00$, $p = .053$, participants were slower to respond on short SOA trials ($M = 763$) than long SOA trials ($M = 720$), and a significant SOA Proportion x SOA interaction, $F(1,38) = 14.54$, $p < .001$ (see Figure 3). This pattern of results replicates Miller and colleagues' Experiment 1, in which they also found a significant SOA Proportion x SOA interaction ($p < .001$), with RT1 increasing on the less frequently occurring SOA trials in each condition. Though RT1 was slower in the Short Frequent condition ($M = 745$) than in the Long Frequent condition ($M = 737$), as was reported by Miller and colleagues, the main effect of SOA Proportion was not significant ($F < 1$).

Figure 5. *SOA Proportion x SOA interaction for Task 1 in Experiment 1. Bars represent 95% within-subjects confidence intervals (Loftus & Masson, 1994).*



A three-way ANOVA consisting of SOA Proportion (Long Frequent vs. Short Frequent) x SOA (Short vs. Long) x Congruency (Incongruent vs. Neutral) revealed no significant main effect of Congruency ($F < 1$), no significant SOA Proportion x Congruency interaction, $F(1,38) = 1.92$, $p = .174$, a marginally significant SOA x Congruency interaction, $F(1,38) = 3.18$, $p = .082$, and no significant SOA Proportion x SOA x Congruency interaction ($F < 1$). RTs for Task 1 can be seen in Table 1.

Table 1. *Task 1 mean RTs and percent error by condition for Experiment 1.*

	RTs				**% Error**			
	Long Frequent		**Short Frequent**		**Long Frequent**		**Short Frequent**	
	50	**1500**	**50**	**1500**	**50**	**1500**	**50**	**1500**
Incongruent	801	687	740	747	4.2	2.4	3.2	1.6
Neutral	772	689	738	757	3.9	2.0	4.0	1.9
Difference	**29**	**-2**	**2**	**-10**	**0.3**	**0.4**	**-0.8**	**-0.3**

Errors

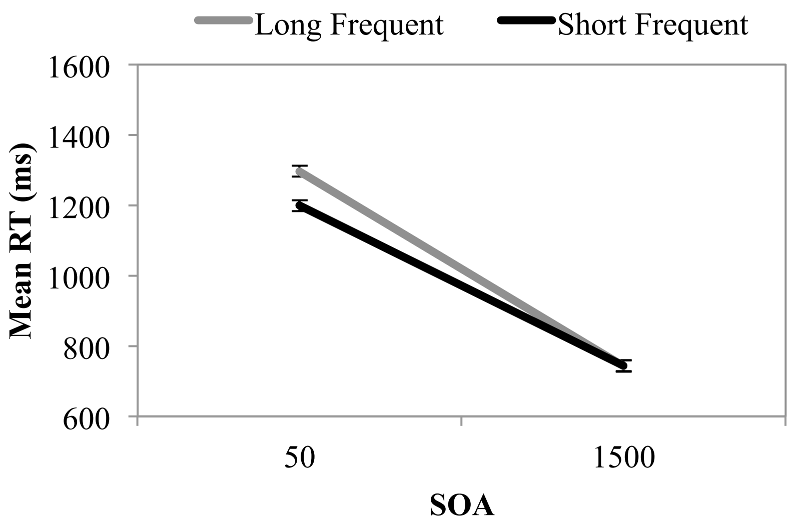
There was a significant main effect of SOA on Task 1 errors, $F(1,38) = 22.27$, $p < .001$. Participants made more errors when the SOA was short ($M = 3.8\%$) as compared to when the SOA was long ($M = 2\%$). There was no significant main effect of SOA Proportion, $F(1,38) = 1.77$, $p = .192$, no significant main effect of Congruency ($F < 1$), no significant SOA Proportion x SOA interaction ($F < 1$), a marginally significant SOA Proportion x Congruency interaction, $F(1,38) = 2.87$, $p = .098$, no significant SOA x Congruency interaction ($F < 1$), and no significant SOA Proportion x SOA x Congruency interaction ($F < 1$, see Table 1).

Task 2: Color Naming

Response Times

As with Task 1, I first looked at the SOA Proportion x SOA interaction (collapsed across Congruency). There was a significant SOA Proportion x SOA interaction on RT2, $F(1,38) = 24.70$, $p < .001$. Consistent with Miller et al.'s (2009) findings in their Experiment 3, RT2 was more affected by the SOA manipulation in the Long Frequent condition than in the Short Frequent condition. More specifically, RT2 was faster on short SOA trials in the Short Frequent condition ($M = 1199$), as compared to the Long Frequent condition ($M = 1297$, see Figure 6). This pattern of results is consistent with participants preparing for parallel processing (a processing mode which aims to benefit short SOA trials) in the Short Frequent condition, and serial processing in the Long Frequent condition,

Figure 6. *SOA Proportion x SOA interaction for Task 2 in Experiment 1. Bars represent 95% within-subjects confidence intervals.*



I then conducted a three-way (SOA Proportion x SOA x Congruency) repeated measures ANOVA (see Table 2). There was a significant main effect of SOA Proportion, $F(1,38) = 11.58$, $p = .002$, participants were slower to respond in the Long Frequent condition ($M = 1020$) than in the Short Frequent condition ($M = 972$), a significant main effect of SOA, $F(1,38) = 577.89$, $p < .001$, participants were slower to respond on short SOA trials ($M = 1248$) than on long SOA trials ($M = 744$), and a significant main effect of Congruency, $F(1,38) = 103.15$, $p < .001$, participants were slower to respond on incongruent trials ($M = 1042$) than on neutral trials ($M = 950$). There was no significant SOA Proportion x Congruency interaction ($F < 1$), and no significant SOA Proportion x SOA x Congruency interaction, $F(1,38) = 1.79$, $p = .189$. There was a marginally significant SOA x Congruency interaction, $F(1,38) = 3.43$, $p = .072$; but note that this interaction is in the wrong direction (over-additive rather than under-additive, see Table 2), and the Scaled JZS Bayes Factor weakly favors the null that there is no interaction (1.2; a Bayes value that is considered inconclusive). More generally, these data are consistent with previous reports by Fagot and Pashler (1992), Magen and Cohen (2002; 2010), and Besner and Reynolds (2016).

Table 2. *Task 2 mean RTs and mean percent error by condition for Experiment 1.*

	RTs				**% Error**			
	Long Frequent		**Short Frequent**		**Long Frequent**		**Short Frequent**	
	50	**1500**	**50**	**1500**	**50**	**1500**	**50**	**1500**
Incongruent	1355	783	1244	786	4.5	1.6	3.3	1.2
Neutral	1239	703	1153	702	1.6	0.7	2.5	0.9
Difference	**116**	**80**	**91**	**84**	**2.9**	**0.9**	**0.8**	**0.3**

Errors

There was a significant main effect of SOA, $F(1,38) = 23.69$, $p < .001$. Participants made more errors on short SOA trials ($M = 3\%$) than on long SOA trials ($M = 1.1\%$). There was a significant main of Congruency, $F(1,38) = 14.85$, $p < .001$. Participants made more errors on incongruent trials (2.7%) than on neutral trials (1.4%). There was also a significant SOA Proportion x Congruency interaction, $F(1,38) = 7.24$, $p = .011$. Participants made more errors on incongruent trials in the Long Frequent condition (3%) than in the Short Frequent condition (2.2%), and made more errors on neutral trials in the Short Frequent condition (1.7%) than the Long Frequent condition (1.1%). There was a significant SOA x Congruency interaction, $F(1,38) = 6.31$, $p = .016$. The SOA manipulation had more of an effect on errors made on incongruent trials (mean error percentage went from 1.4% on long SOA trials to 3.9% on short SOA trials) than errors made on neutral trials (mean error percentage went from 0.8% on long SOA trials to 2.1% on short SOA trials). There was no significant main effect of SOA Proportion ($F < 1$), no significant SOA Proportion x SOA interaction ($F < 1$), and no significant SOA Proportion x

SOA x Congruency interaction, $F(1,38) = 2.29$, $p = .138$. The percent error for each condition can be seen in Table 2.

To summarize the results of Experiment 1, I replicate Miller and colleagues' (2009) findings when I examine the SOA x SOA Proportion interaction (collapsed across Congruency). That is, I found that (1) RT1 increased on the less frequently occurring SOA trials in each SOA Proportion condition, (2) RT2 was more affected by the SOA manipulation in the Long Frequent condition than the Short Frequent condition, and (3) that RT1 was slower in the Short Frequent condition as compared to the Long Frequent condition (but this last effect was not significant in the current investigation). However, I also replicate additivity of SOA and Congruency (the marginally significant interaction was over-additive rather than under-additive, and the Bayes weakly favored the null). Put another way, the size of the Stroop effect in RTs was unaffected by the SOA Proportion manipulation. Though there was over-additivity of Congruency and decreasing SOA in Task 2 errors (the Scaled JZS Bayes Factor of 2.7 weakly favors the alternative), it is important to note that overall RTs on short SOA trials were slower on short SOA trials in the Long Frequent condition, consistent with capacity sharing.

Experiment 2: Semantic Stroop in the Context of PRP

As was noted earlier, it is possible that the *semantic* component of the standard Stroop effect is capacity free, but may go undetected because it is absorbed into the much larger effect of response competition which is capacity limited (See Figure 2). It is therefore important to investigate the semantic Stroop effect in isolation. Experiment 2 was identical to Experiment 1, with the exception that the color-associated words *tomato*, *sky*, *frog*, and *lemon* (from Manwell et al., 2004) were used as the incongruent stimuli instead of color words. This contrast is widely assumed to index the semantic component (e.g., Augustinova and Ferrand, 2014; Manwell et al, 2004; Neely & Kahan, 2001).

Method

Participants. Forty-eight participants from the same participant pool as Experiment 1 participated for course credit. The selection criteria were the same as in Experiment 1.

Results

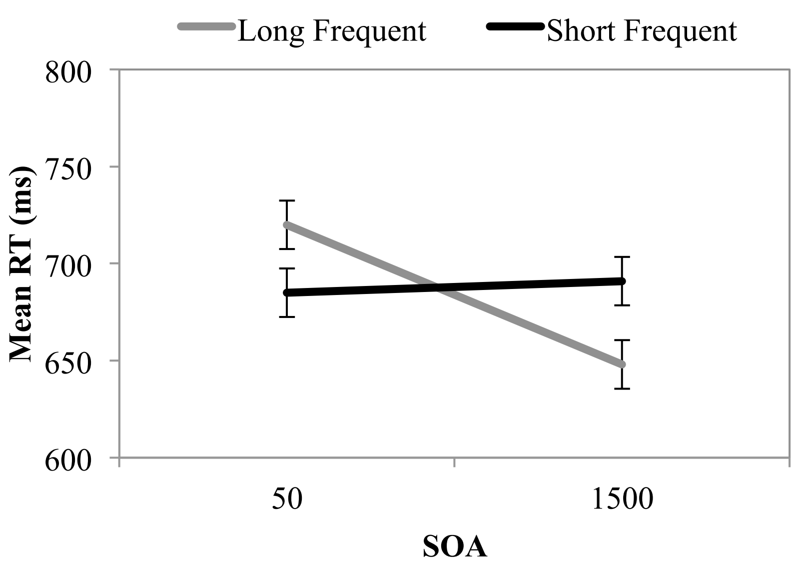
Prior to data analysis, two participants were removed due to a failure to follow instructions. 7.1% of trials were removed due to microphone errors, and 10.4% of trials on which an error was made on either Task 1 or Task 2 were removed. 0.2% of the trials with correct responses were removed because they had an RT of less than 150 ms or greater than 3000 ms. 2.5% of trials with correct responses were removed as RT outliers. 6 participants were error outliers and 2 participants were RT outliers for Task 1, and 1 participant was an RT outlier for Task 2, leaving data from 37 participants for further analysis.

Task 1: Tone Identification

Response Times

For Task 1, there was a significant main effect of SOA, $F(1,36) = 12.36$, $p = .001$, participants were slower to respond on short SOA trials ($M = 702$) than on long SOA trials ($M = 670$), and a significant SOA Proportion x SOA interaction, $F(1,36) = 27.96$, $p < .001$ (see Figure 6), which is again consistent with Miller and colleagues' (2009) Experiment 1 results. As with standard Stroop, there was no significant effect of SOA Proportion involving RT1 ($F < 1$).

Figure 7. *SOA Proportion x SOA interaction for Task 1 in Experiment 2. Bars represent 95% within-subjects confidence intervals*.



A three-way SOA Proportion (Long Frequent vs. Short Frequent) x SOA (Short vs. Long) x Congruency (Incongruent vs. Neutral) ANOVA revealed a significant SOA x Congruency interaction, $F(1,36) = 7.13$, $p = .011$. I believe that this SOA x Congruency interaction is a Type 1 error. The basis for this conclusion is that the Congruency effect is present for *long* SOA trials, in which the tone is presented 1500 ms before the target stimulus, and not on short SOA trials. There were no other significant main effects or interactions (*F*s < 1, see Table 3).

Table 3. *Task 1 RTs and percent error by condition for Experiment 2.*

	RTs				**% Error**			
	Long Frequent		**Short Frequent**		**Long Frequent**		**Short Frequent**	
	50	**1500**	**50**	**1500**	**50**	**1500**	**50**	**1500**
Incongruent	715	655	679	696	6.6	4.5	5.0	4.7
Neutral	724	641	690	686	1.8	2.2	2.0	2.1
Difference	**-9**	**14**	**-11**	**10**	**4.8**	**2.3**	**3.0**	**2.6**

Errors

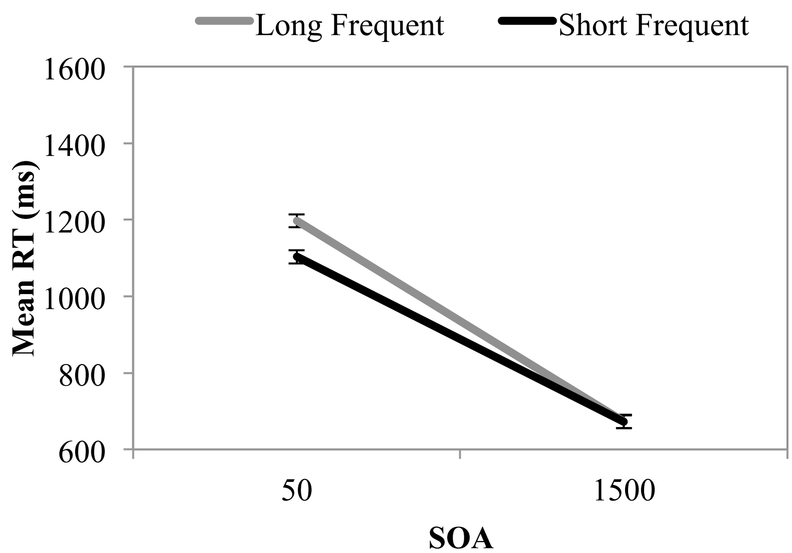
The error data can be seen in Table 3. There was a significant main effect of Congruency, $F(1,36) = 38.46$, $p < .001$, participants made more errors on incongruent trials ($M = 5.2\%$) than on neutral trials ($M = 2.1\%$), and a significant SOA x Congruency interaction, $F(1,36) = 5.33$, $p = .027$. There was a marginally significant main effect of SOA, $F(1,36) = 3.16$, $p = .084$, and a marginally significant SOA Proportion x SOA x Congruency interaction, $F(1,36) = 2.90$, $p = .097$. There was no significant main effect of SOA Proportion ($F < 1$), no significant SOA Proportion x SOA interaction, $F(1,36) = 2.11$, $p = .155$, and no significant SOA Proportion x Congruency interaction, $F(1,36) = 1.19$, $p = .283$.

