SEARCHING FOR E(XECUTIVE CONTROL) IN THE STRENGTH MODEL OF SELF-REGULATION: AN EXAMINATION INTO THE LETTER-CROSSING TASK

A Thesis submitted by

Madeleine McKay Arber, BPsycSci(Hons)

For the award of

Doctor of Philosophy

2018

Abstract

Self-regulation is the effortful process of controlling the self in order to meet goals or standards. The strength model of self-regulation poses that the resource behind self-regulation is limited in processing capacity, resulting in failures over time (i.e. ego-depletion). This theory was generally accepted until recently, when the literature encountered a replication crisis with widespread reported difficulties in replicating the depletion effect. This led to a conceptual crisis questioning whether the effect is legitimate, and if so, what powers this limited reserve. This thesis aims to address three major problems that have arisen in the literature. Since the self-regulation reserve has not been defined beyond a global limited reserve, almost any task can then be employed to induce depletion, provided it is effortful and demanding enough. Because of this, broad definitions of self-regulation measures have been applied. These self-regulation tasks, such as the letter-crossing task, are rarely scored and analysed. Subsequently, there is no established scoring method or knowledge as to what these tasks are measuring. Following the strength model of self-regulation, which implies depletion effects increase with ongoing processing and time, depletion effects should be observable over time on the letter-crossing task and transfer onto an effortful follow-up task. Over three studies (eight experiments), performance under a modified letter-crossing task was scored, analysed, and compared to standardised executive measures (Stroop, OSPAN, ISR, PI-ISR tasks) to address these three problems. Scoring was formed to measure target accuracy, slope of accuracy over time, task completion time, and self-regulation failures in errors on the letter-crossing task. This revealed that accuracy provided the best measure for detecting depletion effects. Direct markers of depletion (functional) provided evidence for depletion transfer effects, whereas indirect markers (cognitive) served as theoretical suggestions for the origin of the self-regulation reserve. A downward performance trend line, a functional marker representing depletion effects, was present across the letter-crossing task, however, this depletion effect did not transfer onto the follow-up tasks as initially predicted by the strength model of self-regulation. Individual differences in letter-crossing performance did predict executive functioning on some (OSPAN and ISR), but not all of the follow-up tasks. These findings suggested that components under each of the executive measures (inhibition: Stroop; updating: OSPAN; binding: ISR; and binding with proactive interference: PI-ISR) were related with letter-crossing ability. While each of the executive functioning tasks were correlated to letter-crossing ability, one executive function did not comprehensively account for letter-crossing ability. Instead, some elements of updating ability (OSPAN) and binding ability (ISR) may employ similar working memory processes to that of the letter-crossing task and, when letter-crossing failures occur, show accountability on these tasks. These findings suggested that a global executive ability can account for letter-crossing ability. The thesis then proposes that self-regulation, as measured by the letter-crossing task, could be explained through higher executive cognitions required for active goal-maintenance, executive control, and working memory.

Certification of Thesis

This Thesis is entirely the work of Madeleine McKay Arber except where otherwise acknowledged. The work is original and has not previously been submitted for any other award, except where acknowledged.

Principal Supervisor: Gerry Tehan

Associate Supervisor: Jessica Marrington

Associate Supervisor: Georgina Anne Tolan

Student and supervisors signatures of endorsement are held at the University.

Acknowledgments

This thesis is dedicated to my loved ones. What I have achieved over the past three years would not have been made possible without your ongoing support in keeping my mental, physical, and spiritual wellbeing in check.

Firstly, I would like to acknowledge my university, the University of Southern Queensland, and the Australian Government for all of the financial support through the Research Training Program (RTS) Scholarship and the help I have received during my time.

I would like to thank my colleagues at the university and intuitions abroad for the lunches and laughs shared, and the experience and wisdom you have imparted on me.

To my fellow HDR comrades: cum silentio patimur.

My sincerest apologies to all of my participants over the years for sitting through some incredibly mind-numbing tasks, which thankfully made for some great research! So thank you all for your data.

Thank you to the greatest of friends for rock-climbing, skateboarding, hiking, dancing in the dark, thrifting, feasting, and travelling to the ends of the earth with me. I am lucky to have you all to remind me how pleasantly painful laughter can be.

A special thanks to my family, most notably, my parents. Your beliefs in my abilities have allowed me to achieve some significant goals. Thank you for understanding my hangry stress-bear condition. I love you all.

A big thank you to my supervisors, for all of your support and encouragement over these years (and for letting me run away to Europe for six months with no questions asked).

Lastly, I would like to thank caffeine for always being there.