Task 2: Color Naming

Response Times

There was a significant SOA Proportion x SOA interaction on RT2, $F(1,36) = 14.87$, $p < .001$. Once again, consistent with Miller et al.'s (2009) findings, RT2 was more affected by the SOA manipulation in the Long Frequent condition than in the Short Frequent condition. RT2 was faster on short SOA trials in the Short Frequent condition (*M* = 1103), as compared to the Long Frequent condition (*M* = 1197, see Figure 8).

Figure 8. *SOA Proportion x SOA interaction for Task 2 in Experiment 2 (collapsed across congruency). Bars represent 95% within-subjects confidence intervals.*



A three-way (SOA Proportion x SOA x Congruency) repeated measures ANOVA (see Table 4) revealed a significant main effect of SOA Proportion, $F(1,36) = 9.84$, $p = .003$, participants were slower to respond in the Long Frequent condition (*M* = 935) than in the Short Frequent condition (*M* = 887), a significant main effect of SOA, $F(1,36) = 468.22$, $p < .001$, participants were slower to respond on short SOA trials (*M* = 1150) than on long SOA trials (*M* = 672), and a significant main effect of Congruency, $F(1,36) = 6.48$, $p = .015$, participants were slower to respond on incongruent trials (*M* = 917) than on neutral trials (*M* = 905). There were no other significant interactions involving color naming RTs (*F*s < 1). In particular, as with standard Stroop, Congruency was additive with SOA ($F < 1$, see Table 4).

How strong is the evidence for a null interaction between SOA x Congruency?

A Bayesian analysis (Rouder, Speckman, Sun, & Morey, 2009) was conducted for the SOA x Congruency interaction. The Bayesian analysis yielded a Scaled JZS Bayes factor (5.5) that positively favored the null (see Wagenmakers, 2007). This means that the evidence for the null (no interaction between SOA and Congruency) is 5 times stronger than the evidence for the alternative hypothesis.

Table 4. *Task 2 RTs and percent error by condition for Experiment 2.*

	RTs				**% Error**			
	Long Frequent		**Short Frequent**		**Long Frequent**		**Short Frequent**	
	50	**1500**	**50**	**1500**	**50**	**1500**	**50**	**1500**
Incongruent	1204	681	1106	677	3.3	2.4	2.1	1.3
Neutral	1190	665	1100	667	1.3	0.7	1.3	0.5
Difference	**14**	**16**	**6**	**10**	**2.0**	**1.7**	**0.8**	**0.8**

Errors

There was a significant main effect of SOA Proportion, $F(1,36) = 4.87$, $p = .034$, participants made more errors in the Long Frequent condition ($M = 2\%$) than the Short Frequent condition ($M = 1.2\%$), a significant main effect of SOA, $F(1,36) = 8.67$, $p = .006$, participants made more errors on short SOA trials ($M = 1.9\%$) than on long SOA trials ($M = 1.1\%$), and a significant main effect of Congruency, $F(1,36) = 15.48$, $p < .001$, participants made more errors on incongruent trials ($M = 2.3\%$) than on neutral trials ($M = 0.9\%$). There was a marginally significant SOA Proportion x Congruency interaction, $F(1,36) = 3.50$, $p = .07$. There were no significant interactions on Task 2 errors (Fs < 1). Percentage error for each condition can be seen in Table 4.

Although the semantic Stroop effect is quite small overall (12 ms), the RT results mirror those yielded in the standard Stroop experiment. The semantic Stroop effect was additive with SOA (the Bayes of 5.5 positively favored the null).

Combined analysis of the PRP experiments

To determine whether increasing the power to detect the three-way (SOA Proportion x SOA x Congruency) interaction would yield different results, the two PRP experiments were combined in one analysis.

SOA and Congruency were still additive when the combined data was considered, $F(1,76) = 1.32$, $p = .254$, the Scaled JZS Bayes Factor (4.3) positively favored the null. This interaction did not vary as a function of type of Stroop ($F(1,76) = 1.20$, $p = .276$, for the SOA x Congruency x Experiment interaction). Additionally, the SOA Proportion x SOA x Congruency interaction was still non-significant in the combined analysis, $F(1,76) = 1.87$, $p = .175$, the Scaled JZS Bayes Factor (3.3) positively favored the null. This interaction also did not vary as a function of type of Stroop ($F(1,76) = 1.21$, $p = .275$).

Discussion

Empirical conclusions

The present two experiments both replicate previous observations, and provide new ones. First, the standard Stroop effect was additive with SOA on RT in Task 2, replicating Fagot and Pashler (1992), Magen and Cohen (2002; 2010), and Besner and Reynolds (2016). Second, the semantic Stroop effect was also additive with SOA in Task 2, replicating Besner and Reynolds (2016). Thirdly, SOA Proportion interacted strongly with SOA (collapsed across Congruency) in both experiments, replicating Miller and colleagues (2009). Finally, I stress that in neither experiment was the three-way interaction of SOA, SOA Proportion, and Congruency significant, nor was the two-way interaction of SOA and Congruency.

Theoretical conclusions

I take these results to support the following conclusions. First, the *absence* of the three-way interaction between SOA, SOA Proportion, and Congruency, coupled with the *absence* of a two-way interaction between SOA and Congruency, and the *presence* of the two-way interaction between SOA Proportion and SOA in both experiments constrains Miller and colleagues' (2009) performance optimization account. In their account, the interaction of SOA Proportion and SOA is taken to imply that when short SOAs occur more frequently than long SOAs, this maximizes the opportunity for parallel processing across Tasks 1 and 2. If this account applies to the operation of *all* processes, then I would have expected to see that the Stroop effect, associated maximally with response competition (the standard version which uses color words), would have been under-additive with SOA in the Short Frequent condition. However, no such effect was observed. Given this failure, the simplest account is that either parallel processing applies only to processes following those where the response competition component in the Stroop task affects performance, and/or that the SOA Proportion effects reflect some process(es) other than parallel versus serial processing across tasks (e.g., anticipation of a short SOA makes it easier to apply some early encoding process, some later response preparation process, or both such processes). No existing data speaks to which of these possibilities should be preferred at present.

Given that there is additivity of SOA and Congruency, coupled with no interaction with SOA Proportion, this is consistent with the hypothesis that response competition in the standard version of the Stroop effect is structurally bottlenecked. That is, this process must wait until some capacity limited process in Task 1 is freed up before response competition can take place, as it too needs capacity to unfold.

Most importantly, I take the additivity of the semantically based Stroop effect (seen with color-associated words in Experiment 2) and SOA to imply that semantic activation is also bottlenecked. That is, in contrast to lexical processing, which has been shown to be largely capacity free in previous work with participants drawn from the same population (Besner et al., 2009; O'Malley, Reynolds, Stolz, & Besner, 2008; Reynolds & Besner, 2006), the mapping from the lexical level to the semantic level appears to require some form of capacity that is also required by Task 1. In short, these results are consistent with this process being structurally bottlenecked.

The overall conclusion is that semantic activation appears to require some form of capacity. This contrasts with the widely held view that such processing is capacity free (see Neely & Kahan, 2001). Converging evidence consistent with the present conclusion is seen in the work by Heyman, Van Rensbergen, Storms, Hutchison, and De Deyne (2015) who report that semantic priming observed when prime and target are related in a forward manner (e.g., *panda-bear*) is eliminated by carrying a demanding visuo-spatial load.[1]

[1] Though it should be noted that this work has yet to be replicated.

Chapter 3: Does semantic activation occur regardless of an individual's intention?

As was noted in the introduction, the automatic processing perspective also posits that semantic activation occurs without intent (Augustinova & Ferrand, 2012; Neely & Kahan, 2001; Posner & Snyder, 1975). In other words, when presented with a word, an individual will process it even if they do not wish to do so. As Neely and Kahan (2001) and Augustinova and Ferrand (2012) conclude:

> "…SA [semantic activation] is indeed automatic in that it is **unaffected by the intention for it to occur** [bold mine]." (Neely & Kahan, 2001, p. 88).

> "…semantic activation in the Stroop task is indeed automatic and ballistic, in the sense that **it occurs without intent** and cannot be prevented…" (Augustinova & Ferrand, 2012, p. 525)

One issue with investigating the role that intention plays in word processing is that in the majority of standard word recognition experiments participants know in advance what task they will be performing on each trial. This is problematic because participants may adopt an experiment wide mental set in which they intend to engage in word processing on every trial, making it difficult to determine whether processing would occur in the absence of intention. Besner and Care (2003) created a paradigm that can be used to address this issue, the Task Set Paradigm.

The Task Set Paradigm (Besner & Care, 2003)

Similarly to the PRP paradigm, participants are first presented with a tone that is either high or low in pitch. However, in this case the tone acts as a cue that indicates which of two tasks participants are to perform on any given trial. As is the case in PRP, the SOA varies, so that on some trials participants know which of the two tasks they are required to perform in advance of the target. This gives participants the opportunity to prepare their mental set. On other trials, participants are presented with the cue and target item closer together in time. This does not give the participant time to prepare in advance. Since the participant does not know which task they will be performing on any given trial, they should not intend to perform one of the two tasks, as this would disadvantage them on 50% of trials.

The interpretation of results in the context of the Task Set Paradigm makes use of the same cognitive slack logic as PRP. If the target is processed without intention, then target processing can unfold during the time taken to decode the cue. In this case, there should be under-additivity of Congruency and decreasing SOA (i.e., the size of the Congruency effect should decrease as SOA decreases), as the effect has been absorbed into the slack created by the time taken to decode the task cue and implement the appropriate set. This pattern of results would support the automatic processing perspective; namely that words are processed regardless of an individual's intention. If the target is not processed without intention, however, then processing will be put on hold until the cue is decoded and the participant has implemented the Task Set. This would yield additivity of Congruency and SOA (the size of the Congruency effect does not change as a function of SOA).[2]

---

[2] Besner and Care (2003) also offer another potential explanation for additivity in this context. It could be the case that processing the cue requires the same resources as target processing, and additivity is observed because of capacity limitations (as is the case in PRP). I will return to this possibility when discussing the results of the two Task Set experiments.

Performance Optimization

Let us revisit the argument outlined by Miller, Ulrich, and Rolke (2009) regarding observations of additivity in PRP. According to Miller et al (2009), additivity could simply reflect the adoption of a strategy in which processing is put on hold to improve performance by reducing TRT. A similar argument could be made in the context of the Task Set Paradigm, given its parallel construction with the PRP paradigm. Perhaps additivity in the context of the Task Set Paradigm occurs as a result of participants putting task processing on hold until the cue has been processed, not simply as a result of task uncertainty, but instead as a result of a strategy in which they aim to reduce their TRT (in this case, the time taken to decode the task cue and respond to the required task on a given trial). If this is the case, then the size of the congruency effect should vary as a function of SOA proportion under Miller et al's (2009) performance optimization account. Congruency should be additive in the Long Frequent condition and under-additive in the Short Frequent condition.

Why would the results differ across these two paradigms?

The results of the PRP experiments in Chapter 2 demonstrated that the SOA Proportion manipulation did not modulate the size of the Congruency effect. This is consistent with the suggestion that the additivity observed in Chapter 2 is due to a processing bottleneck. The results could differ in the context of the Task Set paradigm, however, as there is no explicit Task 1 that must be completed prior to responding to the target stimulus (i.e., there is no overt response to the tone), hence there might not be a processing bottleneck. Additionally, there is evidence from O'Malley and Besner (2011) that processing the cue in the context of Task Set paradigm does not interfere with previous stages of word processing. In these experiments, word frequency was under-additive with decreasing SOA suggesting that lexical processing can unfold while the cue is being decoded.

As in Chapter 2, I used both standard and semantic Stroop to examine the role of intention in semantic activation from print.

Experiment 3: Standard Stroop in the Context of the Task Set Paradigm

The two tasks used in the Task Set experiments were case decision and color naming. Case decision was used as the second task in the present experiments, as it has typically been used as the second task in the context of the Task Set paradigm (e.g., Besner & Care, 2003; O'Malley & Besner, 2011). As with the two PRP experiments (Chapter 2), SOA Proportion was manipulated across blocks to determine whether the size of the standard Stroop effect varies as a result of which SOA is more likely. There are three potential outcomes given the present manipulation in the context of the Task Set Paradigm: (1) the standard Stroop effect is under-additive with decreasing SOA in both SOA Proportion conditions, consistent with the automatic processing account, (2) the standard Stroop effect is additive with SOA in both SOA Proportion conditions, inconsistent with the automatic processing account, but not necessarily consistent with performance optimization (the SOA Proportion x SOA interaction will shed more light on its role), or (3) the standard Stroop effect is under-additive in one SOA proportion condition, and additive in the other, consistent with a performance optimization account (and still inconsistent with the automatic processing account, which posits that semantic activation should be unaffected by such manipulations).

Method

Participants. Forty-eight participants from the same participant pool as Experiments 1 and 2 participated for course credit. The selection criteria were the same as in Experiments 1 and 2.

Stimuli. The stimuli in Experiment 3 were identical to Experiment 1, with the exception that half of the items were presented in uppercase letters (given that one of the tasks was case decision).

The **design** and **apparatus** were identical to Experiment 1.

Procedure. The procedure for the Task Set experiments varied from the PRP experiments in the following ways; (1) The 50 ms tone that played over the speakers following the offset of fixation acted as a cue that indicated which of two tasks participants were required to perform on any given trial; participants did not overtly respond to the cue, and (2) 50% of trials were color naming trials, on which participants were required to name the color that the target item was presented in, and 50% of trials were case decision trials, on which participants were required to indicate, via key press, the case (UPPER versus lower) that the target item was presented in. The two keys were 'g' and 'h', and the key-case correspondence was counterbalanced across participants. Participants were told to respond as quickly and accurately as possible in both tasks.

Results

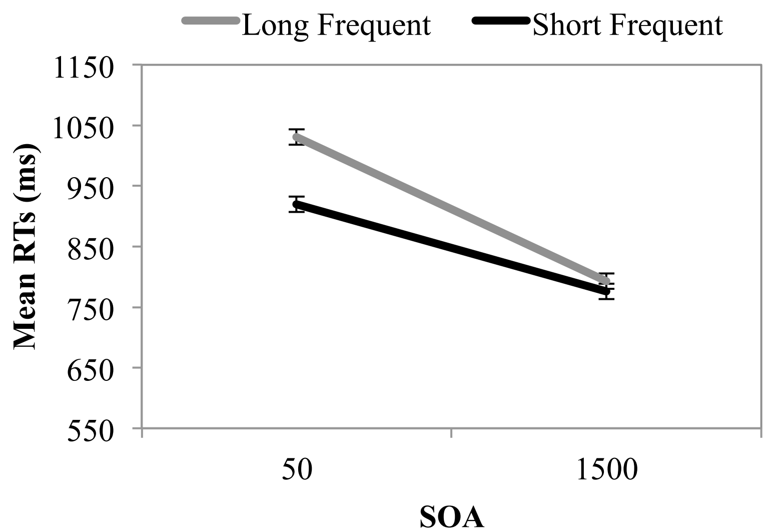
Four participants were removed due to a failure to follow instructions. 2.7% of trials were removed due to microphone errors, 3.7% were removed as a result of an error on the color naming task, and 1.7% were removed as a result of an error on the case decision task. Following these removals, trials with correct responses that had an RT of less than 150 ms or greater than 3000 ms were also removed, resulting in the removal of 0.4% of color naming trials and 0.2% of case decision trials. Of the trials with correct responses, 1.1% were RT outliers in the color naming task and 1.2% were outliers in the case decision task. Additionally, 5 participants were error outliers and 3 participants were RT outliers in the case decision task, and 1 participant was an RT outlier in the color naming task, leaving data from 35 participants for further analysis.

Color Naming Task

Response Times

As with Experiments 1 and 2, I first conducted a repeated measures ANOVA to determine whether there was an SOA Proportion x SOA interaction on color naming response times. As was the case in the previous two experiments, there was a significant SOA Proportion x SOA interaction, $F(1,34) = 25.94$, $p < .001$. The results were once again consistent with Miller et al.'s (2009) expected findings for RT2. Color naming responses times were more affected by the SOA manipulation in the Long Frequent condition than in the Short Frequent condition. Specifically, as is shown in Figure 9, mean RT was faster for short SOA trials in the Short Frequent condition (M = 919), as compared to the Long Frequent condition (M = 1031).

Figure 9. *SOA Proportion x SOA interaction for the Color Naming Task in Experiment 3 (collapsed across congruency). Bars represent 95% within-subjects confidence intervals.*



A three-way (SOA Proportion x SOA x Congruency) repeated measures ANOVA yielded a significant main effect of SOA Proportion, $F(1,34) = 17.41$, $p < .001$, participants were slower to respond in the Long Frequent condition ($M = 911$) than in the Short Frequent condition ($M = 847$), a significant main effect of SOA, $F(1,34) = 223.895$, $p < .001$, participants were slower to respond on short SOA trials ($M = 975$) than on long SOA trials ($M = 784$), and a significant main effect of Congruency, $F(1,34) = 30.01$, $p < .001$, participants were slower to respond on incongruent trials ($M = 908$) than on neutral trials ($M = 851$).

There was also a significant SOA x Congruency interaction, $F(1,34) = 6.08$, $p = .019$. The Stroop effect was smaller on short SOA trials ($M = 22$) than long SOA trials ($M = 80$). There was no significant SOA Proportion x Congruency, $F(1,34) = 1.48$, $p = .232$, or SOA Proportion x SOA x Congruency, $F(1,34) = 1.70$, $p = .202$, interaction. The RTs and percent error for the color naming task are shown in Table 5.

Table 5. *Color naming Task RTs and percent error by condition for Experiment 3.*

	RTs				% Error			
	Long Frequent		**Short Frequent**		**Long Frequent**		**Short Frequent**	
	50	**1500**	**50**	**1500**	**50**	**1500**	**50**	**1500**
Incongruent	1035	833	947	815	4.0	2.2	3.3	2.2
Neutral	1026	751	892	736	3.5	2.2	3.1	2.0
Difference	**9**	**82**	**55**	**79**	**0.5**	**0.0**	**0.2**	**0.2**

Errors

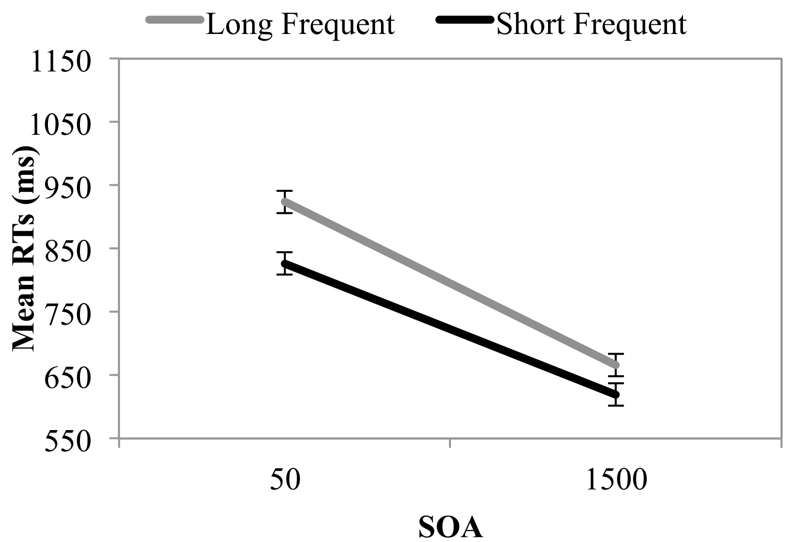
There was a significant main effect of SOA on color naming errors, $F(1,34) = 8.38$, $p = .007$. Participants made more errors on short SOA trials ($M = 3.5$) than on long SOA trials ($M = 2.2$). There were no other main effects or interactions on color naming errors (*F*s < 1).

Case Decision Task

Response Times

The case decision task can also be considered RT2, as it too follows the decoding of the task cue. Following Miller et al's (2009) logic, there should be a significant SOA Proportion x SOA interaction for this task as well. This was indeed the case, $F(1,34) = 6.93$, $p = .013$. As is shown in Figure 10, case decision responses times were more affected by the SOA manipulation in the Long Frequent condition than in the Short Frequent condition; mean RT was faster for short SOA trials in the Short Frequent condition ($M = 826$), as compared to the Long Frequent condition ($M = 923$).

Figure 10. *SOA Proportion x SOA interaction for the Case Decision Task in Experiment 3 (collapsed across congruency). Bars represent 95% within-subjects confidence intervals.*



There was a significant main effect of SOA Proportion, $F(1,34) = 14.14$, $p = .001$, such that participants were slower to respond in the Long Frequent condition ($M = 794$) than in the Short Frequent condition ($M = 722$), a significant main effect of SOA, $F(1,34) = 123.48$, $p < .001$, participants were slower to respond on short SOA trials ($M = 874$) than on long SOA trials ($M = 642$), and a significant main effect of Congruency, $F(1,34) = 14.01$, $p = .001$, participants were slower to respond on incongruent trials ($M = 772$) than on neutral trials ($M = 745$). There was no significant interaction between SOA Proportion x Congruency, or SOA x Congruency, (Fs < 1), or significant SOA Proportion x SOA x Congruency interaction, $F(1,34) = 2.73$, $p = .108$.

Table 6. *Case Decision Task RTs and percent error by condition for Experiment 3.*

	RTs				**% Error**			
	Long Frequent		**Short Frequent**		**Long Frequent**		**Short Frequent**	
	50	**1500**	**50**	**1500**	**50**	**1500**	**50**	**1500**
Incongruent	931	688	842	626	4.2	1.9	4.6	1.7
Neutral	915	643	809	611	2.5	2.1	3.0	0.7
Difference	**16**	**45**	**33**	**15**	**1.7**	**-0.2**	**1.6**	**1.0**

Errors

There was a significant main effect of SOA, $F(1,34) = 20.07$, $p < .001$, and a significant main effect of Congruency, $F(1,34) = 9.56$, $p = .004$, on errors in the case decision task. There was no significant main effect of SOA Proportion ($F < 1$). There was also no significant SOA Proportion x Congruency interaction ($F < 1$), SOA x Congruency interaction, $F(1,34) = 2.18$, $p =$ .149, or SOA Proportion x SOA x Congruency interaction ($F < 1$).

The significant SOA Proportion x SOA interactions found for both tasks (color naming and case decision) are consistent with Miller and colleagues' (2009) performance optimization account. However, in contrast to the results obtained with the PRP paradigm (additivity of Congruency and SOA), Congruency appears to be under-additive with decreasing SOA. Though it looks like there is a trend towards a significant three-way (SOA Proportion x SOA x Congruency) interaction for the color naming task (the Stroop effect decreased by 24 ms in the Short Frequent condition and by 73 ms in the Long Frequent condition), this 49 ms interaction was not significant.

Additionally, there was a significant Stroop effect in the *case decision task*. There are two potential explanations for the presence of this effect. The first is Besner and Care's (2003) alternative explanation for additivity in the context of the Task Set paradigm. They argue:

> "…if target processing leads to the computation of both potential responses in the short SOA condition prior to cue decoding, then the size of the [effect of interest] at the short SOA should be equivalent for [both tasks]." (p. 312)

If cue processing and target processing both require the same resource, and both tasks are computed on each trial, then the effect of the manipulated factor should be present in both tasks. However, if this were the case in the present experiment, then the effect of Congruency should *not* be under-additive with decreasing SOA in any condition, as there should be a bottleneck in processing if both processes require the same resource. Given that we see a significant SOA x Congruency interaction in color naming task (a 58 ms reduction), it is clear that processing *can* unfold at the same time as the cue is being decoded. The second explanation that I propose is that **response set implementation** is intentionally put on hold in some conditions until the cue and target have both been processed to optimize performance (a point that will be revisited in the discussion section of this chapter).

Experiment 4: Semantic Stroop in the Context of the Task Set Paradigm

Remember, the standard Stroop effect does not necessarily index semantic processing. It is therefore important to also examine the size of the semantic Stroop effect in the context of the Task Set Paradigm.

Method

Participants. Forty-eight participants from the same participant pool as Experiments 1, 2, and 3 participated for course credit. The selection criteria were the same as in earlier experiments.

Stimuli. The stimuli only differed from Experiment 3, in that the incongruent stimuli were color-associated words (*tomato, sky, frog, lemon*) instead of color words (*red, blue, green, yellow*).

The **design**, **apparatus**, and **procedure** were identical to Experiment 3.

Results

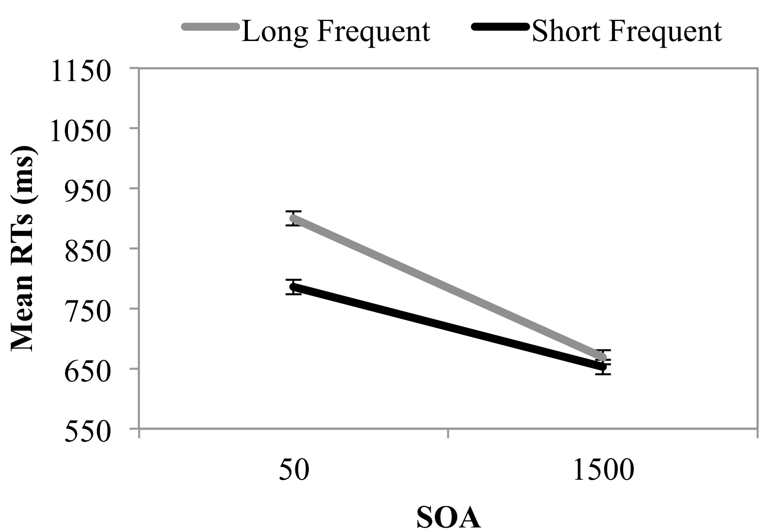
2.7% of trials were removed due to microphone errors, 3.7% were removed as a result of an error on the color naming task, and 1.7% were removed as a result of an error on the case decision task. Following these removals, trials with correct responses that had an RT of less than 150 ms or greater than 3000 ms were also removed (0.4% of color naming trials and 0.2% of case decision trials). Of the trials with correct responses, 1.1% were RT outliers in the color naming task and 1.2% were outliers in the case decision task. At the participant level, 4 were error outliers and 2 were RT outliers in the case decision task, leaving data from 42 participants for further analysis.

Color Naming Task

Response Times

There was a significant SOA Proportion x SOA interaction on color naming response times, $F(1,41) = 27.56$, $p < .001$. As with all of the experiments reported thus far, the size of the SOA effect was more pronounced in the Long Frequent condition. Response times were slower on short SOA trials in the Long Frequent condition ($M = 900$) than on short SOA trials in the Short Frequent condition ($M = 785$). These results are shown in Figure 11.

Figure 11. *SOA Proportion x SOA interaction for the Color Naming Task in Experiment 4 (collapsed across congruency). Bars represent 95% within-subjects confidence intervals.*



A three-way (SOA Proportion x SOA x Congruency) repeated measures ANOVA revealed a significant main effect of SOA Proportion, $F(1,41) = 27.39$, $p < .001$, whereby participants were slower to respond in the Long Frequent condition ($M = 784$) than in the Short Frequent condition ($M = 719$), a significant main effect of SOA, $F(1,41) = 181.78$, $p < .001$, in that participants were slower to respond on short SOA trials (*M* = 843) than on long SOA trials (*M* = 660), and a significant main effect of Congruency, *F*(1,41) = 5.94, *p* = .019, participants were slower to respond on incongruent trials (*M* = 758) than on neutral trials (*M* = 745).

In addition to the significant SOA Proportion x SOA interaction mentioned previously, there was a significant SOA Proportion x Congruency interaction, *F*(1,41) = 6.69, *p* = .013, and a marginally significant SOA Proportion x SOA x Congruency interaction, *F*(1,41) = 2.93, *p* = .095. The effect of Congruency was smaller in the Long Frequent condition (*M* = -2) than in the Short Frequent condition (*M* = 26). Additionally, the effect of Congruency is trending toward being under-additive with decreasing SOA in the Long Frequent condition (a reduction from 11 ms to -15 ms; though this reduction is not significant, *F*(1,41) = 2.64, *p* = .112), and not in the Short Frequent condition (an increase in the size of the effect from 22 ms at the long SOA to 31 ms at the short SOA). There was no significant SOA x Congruency interaction (*F* < 1).

Table 7. *Color Naming Task RTs and percent error by condition for Experiment 4.*

	RTs				**% Error**			
	Long Frequent		**Short Frequent**		**Long Frequent**		**Short Frequent**	
	50	**1500**	**50**	**1500**	**50**	**1500**	**50**	**1500**
Incongruent	892	674	801	663	2.2	1.6	2.3	1.2
Neutral	907	664	770	641	3.1	0.8	1.5	1.2
Difference	**-15**	**11**	**31**	**22**	**-0.9**	**0.8**	**0.8**	**0.0**

Errors

There was a significant main effect of SOA on color naming errors, *F*(1,41) = 11.96, *p* = .001, participants made more errors on short SOA trials (*M* = 2.3) than on long SOA trials (*M* = 1.2), and a marginally significant SOA Proportion x SOA x Congruency interaction, *F*(1,41) = 2.98, *p* = .092.

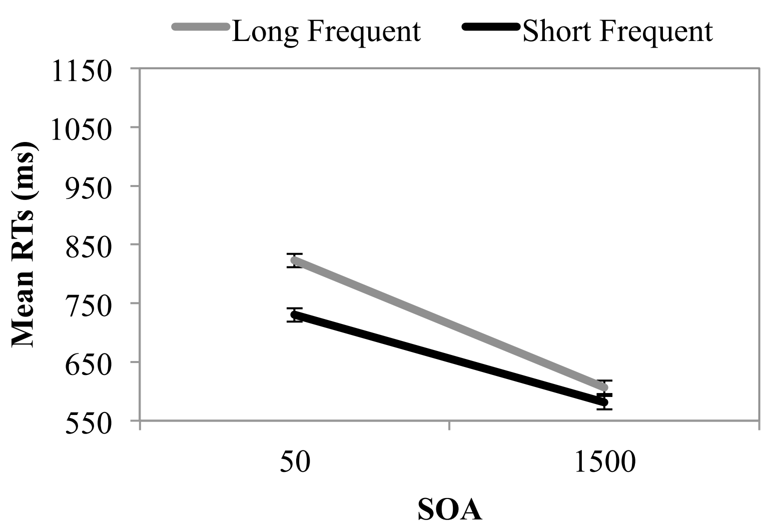
There was no significant main effect of SOA Proportion, $F(1,41) = 1.01$, $p = .322$, or Congruency ($F < 1$) on color naming errors, nor was there a significant SOA Proportion x SOA interaction, $F(1,41) = 2.20$, $p = .146$, SOA Proportion x Congruency interaction ($F < 1$), or SOA x Congruency interaction ($F < 1$).

Case Decision Task

Response Times

Consistent with the performance optimization account, as is evident in Figure 12, there was a significant SOA Proportion x SOA interaction, $F(1,41) = 14.43$, $p < .001$. Response times were slower on the short SOA trials in the Long Frequent condition ($M = 822$) than in the Short Frequent condition ($M = 730$).

Figure 12. *SOA Proportion x SOA interaction for the Case Decision Task in Experiment 4 (collapsed across congruency). Bars represent 95% within-subjects confidence intervals.*



A three factor (SOA Proportion x SOA x Congruency) repeated measures ANOVA revealed a significant main effect of SOA Proportion, $F(1,41) = 22.32$, $p < .001$. Participants were slower to respond in the Long Frequent condition ($M = 714$) than in the Short Frequent condition ($M = 655$). There was a significant main effect of SOA, $F(1,41) = 226.86$, $p < .001$. Participants were slower to respond on short SOA trials ($M = 775$) than on long SOA trials ($M = 593$). Interestingly, as with standard Stroop, there was a significant main effect of Congruency, $F(1,41) = 6.68$, $p = .013$. Participants were slower to respond on incongruent trials ($M = 691$) than on neutral trials ($M = 677$).

There was a significant SOA Proportion x Congruency interaction, $F(1,41) = 4.74$, $p = .035$. The Congruency effect was smaller in the Long Frequent condition ($M = 4$) than in the Short Frequent condition ($M = 25$). There was no significant SOA x Congruency interaction ($F < 1$), or SOA Proportion x SOA x Congruency interaction, $F(1,41) = 2.78$, $p = .103$.

Table 8. *Case Decision Task RTs and percent error by condition for Experiment 4.*

	RTs				**% Error**			
	Long Frequent		**Short Frequent**		**Long Frequent**		**Short Frequent**	
	50	**1500**	**50**	**1500**	**50**	**1500**	**50**	**1500**
Incongruent	820	612	746	589	3.2	3.0	3.3	2.8
Neutral	824	600	713	572	3.6	1.8	3.7	1.8
Difference	**-4**	**12**	**33**	**17**	**-0.4**	**1.2**	**-0.4**	**1.0**

Errors

There was a significant main effect of SOA, $F(1,41) = 7.47$, $p = .009$. Participants made more errors on short SOA trials ($M = 3.4$) than on long SOA trials ($M = 2.3$). There was no significant main effect of SOA Proportion ($F < 1$), or of Congruency, $F(1,41) = 1.07$, $p = .308$, on case decision errors.

There was a marginally significant SOA x Congruency interaction on errors, $F(1,41) = 3.56$, $p = .066$. No other interactions were significant (Fs < 1).

What is the evidence for the null three-way interaction in standard Stroop, and the marginally significant three-way interaction in semantic Stroop?

The Scaled JZS Bayes Factor weakly favored the null for standard Stroop (2.5) and was inconclusive for semantic Stroop (1.6 in favor of the null).

Combined analysis of the Task Set experiments

The pattern of results in both the standard and semantic Stroop Task Set experiments showed a decrease in the size of the Congruency effect in the Long Frequent condition, and not in the Short Frequent condition. However, the three-way interaction was not significant in the standard Stroop experiment, and was only marginally significant in the semantic Stroop experiment. To increase the power to detect the three-way interaction, the data from both Task Set experiments were considered in a combined analysis.

When the data was combined, there was a significant SOA Proportion x SOA x Congruency interaction, $F(1,75) = 4.16$, $p = .045$ interaction, which did not vary as a function of Experiment, which was included as a between-subjects factor ($F < 1$ for the four-way interaction). The Congruency effect was under-additive with decreasing SOA in the Long Frequent condition (a decrease from 46 ms on long SOA trials to -3 ms on the short SOA trials; $F(1,75) = 9.29$, $p = .003$; the scaled JZS Bayes Factor of 8.8 favors the alternative), and additive with SOA in the Short Frequent condition (the Congruency effect was 51 ms on long SOA trials and 43 ms on short SOA trials; $F < 1$; the scaled JZS Bayes Factor of 6.8 favors the null).

Discussion

Empirical conclusions

Though the SOA Proportion x SOA interactions discussed in Chapter 3 again fit with the logic outlined in Miller et al (2009), the other results of the Task Set experiments are unexpected. When I conducted a combined analysis of the data, I found that Congruency was *additive* with SOA in the Short Frequent condition and *under-additive* with decreasing SOA in the Long Frequent condition. If the Long Frequent condition is supposed to promote serial processing, and the Short Frequent condition is supposed to promote parallel processing, then I would expect to see the opposite pattern of results; Congruency should be additive with SOA in the Long Frequent condition, and under-additive with decreasing SOA in the Short Frequent condition.

Additionally, there was a significant effect of Congruency in the case decision task in both the standard and semantic Stroop experiments. This will be revisited in the theoretical conclusions section.

Theoretical conclusions

As was mentioned in the *Performance Optimization* section of the introduction to Chapter 3, as with Miller et al's PRP logic, additivity in the Task Set paradigm could reflect the adoption of a strategy in which participants aim to improve performance by reducing their TRT. In the context of the Task Set paradigm, target processing may be put on hold until the cue has been processed to reduce the time taken to decode the task cue and respond to the required secondary task.

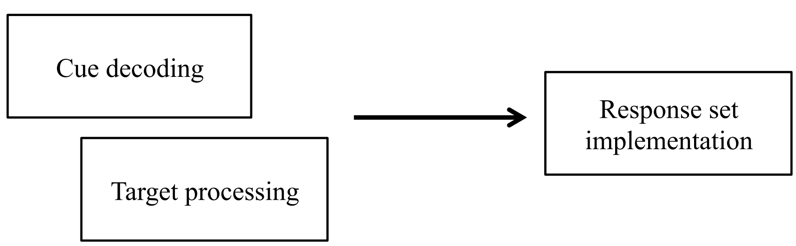
The under-additivity of Congruency and decreasing SOA in the Long Frequent condition contradicts Miller and colleagues' (2009) performance optimization account, at least at the level of semantic processing. According to Miller et al, the Long Frequent condition promotes serial processing. In other words, the cue should be decoded *before* target processing begins if participants are trying to reduce their TRT in this condition. There is, however, still evidence that supports their account in terms of overall RTs (collapsed across Congruency).

Looking at the SOA Proportion x SOA interaction, it is clear that participants are slowest to respond on short SOA trials in the Long Frequent condition. This is consistent with Miller and colleagues (2009) who attribute this slowing to strategic processing; participants are prepared to process serially in the Long Frequent condition, and are therefore slower to respond on short SOA trials than they are in the Short Frequent condition (in which they are prepared for parallel processing). It is possible, however, that this slowing reflects a disadvantage associated with an unexpected short SOA trial (as compared to a condition where the majority of trials are short SOA trials). Contrary to Miller and colleagues' hypothesis, participants may be prepared to process in parallel in the Long Frequent condition in the present experiments because of the negative effect of unexpected short SOA trials on performance in a task (Stroop) in which the word stimuli interfere with participants' goals on half of trials. Alternatively, it could be that the increase in RTs on short SOA trials in the Long Frequent condition allows the Congruency effect to be absorbed into the extra time generated by this slowing. Regardless of the reason for this absorption of the Congruency effect on short SOA trials in the Long Frequent condition, there is clearly evidence that the SOA Proportion manipulation affects the SOA x Congruency interaction. In other words, SOA Proportion influences how participants process stimuli in the Task Set paradigm.

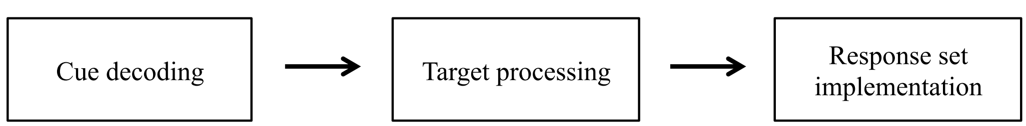
The presence of the Congruency effect in the case decision task, combined with evidence of under-additivity of Congruency and decreasing SOA for the color naming task in the Long Frequent condition, suggests that response set implementation is put on hold on short SOA trials in the Short Frequent condition until cue and target processing are complete, and that this is not due to a resource being shared by cue and target processing (as there would be no under-additivity in any condition in this case). I propose that processing unfolds as follows: on short SOA trials in the Long Frequent condition, (1) cue decoding begins, (2) target processing begins upon the presentation of the target stimulus (in parallel with the decoding of the cue; allowing for absorption of the Congruency effect), and (3) a response set is implemented once the previous two stages are complete (see Figure 13a), and on short SOA trials in the Short Frequent condition (and on long SOA trials in both conditions), (1) cue decoding occurs first, (2) followed by target processing (which begins after cue decoding is complete), and, finally, (3) a response set is implemented (once the previous two stages are complete; see Figure 13b).

Figure 13. *A depiction of the proposed order of processing in the Task Set experiments as a function of SOA and SOA Proportion.*

a) Processing on short SOA trials in the Long Frequent condition



b) Processing on short SOA trials in the Short Frequent condition (and on long SOA trials in both conditions)



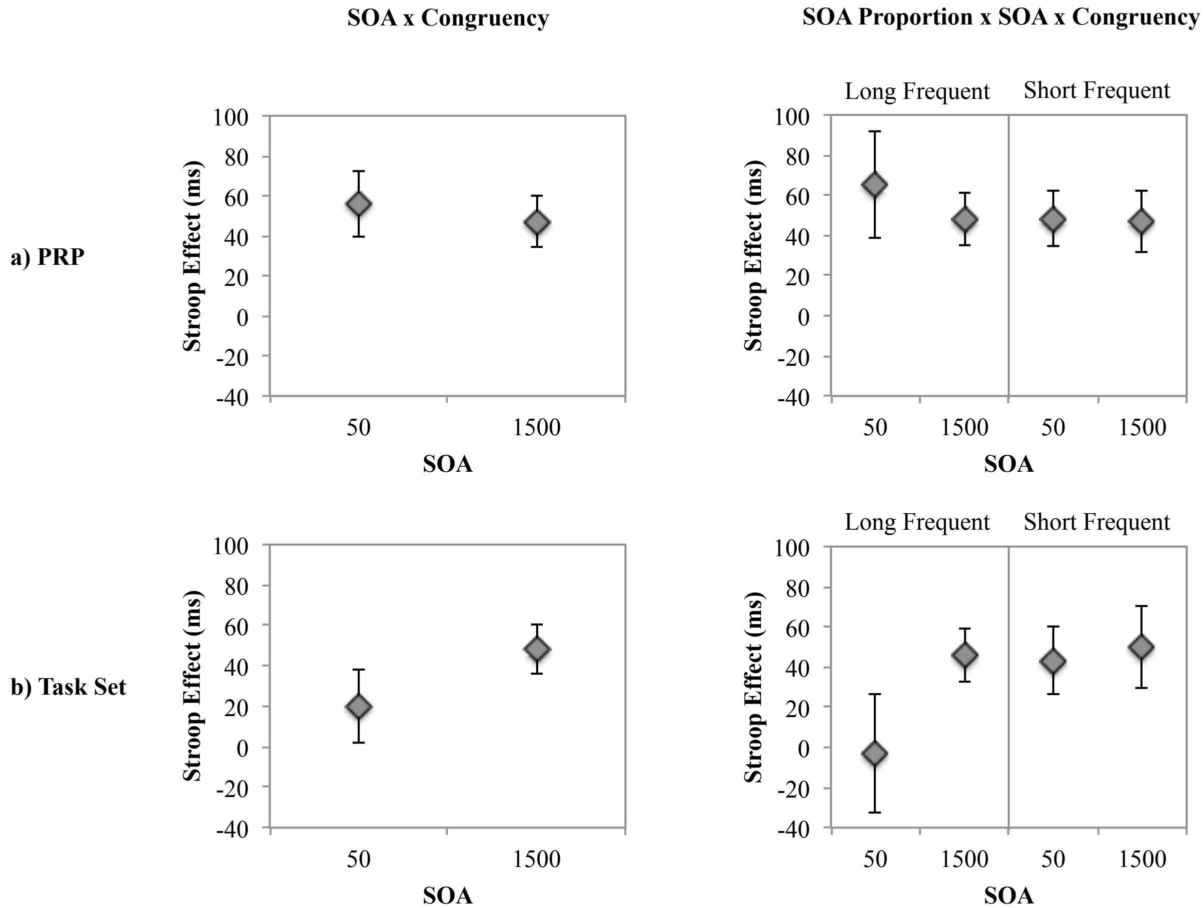
The results of the Task Set experiments suggest that processing of the target is put on hold on short SOA trials in the Short Frequent condition, which conflicts with the automatic processing perspective. Additionally, if processing is indeed not subject to interference by other processes, then the SOA Proportion manipulation should not affect the under-additivity of Congruency and decreasing SOA, yet this is what is seen in the present experiments.

It is unclear from these results, however, whether the additivity observed in the Short Frequent condition reflects performance optimization, or whether the ratio of short SOA trials is not optimal for promoting parallel processing at the level of semantics in this condition. The latter possibility could be examined in a future experiment in which the ratio of short/long SOA trials in each SOA Proportion condition are modified (e.g., Long Frequent, 90% long SOA trials and 10% short SOA trials; Short Frequent, 10% long SOA trials and 90% short SOA trials).

Cross-Paradigm Comparison

The results of the Task Set experiments differ from those obtained in the context of PRP. In the PRP experiments, SOA and Congruency were additive, and this additivity was unaffected by the SOA Proportion manipulation (see Figure 14a). In contrast, there was a significant SOA x Congruency interaction, and a significant SOA Proportion x SOA x Congruency interaction in Task Set (see Figure 14b). To determine whether the difference across these two paradigms was significant, I conducted a four-way (SOA Proportion x SOA x Congruency x Paradigm) repeated measures ANOVA (collapsed across the two Stroop manipulations).

Figure 14. *The SOA x Congruency and SOA Proportion x SOA x Congruency interaction for the PRP and Task Set experiments (collapsed across type of Stroop).*



This analysis revealed that there was significant difference between the two paradigms in terms of the SOA x Congruency interaction (the three-way including Paradigm as a between subjects factor was significant), $F(1,153) = 7.08$, $p = .009$, and the SOA Proportion x SOA x Congruency interaction (the four-way including Paradigm as a between-subjects factor was significant), $F(1,153) = 6.08$, $p = .015$.

This significant difference in the pattern of results across these two paradigms yields an interesting question. Why is there a significant difference between these two paradigms if participants are also required to decode the tone in the Task Set Paradigm prior to performing a secondary task? One hypothesis is that the difference across these two paradigms is driven by the overt response to the tone in the context of PRP. This possibility was explored in Experiments 5, 6, and 7.

Chapter 4: Does the overt response to the tone in PRP contribute to the additivity of congruency and SOA?

As was mentioned previously, it could be argued that the decoding of the task cue in the Task Set Paradigm is equivalent to a PRP "Task 1" to some degree. Participants must perform this primary task to determine which of two possible tasks they will be performing on a trial-by-trial basis. It is therefore useful to determine why the SOA Proportion manipulation influences the size of the Stroop effects in the context of the Task Set Paradigm, and not in PRP. One possibility is that the difference in results across these two paradigms is a consequence of the requirement that participants make an overt response to the tone in the context of PRP. The present chapter investigates this possibility using a modified version of the Task Set Paradigm.

Experiment 5: Standard Stroop in the Overt Response Task Set Paradigm

The Task Set experiment outlined here only differs from Experiment 3 in that it requires participants to provide an overt response to the tone prior to responding to the target stimulus. This modification will allow me to determine whether it is the overt response to the tone in PRP that yields additivity of Congruency and SOA in both SOA Proportion conditions.

Method

Participants. Using the same selection criteria as in the previous experiments, forty-eight participants from the same participant pool participated for course credit.

The **stimuli**, **design**, and **apparatus** were identical to Experiment 3.

Procedure. The procedure differed from Experiment 3 in that (1) participants were required to overtly respond to the cue tone via key press ('g' and 'h' were used as the response keys, and the key-tone correspondences were counterbalanced across participants) prior to performing the task indicated by the tone, and (2) the case decision task response was now vocal (i.e., the participant said "upper" or "lower") to avoid any potential interference caused by both the tone identification task and case decision task having manual responses (if additivity is observed in this context, it would be unclear whether it simply reflects motor based structural interference; McLeod, 1977a; 1977b; ; Miller, Ulrich, & Rolke, 2009). Participants were told to respond as quickly and accurately as possible to both tasks.

Results

One participant was removed due to a failure to follow instructions. 3.2% of trials were removed due to microphone errors, 1.8% were removed as a result of an error on the color naming task, and 0.9% were removed as a result of an error on the case decision task. Following these removals, trials with correct responses that had an RT of less than 150 ms or greater than 3000 ms were also removed, resulting in the removal of 2.5% of color naming trials and 2.2% of case decision trials. Of the remaining trials with correct responses, 1.0% were RT outliers in the color naming task and 1.3% were outliers in the case decision task. Additionally, 7 participants were error outliers and 1 participant was an RT outlier in the case decision task, and 1 participant was an RT outlier in both the case decision and the color naming task, leaving data from 38 participants for further analysis.

Task 1: Tone Identification

Response Times

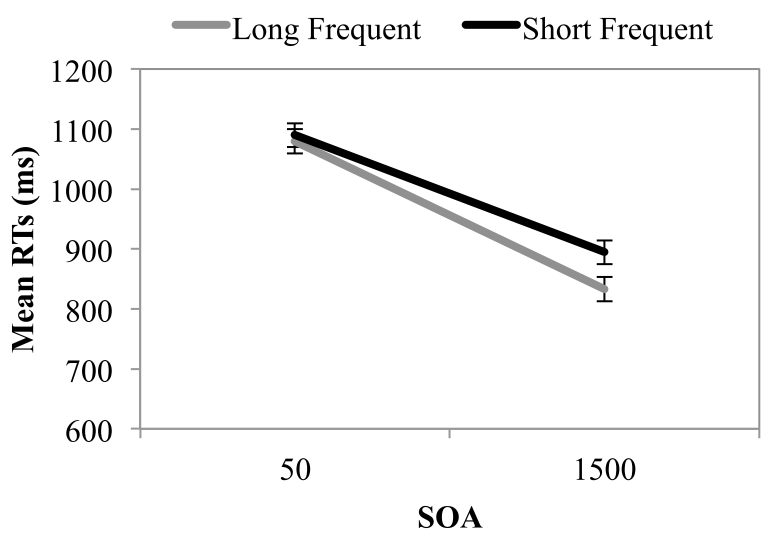
Unfortunately, it appears that participants in this experiment grouped their responses on short SOA trials. As the beginning of a new trial is triggered by a vocal response, the majority of the key presses in response to the tone on short SOA trials were not recorded. It was initially unclear as to whether this tendency was due to the initial instructions not properly emphasizing that the tone should be responded to prior to performing the second task, or whether the addition of an overt response to the cue was too difficult for participants. I will revisit this issue in Experiment 7.

Task 2: Color Naming Task

Response Times

There was a significant SOA Proportion x SOA interaction, $F(1,37) = 6.61$, $p = .014$. The results were consistent with Miller et al.'s (2009) expected findings for RT2, and replicate their Experiment 2 findings. Color naming responses times were more affected by the SOA manipulation in the Long Frequent condition than in the Short Frequent condition. In this case, however, as is shown in Figure 15, mean RT was faster for long SOA trials in the Long Frequent condition (*M* = 832), as compared to the Short Frequent condition (*M* = 894), and was virtually the same size for short SOA trials in both SOA Proportion conditions (*M* = 1079 in the Long Frequent condition; *M* = 1089 in the Short Frequent condition). Given the results of Experiments 1-4, the lack of a difference for short SOA trials is surprising. Perhaps the difficulty of the combined paradigms caused participants to adopt a different strategy than in previous experiments, which involved adapting to long SOA trials as opposed to short SOA trials.

Figure 15. *SOA Proportion x SOA interaction for the Color Naming Task in Experiment 5 (collapsed across congruency). Bars represent 95% within-subjects confidence intervals.*



A three-way (SOA Proportion x SOA x Congruency) ANOVA yielded a significant main effect of SOA, $F(1,37) = 97.75$, $p < .001$, participants were slower to respond on short SOA trials (*M* = 1071) than on long SOA trials (*M* = 857), and a significant main effect of Congruency,

$F(1,37) = 39.84$, $p < .001$ participants were slower to respond on incongruent trials ($M = 1003$) than on neutral trials ($M = 926$). There was no significant main effect of SOA Proportion, $F(1,37) = 2.13$, $p = .152$.

There was no significant SOA Proportion x Congruency interaction ($F < 1$), SOA x Congruency interaction, $F(1,37) = 1.23$, $p = .276$, or SOA Proportion x SOA x Congruency interaction ($F < 1$). The Scaled JZS Bayes factor (3.3) for the SOA x Congruency interaction positively favors the null hypothesis of no interaction. Congruency and SOA are additive. The RTs and percent error for the color naming task are shown in Table 9.

Table 9. *Color naming Task RTs and percent error by condition for Experiment 5.*

	RTs				**% Error**			
	Long Frequent		**Short Frequent**		**Long Frequent**		**Short Frequent**	
	50	**1500**	**50**	**1500**	**50**	**1500**	**50**	**1500**
Incongruent	1104	870	1108	928	3.1	2.0	2.2	2.0
Neutral	1032	780	1040	851	0.7	1.1	0.8	1.5
Difference	**72**	**90**	**68**	**77**	**2.4**	**0.9**	**1.4**	**0.5**

Errors

There was a significant main effect of Congruency on color naming errors, $F(1,37) = 14.56$, $p < .001$. Participants made more errors on incongruent trials ($M = 2.3$) than on neutral trials ($M = 1.0$). There were no other main effects on color naming errors (Fs < 1).

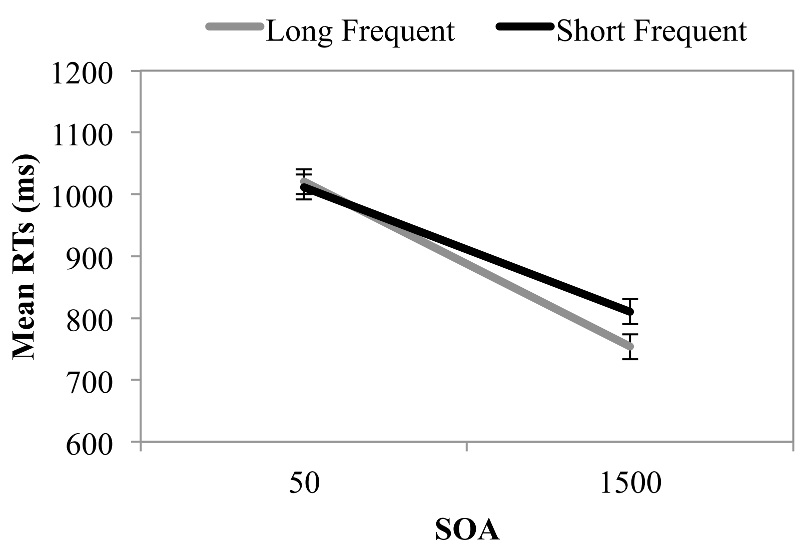
There was no significant SOA Proportion x SOA interaction ($F < 1$), SOA Proportion x Congruency interaction, $F(1,37) = 1.58$, $p = .216$, SOA x Congruency interaction, $F(1,37) = 2.43$, $p = .128$, or SOA Proportion x SOA x Congruency interaction ($F < 1$).

Task 2: Case Decision Task

Response Times

As was the case for the color naming task, there was a significant SOA Proportion x SOA interaction on case decision response times, $F(1,37) = 9.66$, $p = .004$. Case decision responses times were more affected by the SOA manipulation in the Long Frequent condition than in the Short Frequent condition (see Figure 16). Once again, mean RT was faster for long SOA trials in the Long Frequent condition ($M = 753$), as compared to the Short Frequent condition ($M = 810$), and was almost the same size in both SOA Proportion conditions for short SOA trials ($M = 1020$ in the Long Frequent condition and $M = 1011$ in the Short Frequent condition).

Figure 16. *SOA Proportion x SOA interaction for the Case Decision Task in Experiment 5 (collapsed across congruency). Bars represent 95% within-subjects confidence intervals.*



There was a significant main effect SOA, $F(1,37) = 124.73$, $p < .001$, participants were slower to respond on short SOA trials ($M = 1016$) than on long SOA trials ($M = 782$), and a significant main effect of Congruency, $F(1,37) = 37.62$, $p < .001$, participants were slower to respond on incongruent trials (*M* = 921) than on neutral trials (*M* = 877). There was no significant main effect of SOA Proportion, and none of the other interactions were significant (*F*s < 1). RTs and percent error in the case decision task are presented in Table 10.

Table 10. *Case Decision Task RTs and percent error by condition for Experiment 5.*

	RTs				**% Error**			
	Long Frequent		**Short Frequent**		**Long Frequent**		**Short Frequent**	
	50	**1500**	**50**	**1500**	**50**	**1500**	**50**	**1500**
Incongruent	1040	773	1040	830	1.4	0.8	0.8	0.5
Neutral	1000	733	983	790	0.0	0.6	0.7	0.4
Difference	**40**	**40**	**57**	**40**	**1.4**	**0.2**	**0.1**	**0.1**

Errors

There was a marginally significant main effect of Congruency, $F(1,37) = 4.00$, $p = .053$, and a marginally significant SOA Proportion x Congruency interaction, $F(1,37) = 3.34$, $p = .076$, on errors in the case decision task. Neither of the other two main effects were significant (*F*s < 1), nor was the SOA Proportion x SOA interaction ($F < 1$), SOA x Congruency interaction, $F(1,37) = 2.82$, $p = .102$, or SOA Proportion x SOA x Congruency interaction, $F(1,37) = 2.09$, $p = .157$.

Congruency was additive with SOA in both the color naming task (Scaled JZS Bayes factor positively favors the null, 3.3) and the case decision task (Scaled JZS Bayes Factor positively favors the null, 4.8). The presence of the Congruency effect for *both tasks*, combined with the lack of under-additivity with decreasing SOA in any condition, suggests that semantic processing was put on hold due to capacity limitations, and that participants only implemented a response set *following* cue and target processing (Besner & Care, 2003). Further, these results imply that the capacity limitations observed for semantic processing in the context of PRP are due to the overt response to Task 1.

Experiment 6: Semantic Stroop in the Overt Response Task Set Paradigm

Experiment 6 was identical to Experiment 5, with the exception that semantic Stroop was used.

Method

Participants. Forty-eight participants from the same participant pool as the previous experiments participated for course credit. The selection criteria were the same as in earlier experiments.

Stimuli. The stimuli only differed from Experiment 5, in that the incongruent stimuli were color-associated words (*tomato, sky, frog, lemon*) instead of color words (*red, blue, green, yellow*).

The **design**, **apparatus**, and **procedure** were identical to Experiment 5.

Results

Four participants were removed prior to analysis due to a failure to follow instructions. 0.7% of trials were removed due to microphone errors, 1.7% were removed as a result of an error on the color naming task, and 1.9% were removed as a result of an error on the case decision task. Following these removals, trials with correct responses that had an RT of less than 150 ms or greater than 3000 ms were also removed (1.2% of color naming trials and 0.9% of case decision trials). Of the remaining trials with correct responses, 1.2% were RT outliers in the color naming task and 1.2% were outliers in the case decision task. At the participant level, 5 were error outliers in the case decision task and 1 was an RT outlier in both the case decision and color naming task, leaving data from 38 participants for further analysis.

Task 1: Tone Identification

Response Times

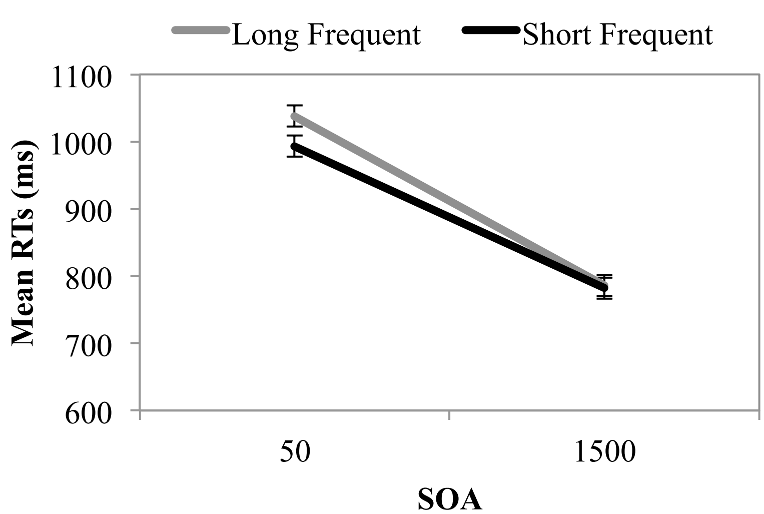
Like what was found in Experiment 5, the majority of Task 1 responses were not recorded, suggesting that the issue with response grouping was also present in the semantic version of the overt response Task Set experiment.

Task 2: Color Naming Task

Response Times

There was a significant SOA Proportion x SOA interaction on color naming response times, $F(1,37) = 4.49$, $p = .041$. Response times were again slower on the short SOA trials in the Long Frequent condition ($M = 1038$) than in the Short Frequent condition ($M = 993$). These results are shown in Figure 17.

Figure 17. *SOA Proportion x SOA interaction for the Color Naming Task in Experiment 6 (collapsed across congruency). Bars represent 95% within-subjects confidence intervals.*



Additionally, there was a significant main effect of SOA, $F(1,37) = 141.26$, $p < .001$, participants were slower to respond on short SOA trials ($M = 1015$) than on long SOA trials ($M = 784$), and a marginally significant main effect of Congruency, $F(1,37) = 3.88$, $p = .056$, participants were slower to respond on incongruent trials ($M = 909$) than on neutral trials ($M = 890$).

The effect of SOA Proportion was in the expected direction, participants were slower to respond in the Long Frequent condition ($M = 912$) than the Short Frequent condition ($M = 887$), but this difference was not significant, $F(1,37) = 2.53$, $p = .120$.

Mirroring the results of Experiment 5, there was no significant SOA x Congruency interaction ($F < 1$); the Scaled JZS Bayes Factor positively favors the null hypothesis of no interaction (3.3). There was also no significant SOA Proportion x Congruency interaction, $F(1,37) = 1.54$, $p = .223$, or SOA Proportion x SOA x Congruency interaction, $F(1,37) = 1.53$, $p = .225$.

Table 11. *Color Naming Task RTs and percent error by condition for Experiment 6.*

	RTs				**% Error**			
	Long Frequent		**Short Frequent**		**Long Frequent**		**Short Frequent**	
	50	**1500**	**50**	**1500**	**50**	**1500**	**50**	**1500**
Incongruent	1061	790	996	787	2.4	1.0	1.5	0.9
Neutral	1015	781	989	776	2.8	1.1	1.5	1.9
Difference	**46**	**9**	**7**	**11**	**-0.4**	**-0.1**	**0.0**	**-1.0**

Errors

There was a significant main effect of SOA on color naming errors, $F(1,37) = 6.99$, $p = .012$, in that participants made more errors on short SOA trials ($M = 2.1$) than on long SOA trials

(M = 1.1), and a marginally significant SOA Proportion x SOA interaction, $F(1,37) = 4.01$, p = .053.

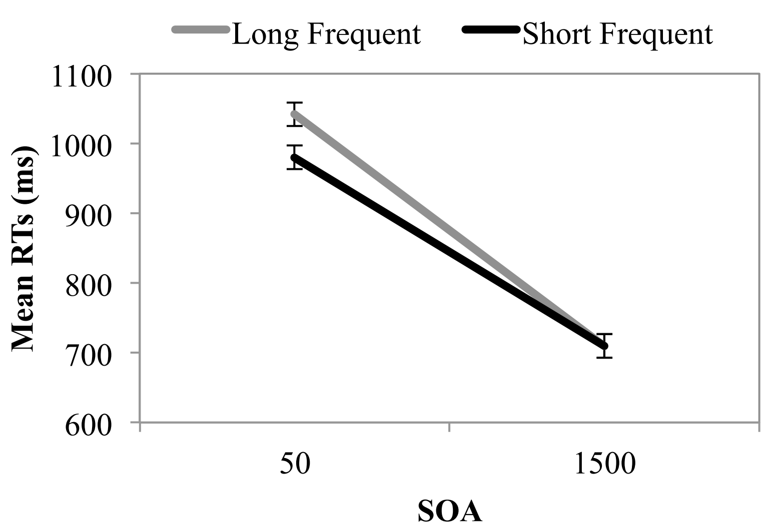
There was no significant main effect of SOA Proportion, $F(1,37) = 1.84$, $p = .184$, or Congruency, $F(1,37) = 1.99$, $p = .167$, on color naming errors, nor was there a significant SOA Proportion x Congruency interaction, ($F < 1$), SOA x Congruency interaction, ($F < 1$), or SOA Proportion x SOA x Congruency interaction, $F(1,37) = 2.06$, $p = .160$.

Task 2: Case Decision Task

Response Times

As is shown in Figure 18, there was a significant SOA Proportion x SOA interaction, $F(1,37) = 8.73$, $p = .005$. Response times were slower on the short SOA trials in the Long Frequent condition (M = 1042) than on short SOA trials in the Short Frequent condition (M = 980).

Figure 18. *SOA Proportion x SOA interaction for the Case Decision Task in Experiment 6 (collapsed across congruency). Bars represent 95% within-subjects confidence intervals.*



A three factor (SOA Proportion x SOA x Congruency) repeated measures ANOVA revealed a marginally significant main effect of SOA Proportion, $F(1,37) = 3.76$, $p = .060$. Participants were slower to respond in the Long Frequent condition ($M = 875$) than in the Short Frequent condition ($M = 844$). There was a significant main effect of SOA, $F(1,37) = 196.19$, $p < .001$. Participants were slower to respond on short SOA trials ($M = 1011$) than on long SOA trials ($M = 709$). There was no significant effect of Congruency, $F(1,37) = 1.92$, $p = .174$.

There was no significant SOA Proportion x Congruency interaction, ($F < 1$) , SOA x Congruency interaction $F(1,37) = 1.07$, $p = .307$, or SOA Proportion x SOA x Congruency interaction, $F(1,37) = 1.78$, $p = .190$.

Table 12. *Case Decision Task RTs and percent error by condition for Experiment 6.*

	RTs				**% Error**			
	Long Frequent		**Short Frequent**		**Long Frequent**		**Short Frequent**	
	50	**1500**	**50**	**1500**	**50**	**1500**	**50**	**1500**
Incongruent	1043	715	995	708	0.6	1.7	1.7	2.2
Neutral	1040	703	964	710	0.9	1.5	1.6	1.3
Difference	**3**	**12**	**31**	**-2**	**-0.3**	**0.2**	**0.1**	**0.9**

Errors

There was no significant main effect of SOA Proportion, $F(1,37) = 2.39$, $p = .130$, SOA, $F(1, 37) = 2.08$, $p = .157$, or Congruency, ($F < 1$), nor was there a significant SOA Proportion x SOA interaction, $F(1,37) = 1.42$, $p = .240$, SOA Proportion x Congruency interaction ($F < 1$), SOA x Congruency interaction ($F < 1$), or SOA Proportion x SOA x Congruency interaction ($F < 1$).

Experiment 7: Semantic Stroop in the Overt Response Task Set Paradigm 2

This experiment aimed to determine whether the response grouping evident on short SOA trials in Experiments 5 and 6 was due to the combination of the two paradigms being too difficult for participants, or whether it was a result of the instructions. If changing the instructions does fix the issue, then it will allow me to see whether the RT1 results also mirror the PRP experiments, as well as ensure that the results do not vary when Task 1 errors and outliers are removed prior to further analysis.

Method

Participants. 24 participants were gathered using the same criteria and participant pool as in previous experiments. Participants were awarded course credit in exchange for their participation.

The **design**, **apparatus**, and **stimuli** were the same as in Experiment 6.

Procedure. The procedure was identical to Experiment 6, with the exception that the initial instructions included the following, "Give priority to identifying the tone! It is the most important task." Additionally, the experiment was coded to generate an error message whenever the tone task was not responded to prior to the microphone being triggered by the vocal response to the target stimulus. The message alternated between "Give priority to the tone task! Respond to it as quickly as possible!" and "Wrong Order".

Results

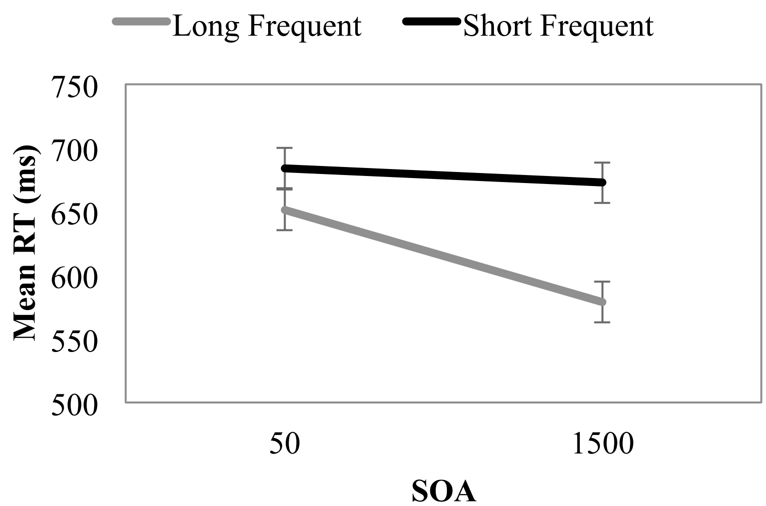
Three participants did not complete the experiment and 2 were removed prior to analysis due to a failure to follow instructions, leaving 19 for further analysis. 7.9% of trials were removed due to microphone errors. Following the removal of microphone errors, and 3.9% of trials on which an error was made on either Task 1 (tone identification) or the Task 2 (color naming or case decision) were removed. 3.7% of the trials with correct responses were removed because they had an RT of less than 150 ms or greater than 3000 ms. 2.2% of trials with correct responses were removed as RT outliers. One participant was an error outlier on the tone task, and one participant was an RT outlier on both the tone task and the color naming task, leaving 17 participants for analysis.

Task 1: Tone Identification

Response Times

RT1 was slower in the Short Frequent condition (M = 678) than in the Long Frequent condition (M = 615), as was reported by Miller and colleagues, $F(1,16) = 8.02$, $p = .012$. Additionally, there was a significant main effect of SOA, $F(1,16) = 16.93$, $p = .001$, participants were slower to respond on short SOA trials (M = 667) than on long SOA trials (M = 626), and a significant SOA Proportion x SOA interaction, $F(1,16) = 11.85$, $p = .003$ (see Figure 19), which is consistent with Miller and colleagues' (2009) Experiment 2 results for RT1.

Figure 19. *SOA Proportion x SOA interaction for Task 1 in Experiment 7. Bars represent 95% within-subjects confidence intervals.*



A three-way SOA Proportion (Long Frequent vs. Short Frequent) x SOA (Short vs. Long) x Congruency (Incongruent vs. Neutral) ANOVA revealed a significant main effect of Congruency, $F(1,16) = 6.91$, $p = .018$. There was a marginally significant SOA Proportion x Congruency interaction, $F(1,16) = 3.83$, $p = .068$, and no significant SOA x Congruency interaction ($F < 1$). The SOA Proportion x SOA x Congruency interaction was significant, $F(1,16) = 7.65$, $p = .014$ (see Table 13).

Table 13. *Task 1 mean RTs and percent error by condition for Experiment 7.*

	RTs				**% Error**			
	Long Frequent		**Short Frequent**		**Long Frequent**		**Short Frequent**	
	50	**1500**	**50**	**1500**	**50**	**1500**	**50**	**1500**
Incongruent	652	577	695	694	1.0	1.0	1.6	0.6
Neutral	649	580	673	650	2.5	0.7	1.5	0.3
Difference	**3**	**-3**	**22**	**44**	**-1.5**	**0.3**	**0.1**	**0.3**

Errors

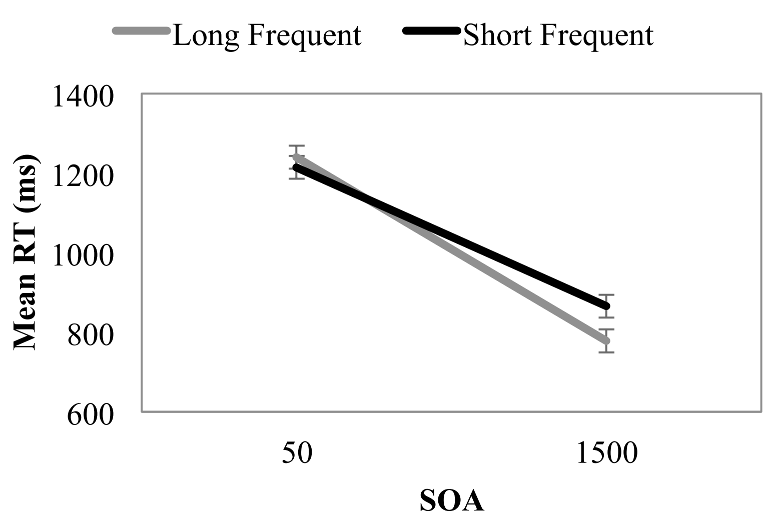
The main effect of SOA on Task 1 errors was significant, $F(1,16) = 10.12$, $p = .006$. Participants made more errors on short SOA trials ($M = 1.7$) than on long SOA trials ($M = 0.6$). There was no significant main effect of SOA Proportion, $F(1,16) = 1.34$, $p = .264$, or Congruency on Task 1 errors ($F < 1$), nor were there any significant interactions on Task 1 errors (Fs < 1).

Task 2: Color Naming Task

Response Times

There was a significant SOA Proportion x SOA interaction on color naming response times, $F(1,16) = 13.22$, $p < .001$. Response times were slower on long SOA trials in the Short Frequent condition ($M = 866$) than in the Long Frequent condition ($M = 778$), and were slower on the short SOA trials in the Long Frequent condition ($M = 1240$) as compared to the Short Frequent condition ($M = 1215$). These results are shown in Figure 20.

Figure 20. *SOA Proportion x SOA interaction for the Color Naming Task in Experiment 7 (collapsed across congruency). Bars represent 95% within-subjects confidence intervals.*



A three-way repeated measures ANOVA revealed a significant main effect of SOA, $F(1,16) = 173.58$, $p < .001$, participants were slower to respond on short SOA trials ($M = 1227$) than on long SOA trials ($M = 821$), and a marginally significant main effect of Congruency, $F(1,16) = 3.14$, $p = .096$, participants were slower to respond on incongruent trials ($M = 1040$) than on neutral trials ($M = 1008$). There was no significant main effect of SOA Proportion or SOA Proportion x Congruency interaction (*F*s < 1) on color naming RTs.

Mirroring the results of Experiments 5 and 6, there was no significant SOA x Congruency interaction ($F < 1$); the Scaled JZS Bayes Factor positively favors the null (3.9), or SOA Proportion x SOA x Congruency interaction ($F < 1$); the Scaled JZS Bayes Factor positively favors the null (3.3).

Table 14. *Color Naming Task RTs and percent error by condition for Experiment 7.*

	RTs				% Error			
	Long Frequent		**Short Frequent**		**Long Frequent**		**Short Frequent**	
	50	**1500**	**50**	**1500**	**50**	**1500**	**50**	**1500**
Incongruent	1253	797	1237	874	1.9	1.5	2.1	0.5
Neutral	1227	757	1192	856	1.5	1.2	1.3	2.2
Difference	**26**	**40**	**45**	**18**	**0.5**	**0.3**	**0.8**	**-1.7**

Errors

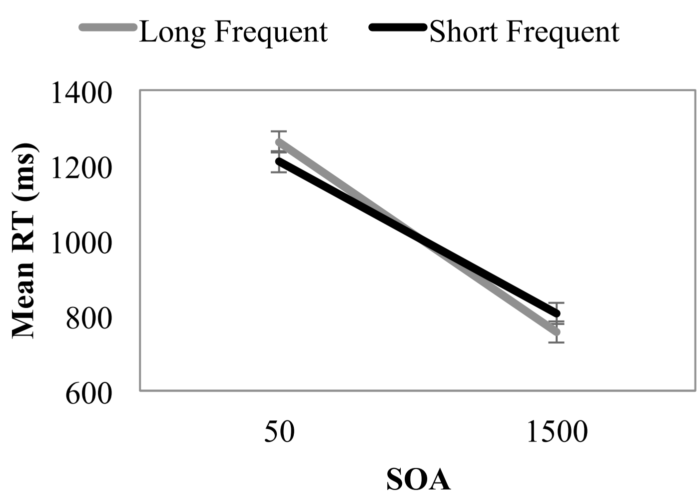
There were no significant main effects on color naming RTs ($F < 1$). There was no significant SOA Proportion x SOA interaction ($F < 1$), SOA Proportion x Congruency interaction, ($F < 1$), SOA x Congruency interaction, $F(1,16) = 2.32$, $p = .147$, or SOA Proportion x SOA x Congruency interaction, $F(1,16) = 1.85$, $p = .193$.

Task 2: Case Decision Task

Response Times

As is shown in Figure 21, there was a significant SOA Proportion x SOA interaction, $F(1,16) = 5.30$, $p = .035$. Response times were slower on the short SOA trials in the Long Frequent condition ($M = 805$) than in the Short Frequent condition ($M = 756$), and were slower on the short SOA trials in the Long Frequent condition ($M = 1262$) as compared to the Short Frequent condition (1210).

Figure 21. *SOA Proportion x SOA interaction for the Case Decision Task in Experiment 7 (collapsed across congruency). Bars represent 95% within-subjects confidence intervals.*



A three factor (SOA Proportion x SOA x Congruency) repeated measures ANOVA revealed a significant main effect of SOA, $F(1,16) = 210.14$, $p < .001$. Participants were slower to respond on short SOA trials ($M = 1235$) than on long SOA trials ($M = 780$). There was no significant main effect of SOA Proportion or Congruency (Fs < 1).

There was no significant SOA Proportion x Congruency interaction, ($F < 1$) , SOA x Congruency interaction ($F < 1$), or SOA Proportion x SOA x Congruency interaction, $F(1,16) = 1.25$, $p = .279$.

Table 15. *Case Decision Task RTs and percent error by condition for Experiment 7.*

	RTs				% Error			
	Long Frequent		**Short Frequent**		**Long Frequent**		**Short Frequent**	
	50	**1500**	**50**	**1500**	**50**	**1500**	**50**	**1500**
Incongruent	1265	751	1207	823	5.3	1.1	1.5	0.5
Neutral	1258	761	1212	788	3.4	1.5	2.0	0.0
Difference	**7**	**-10**	**-5**	**35**	**1.9**	**-0.4**	**-0.5**	**0.5**

Errors

There was a significant main effect of SOA Proportion, $F(1,16) = 8.62$, $p = .010$, in which participants made more errors in the Long Frequent condition ($M = 2.8$) than the Short Frequent condition ($M = 1.0$). There was also a significant main effect of SOA, $F(1,16) = 10.12$, $p = .006$, in which participants made more errors in the short SOA trials ($M = 3.1$) than on long SOA trials ($M = 0.8$). There was no significant main effect of Congruency ($F < 1$), SOA Proportion x SOA interaction, $F(1,16) = 1.96$, $p = .181$, SOA Proportion x Congruency interaction ($F < 1$), SOA x Congruency interaction ($F < 1$), or SOA Proportion x SOA x Congruency interaction, $F(1,16) = 1.58$, $p = .227$.

Combined analysis of the overt response Task Set Experiments

As with the previous sets of experiments, a combined analysis of the overt Task Set experiments was conducted.

SOA and Congruency were additive. There was no significant SOA x Congruency interaction ($F < 1$), and the Scaled JZS Bayes Factor (8.3) favored the null. The SOA x Congruency interaction did not vary significantly as a function of type of Stroop, $F(1,92) = 2.76$, $p = .100$. The SOA Proportion x SOA x Congruency interaction was also not significant (Fs < 1, the Scaled JZS Bayes Factor of 8.7 favored the null). This interaction did not vary as a function of Stroop Type ($F < 1$).

Discussion

Empirical Conclusions

For both standard Stroop (Experiment 5) and semantic Stroop (Experiments 6 and 7), the SOA Proportion x SOA interaction was significant for both the color naming task and the case decision task. These results are in line with Miller and colleagues' (2009) performance optimization account, which suggests that responses on long SOA trials will be shorter in the Short Frequent condition than the Long Frequent condition if participants are preparing for parallel processing in the former case and serial in the latter. However, it should be noted that the pattern in Experiments 5 and 7 differed from previous experiments. There was a greater difference in RTs (collapsed across Congruency) on long SOA trials as a function of SOA Proportion in Experiments 5 and 7. In contrast, the difference in RTs was greater on *short SOA* trials as a function of SOA Proportion in the previous experiments. This difference could reflect a shift in strategy in the present context to one that has a greater effect on long SOA trials.

There was no significant SOA x Congruency interaction in either Experiment 5 or Experiment 6, and the Bayes factor for both of these comparisons positively favored the null, suggesting that these two factors were additive. One issue with both Experiment 5 and Experiment 6, however, was that the tone task responses were not coded on the majority of short SOA trials. Experiment 7 was conducted to determine whether this issue was being driven by task difficulty or experiment instructions.

Experiment 7 was identical to Experiment 6, with the exception that a greater emphasis was placed on responding to Task 1 (tone identification) prior to responding to Task 2 (color naming or case decision). With the new instructions, participants gave priority to responding to the tone prior to the target stimulus for Task 2. Importantly, the results for Task 1 were consistent with Miller et al's (2009), and the removal of tone task outlier trials and participants did not alter the form of the SOA x Congruency interaction. Once again, SOA and Congruency were additive factors.

Theoretical Conclusions

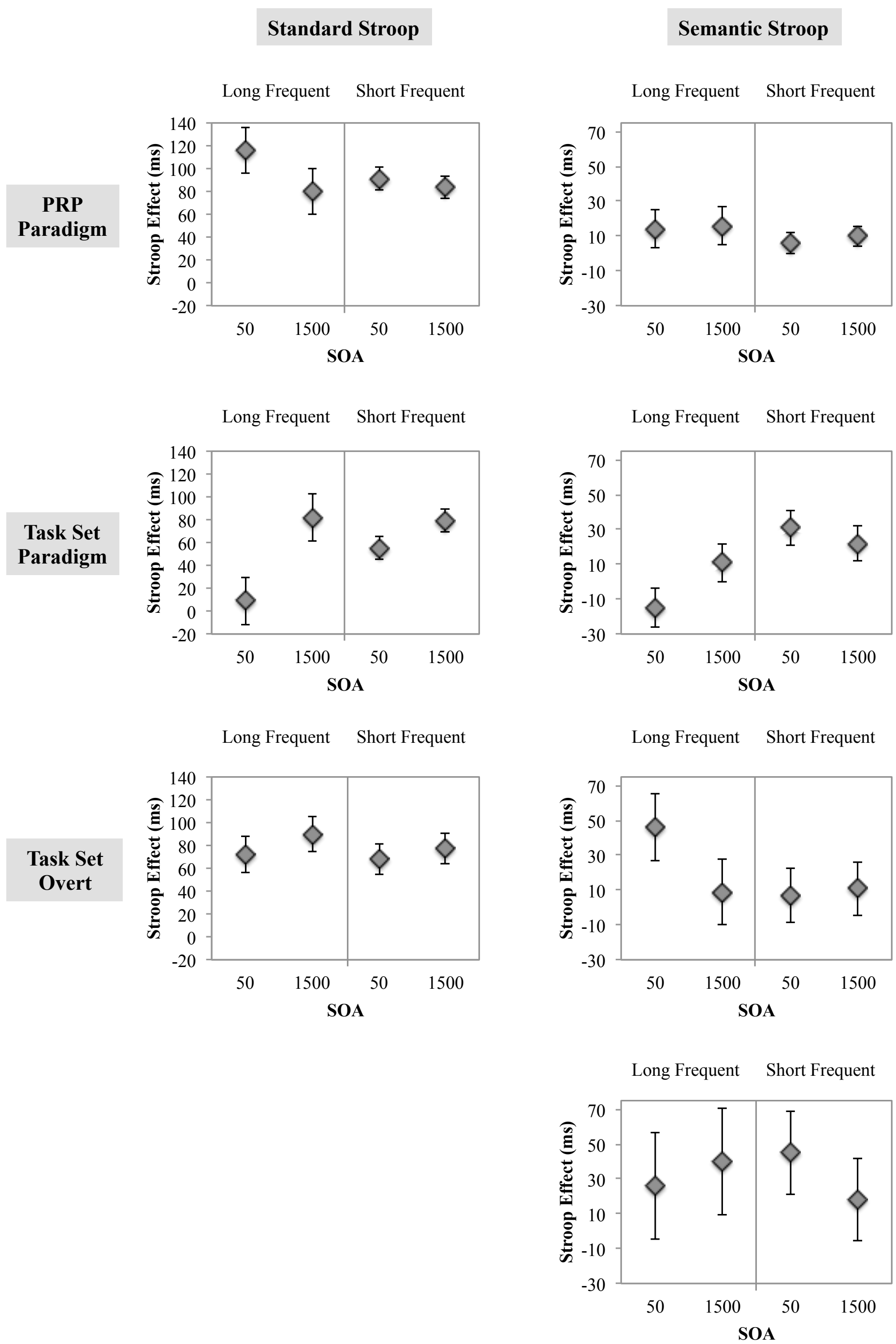
The results of the three overt response Task Set experiments are consistent with Miller and colleagues' (2009) performance optimization account; the SOA Proportion x SOA interaction was significant for Task 1 and both secondary tasks. For Experiments 5 and 7, this pattern of results differed from previous experiments, in that the SOA Proportion manipulation had more of an effect on long SOA trials than short SOA trials. The different results for the SOA Proportion x SOA interaction could be driven by a change in strategy as a result of the increased difficulty of the combined paradigms.

Both the SOA Proportion x SOA x Congruency interaction and the SOA x Congruency interaction were not significant in any of the three overt response Task Set experiments. These results mirror those obtained in the PRP paradigm experiments, suggesting that the bottleneck on semantic processing in Experiments 1 and 2 results from the overt response to the tone.

General Discussion

I have covered considerable ground here, so a summary of the results across these 7 experiments is useful (see Figure 22).

Figure 22**.** *A summary of the SOA x Congruency interactions in the color naming tasks.*



Experiment 1 examined standard Stroop in the context of PRP, and Experiment 2 semantic Stroop. The inclusion of an SOA Proportion manipulation aimed to investigate whether performance optimization plays a role in this context (as suggested by Miller et al, 2009) and whether there would be evidence of performance optimization at the level of lexical (standard Stroop) and semantic (semantic Stroop) processing. The significant SOA Proportion x SOA interaction on RTs (collapsed across Congruency) for Task 1 and Task 2 in both experiments supported Miller and colleagues' performance optimization account. However, the SOA Proportion manipulation did not modulate the size of the SOA x Congruency interaction. The SOA x Congruency interaction was in the wrong direction in the standard Stroop experiment (over-additive rather than additive, with a Scaled JZS Bayes factor of 1.2 in favor of the null that there is no interaction), and was additive in the semantic Stroop experiment (the Scaled JZS Bayes factor of 5.5 positively favored the null that there is no interaction). These results (a replication of Miller et al's results, combined with no effect of SOA Proportion on the size of the Congruency effect as a function of decreasing SOA) suggest that *semantic* activation from print is bottlenecked, and that some other factors may be subject to performance optimization.

An important question that followed from the first two experiments was whether performance optimization also plays a role in Task Set experiments. Following Miller et al's (2009) logic, it is conceivable that participants in the Task Set paradigm are also adopting a strategy that aims to reduce their TRT. This possibility was explored in Experiment 3 (standard Stroop) and Experiment 4 (semantic Stroop). As would be expected following a performance optimization account, the SOA Proportion x SOA interaction was significant for both color naming trials and case decision trials. This pattern of results is consistent with Miller et al's, suggesting that performance optimization plays a role in this context. In contrast to the PRP experiments, however, the SOA Proportion manipulation influenced the SOA x Congruency interaction. Congruency was *under-additive* with decreasing SOA in the Long Frequent condition (the scaled JZS Bayes Factor of 8.8 positively favored the alternative), and *additive* with SOA in the Short Frequent condition (the scaled JZS Bayes Factor of 6.8 positively favored the null). This result was surprising, as Miller and colleagues would predict the reverse; that the Short Frequent condition should promote parallel processing and the Long Frequent condition should promote serial processing. A potential explanation for this pattern of results is that participants prepare for parallel processing in the Long Frequent condition to reduce the disadvantage on performance that results from an unexpected short SOA trial. Alternatively, it could be that the unexpected short SOA trials are slowing participants down enough to allow the effect of Congruency to be absorbed into the extra time that is generated.

In the Task Set experiments, the *presence* of the three-way SOA Proportion x SOA x Congruency interaction, along with *under-additivity* of Congruency and SOA in the Long Frequent condition, and *additivity* of SOA and Congruency in the Short Frequent condition implies that even though semantic processing is able to unfold at the same time as the cue is being decoded, it *can* be put on hold under certain conditions. Additionally, Stroop effects were present in case decision RTs in both of the Task Set experiments. This implies that response set implementation is put on hold until the cue and target have both been processed. Combined, these results demonstrate that semantic processing *can* unfold at the same time as the cue is being decoded, but that whether processing occurs in parallel varies as a function of other factors (e.g., attempts to optimize performance).

The results of the Task Set experiments open the door to some interesting possibilities. Perhaps some processes that appear intention free could in fact be put on hold under the right conditions. If the SOA Proportion manipulation affects whether participants aim to optimize performance by putting processing on hold, then it is possible that some effects that have been shown to be under-additive with decreasing SOA in a typical Task Set experiment, such as word frequency (O'Malley & Besner, 2011), could be additive in a condition in which there are more short SOA trials than long SOA trials. Relatedly, it is possible that effects that have been shown to be additive in Task Set experiments could be made under-additive as a result of SOA Proportion manipulations.

The different patterns of results obtained in these two paradigms yielded an interesting follow-up question. If cue decoding in the Task Set paradigm is equivalent to a Task 1, then why was Congruency under-additive in the Long Frequent condition in this context? Did this difference result from task uncertainty or the lack of an overt response to the tone? Experiments 5, 6, and 7 addressed these questions by looking at whether adding an overt response to the cue in the context of Task Set would yield additivity of SOA and Congruency. This was indeed the case; neither the SOA Proportion x SOA x Congruency interaction, or the SOA x Congruency interaction were significant. SOA and Congruency were additive in all three experiments, and the Scaled JZS Bayes Factors positively favored the null (Experiment 5, 3.3; Experiment 6, 3.3; Experiment 7, 3.9). It seems clear from these results that semantic processing is bottlenecked by the overt response to the tone in the Task Set paradigm. The tone must also be decoded in Task Set, and a decision about which task to perform must be made based on the information gained from the tone, yet the results differ when an overt response is selected. This difference is likely driven by the need to select and execute a response to the tone task prior to performing the secondary task.

The results of the three sets of experiments outlined in my dissertation are straightforward: contrary to the standardly held view, semantic processing is not automatic. However, it is useful to bear in mind that it is important to define "automaticity" whenever experimental work is carried out and theoretical claims are made. The conclusions drawn here (re: automatic processing) apply to the issue of capacity limitations, intention, and performance optimization. Other questions, such as whether spatial attention is involved in earlier processes, whether earlier processing (e.g., lexical) can be interfered with by other processes, whether sublexical phonological processing (as opposed to lexical processing) is capacity limited, are distinct issues that have been the subject of other investigations (e.g., Besner et al., 2005; Labuschagne & Besner, 2015; Lachter et al., 2004; Lien et al., 2010; O'Malley et al., 2008; Reynolds & Besner, 2006; Robidoux & Besner, 2015; Ruthruff, Allen, Lien, & Grabbe, 2008). The emerging story is that various aspects of visual word recognition are not automatic in many ways, but certainly not all ways (e.g., for a brief review see Besner, Risko, Stolz, White, Reynolds, O'Malley & Robidoux, 2016).

To the future

It would be useful if further investigations looked more extensively at (a) skilled versus less skilled readers with regard to what forms of attention limit performance, and (b) to what extent other measures of performance (e.g., fMRI/ ERP's) converge with the results of behavioural measures, or provide evidence for automaticity.

References

Ashcraft, M. H., & Klein, R. (2010). *Human Memory and Cognition*. New York, NY: Harper Collins.

Augustinova, M., & Ferrand, L. (2012). Suggestion does not de-automatize word reading: evidence from the semantically based Stroop task. *Psychonomic Bulletin & Review*, *19*, 521-527. doi: 10.3758/s13423-012-0217-y

Augustinova, M., & Ferrand, L. (2014). Automaticity of word reading: Evidence from the semantic Stroop paradigm. *Current Directions in Psychological Science*, *23*(5), 343-348. doi: 10.1177/0963721414540169

Bargh, J. A. (1989). Conditional automaticity: Varieties of automatic influence in social perception and cognition. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended Thought* (3-51). New York, NY: Guilford Press.

Baror, S. & Bar, M. (2016). Associative activation and its relation to exploration and exploitation in the brain. *Psychological Science*, *27*(6), 776-789. doi: 10.1177/0956797616634487

Besner, D. (2001). The myth of ballistic processing: Evidence from Stroop's paradigm. *Psychonomic Bulletin & Review*, *8*(2), 324-330. doi: 10.3758/BF03196168

Besner, D., & Care, S. (2003). A paradigm for exploring what the mind does while deciding what it should do. *Canadian Journal of Experimental Psychology*, *57*(4), 311-320. doi: 10.1037/h0087434

Besner, D., & Reynolds, M. (2016). Is semantic activation from print capacity limited? Evidence from the psychological refractory period paradigm. *Psychonomic Bulletin & Review*. doi: 10.3758/s13423-016-1178-3

Besner, D., Reynolds, M., & O'Malley, S. (2009). When under-additivity of factor effects in the psychological refractory period paradigm implies a bottleneck: Evidence from psycholinguistics. *The Quarterly Journal of Experimental Psychology*, *62*(11), 2222-2234.

Besner, D., Risko, E. F., & Sklair, N. (2005). Spatial attention as a necessary preliminary to early processes in reading. *Canadian Journal of Experimental Psychology*, *59* (2), 99-108. doi: 10.1037/h0087465

Besner, D., Risko, E. F., Stolz, J. A., White, D., Reynolds, M., O'Malley, S., & Robidoux, S. (in press). Varieties of attention: Their roles in visual word identification. *Current Directions in Psychological Science*. doi: 10.1177/0963721416639351

Besner, D., & Stolz, J. A. (1999). What kind of attention modulates the Stroop effect? *Psychonomic Bulletin & Review*, *6*(1), 99-104. doi: 10.3758/BF03210815

Brown, T. L., Gore, C. L., & Carr, T. H. (2002). Visual attention and word recognition in Stroop color naming: Is word recognition "automatic"? *Journal of Experimental Psychology*, *131*, 220-240. doi: 10.1037/0096-3445.131.2.220

Bryan, W. L., & Harter, N. (1899). Studies on the telegraphic language. The acquisition of a hierarchy of habits. *The Psychological Review*, *6*(4), 345-375.

Cattell, J. M. (1886). The time it takes to see and name objects. *Mind*, *11*(41), 63-65.

Dalrymple-Alford, E. C. (1972). Associative facilitation and interference in the Stroop color-word task. *Perception & Psychophysics*, *11*(4), 274-276.

Eysenck, M. W., & Keane, M. T. (2015). *Cognitive Psychology: A Student's Handbook*. New York, NY: Taylor & Francis.

Fagot, C., & Pashler, H. (1992). Making two responses to a single object: Implications for the central attentional bottleneck. *Journal of Experimental Psychology: Human Perception and Performance*, *18*(4), 1058-1079.

Galotti, K. M., Fernandes, M. A., Fugelsang, J., & Stolz, J. (2010). *Cognitive Psychology: In and Out of the Laboratory.* Toronto: Nelson Education Ltd.

Goldstein, E. B. (2011). *Cognitive Psychology: Connecting Mind, Research, and Everyday Experience*. Belmont, CA: Wadsworth Cengage Learning.

Heyman, T., Van Rensbergen, B., Storms, G., Hutchison, K. A., & De Deyne, S. (2015). The influence of working memory load on semantic priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *41*(3), 911-920. http://dx.doi.org/10.1037/xlm0000050

James, W. (1890). *The Principles of Psychology*. New York: Holt, Rinehart & Winston.

Klein, G. S. (1964). Semantic power measured through the interference of words with color-Naming. *The American Journal of Psychology*, *77*(4), 576-588.

Laberge, D., & Samuels, J. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology*, *6*, 293-323. doi: https://doi.org/10.1016/0010-0285(74)90015-2

Labuschagne, E. M., & Besner, D. (2015). Automaticity revisited: When print doesn't activate semantics. *Frontiers in Psychology*, *6*, 1-7. doi: 10.3389/fpsyg.2015.00117

Lachter, J., Forster, K. I., & Ruthruff, E. (2004). Forty-five years after Broadbent (1958): Still no identification without attention. *Psychological Review*, *111*(4), 880-913. doi: 10.1037/0033-295X.111.4.880

Lien, M., Ruthruff, E., Kouchi, S., & Lachter, J. (2010). Even frequent and expected words are not identified without spatial attention. *Attention, Perception, & Psychophysics*, *72*(4), 973-988. doi: 10.3758/APP.72.4.973

Loftus, G. R., & Masson, M. E. (1994). Using confidence intervals in within-subjects designs. *Psychonomic Bulletin & Review*, *1*(4), 476-490.

Logan, G. D. (1985). Skill and automaticity: Relations, implications, and future directions. *Canadian Journal of Psychology*, *39*(2), 367-386.

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, *95*(4), 492-527.

Logan, G. D. (1997). Automaticity and reading: Perspectives from the instance theory of automatization. *Reading & Writing Quarterly*, *13*(2), 123-146. doi: 10.1080/1057356970130203

MacLeod, C. M. (1991). Half a century of research on the Stroop effect: An integrative review. *Psychological Bulletin*, *109*(2), 163-203. doi:10.1037/0033-2909.109.2.163

Magen, H., & Cohen, A. (2002). Action-based and vision-based selection of input: Two sources of control. *Psychological Research*, *66*, 247-259. doi: 10.1007/s00426-002-0099-0

Magen, H., & Cohen, A. (2010). Modularity beyond perception: Evidence from the PRP paradigm. *Journal of Experimental Psychology: Human Perception and Performance*, *36*(2), 395-414. doi: 10.1037/a0017174

Manwell, L. A., Roberts, M. A., & Besner, D. (2004). Single letter coloring and spatial cuing eliminates a semantic contribution to the Stroop effect. *Psychonomic Bulletin & Review*, *11*(3), 458-462. doi: 10.3758/BF03196595

McLeod, P. (1977). A dual task response modality effect: Support for multiprocessor models of attention. *The Quarterly Journal of Experimental Psychology*, *29*(4), 651-667. doi: 10.1080/14640747708400639

McLeod, P. (1977). Does probe RT measure central processing demand? *The Quarterly Journal of Experimental Psychology*, *30*, 83-89. doi: http://dx.doi.org/10.1080/14640747808400656

Miller, J., Ulrich, R., & Rolke, B. (2009). On the optimality of serial and parallel processing in the psychological refractory period paradigm: Effects of the distribution of stimulus onset asynchronies. *Cognitive Psychology*, *58*(3), 273-310. doi: 10.1016/j.cogpsych.2006.08.003

Moors, A., & De Houwer, J. Automaticity: A theoretical and conceptual analysis. *Psychological Bulletin*, *132*(2), 297-326. doi: 10.1037/0033-2909.132.2.297

Neely, J. H. (1977). Semantic priming and retrieval from lexical memory: Roles of inhibitionless spreading activation and limited-capacity attention. *Journal of Experimental Psychology: General*, *106*(3), 226-254.

Neely, J. H., & Kahan, T. (2001). “Is semantic activation automatic? A critical re-evaluation,” in *The Nature of Remembering: Essays in Honor of Robert G. Crowder*, H.L. III. Roediger, J. S. Nairne, I. Neath, & A. M. Surprenant (Eds.), Washington, DC: American Psychological Association, 69-93.

O’Malley, S., & Besner, D. (2011). Lexical processing while deciding what task to perform: Reading aloud in the context of the task set paradigm. *Consciousness and Cognition*, *20*, 1594-1603.

O'Malley, S., Reynolds, M. G., Stolz, J. A., & Besner, D. (2008). Reading aloud: Spelling-sound translation uses central attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(2), 422-429. doi: 10.1037/0278-7393.34.2.422

Pashler, H. (1994). Dual task interference in simple tasks: Data and theory. *Psychological Bulletin*, *116*, 220-244.

Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R. L. Solso (Ed.), *Information Processing and Cognition: The Loyola Symposium* (pp. 55-85). Hillsdale, NJ: Lawrence Erlbaum Associates.

Reisberg, D. (1997). *Cognition: Exploring the Science of the Mind*. New York, NY: W. W. Norton & Company.

Reynolds, M., & Besner, D. (2006). Reading aloud is not automatic: Processing capacity is required to generate a phonological code from print. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(6), 1303-1323. doi: 10.1037/0096-1523.32.6.1303

Risko, E. F., Schmidt, J. R., & Besner, D. (2006). Filling a gap in the semantic gradient: Color associates and response set effects in the Stroop task. *Psychonomic Bulletin & Review*, *13*(2), 310-315.

Robidoux, S., & Besner, D. (2015). Conflict resolved: On the role of spatial attention in reading and color naming tasks. *Psychonomic Bulletin & Review*, *22*(6), 1709-1716. doi: 10.3758/s13423-015-0830-7

Rouder, J. N., Speckman, P. L., Sun, D., & Morey, R. D. (2009). Bayesian *t* tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review, 16*, 225–237. doi: 10.3758/PBR.16.2.225

Ruthruff, E., Allen, P. A., Lien, M., & Grabbe, J. (2008). Visual word recognition without central attention: Evidence for greater automaticity with greater reading ability. *Psychonomic Bulletin & Review*, *15*(2), 337-343. doi: 10.3758/PBR.15.2.337

Shiffrin, R. M., Schneider, W. (1977). Controlled and automatic human information processing. II. Perceptual learning, automatic attending and a general theory. *Psychological Review*, *84*(2), 127-190.

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, *18*, 643-662.

Thorne, F. C. (1955). Habit formations. *Principals of psychological examining: A systematic textbook of applied integrative psychology*, (pp. 373-391). Brandon, VT: Journal of Clinical Psychology.

Wagenmakers, E. J. (2007). A practical solution to the pervasive problems of *p* values. *Psychonomic Bulletin & Review*, *14*(5), 779-804.

Waechter, S., Besner, D., & Stolz, J. A. (2011). Basic processes in reading: Spatial attention as a necessary preliminary to orthographic and semantic processing. *Visual Cognition*, *19*(2), 171-202. doi: 10.1080/13506285.2010.517228

White, D., & Besner, D. (2016). Attentional constraints on semantic activation: Evidence from Stroop's paradigm. *Acta Psychologica*. doi: 10.1016/j.actpsy.2016.08.008