For the enjoyment of the reader, I have totalled the amount of e-vowel occurrences in this thesis:

- ‘ae’ combinations occurred 1,510 times
- ‘ee’ combinations occurred 819 times
- ‘ie’ combinations occurred 1,191 times
- ‘oe’ combinations occurred 155 times, and finally
- ‘ue’ combinations occurred 541 times.

Table of Contents

List of Figures

List of Tables

Abbreviations

ACC: Anterior Cingulate Cortex

ANOVA: Analysis Of Variance

Blk-# Targets: Block-# Targets

Cf: crystallised intelligence

CFA: Confirmatory Factor Analysis

dPFC: dorsal Parietal Frontal Cortex

EA: Executive Attention

EC: Executive Control

EF: Executive Function

Gf: fluid intelligence

IS: Irrelevant Speech/Sounds

ISR: Immediate Serial Recall task

OSPAN: Operation Span task

PI-ISR: Immediate Serial Recall with Proactive Interference

PM: Primary Memory

RM-ANOVA: Repeated Measures ANOVA

SM: Secondary Memory

SEM: Structural Equation Modeling

WM: Working Memory

WMC: Working Memory Capacity

Chapter 1. Introduction to Self-Regulation and the Depletion Effect

1.1. Overview of Thesis

An individual's ability to control their own actions, behaviours, and thoughts is largely understood to be self-control, or self-regulation among researchers. Ego-depletion is the resulting decline of performance after initial self-regulation processing. This thesis aims to contribute to the field of knowledge by expanding upon the current understanding of the depletion effect and the background processes that contribute to self-regulation. Due to the broad definition of self-regulation, almost any task can be applied to induce self-regulatory depletion provided it is effortful and demanding enough. For instance, a task requiring no physical exertion, such as withholding the act of eating a biscuit, is allegedly just as exhausting to the self-regulation reserve as a hand-grip measure requiring constant physical exertion over a period of time. Self-control theory has been extensively researched in social psychology for the past two decades, however, only recently has the literature started repeatedly publishing null effects when applying the standard self-regulation measures, calling for a replication crisis in the literature. This replication crisis has led the field to collectively challenge the legitimacy of the depletion effect and question what is being measured, if not self-regulation ability.

This thesis is important because depletion in self-regulation has a detrimental impact on an individual. Therefore, understanding this detriment may help to identify why there are fluctuations in willpower, and why some individuals appear to have more self-control than others do. Secondly, the self-regulation research field has contributed to extending the base of knowledge in other fields. Fields such as crime, dietary behaviour, and animal behaviour, have incorporated self-control into their theories to explain unlawful tendencies, eating disorders, and animal intelligence, without taking into consideration the validity and reliability of the supposed self-regulation depletion effects. This research is critical in addressing these practical and conceptual issues within the self-regulation field in order for the self-regulation theory to be applicable to other fields.

Over three studies, the thesis will address some of the issues that have been identified by the literature as necessary in order to make significant advancements to the field. The thesis focuses on one experimental task that has been widely used to supposedly deplete self-regulation resources, namely the letter-crossing task. The key research question this thesis is attempting to understand is what the letter-crossing task is measuring. While this task, among other self-regulation measures, is believed to induce self-regulation depletion, it has largely been accepted without question as to the validity of what the task achieves. Aside from the ability to cross out letters regarding a ruleset, this thesis will identify what underlying abilities the letter-crossing task is measuring. In addition, as the letter-crossing task has rarely been scored, particularly throughout the course of resource depletion, the thesis will examine the best scoring method for this purported depletion-inducing measure. This letter-crossing measure will be compared against validated and reliable executive functioning measures, in order to determine what executive functions are required for successful self-regulation processing under the letter-crossing task. The secondary research question of this thesis is then to identify whether self-regulation processing can be explained by executive functioning processing. That is to say, this thesis will determine whether self-regulation phenomena under the letter-crossing task, including depletion effects, can be best accounted for by complex cognition, thereby reducing the need for theoretical self-regulation models. Importantly, while the thesis is undecided regarding the validity of the depletion effect, the assumptions underpinning the dominant model of the depletion effect, the strength model, and the method employed in inducing the effect is being questioned. The theory behind self-regulation depletion is also examined, as this thesis sets out to explain the phenomenon through higher cognition processes alone.

The thesis is divided into seven chapters. The first chapter outlines the rich history of self-regulation research and introduces the ego-depletion effect and the task in question; the letter-crossing task. The second chapter reveals the complexities within executive functioning abilities and the theoretical overlap between executive functioning and self-regulation processing. The third chapter re-establishes this theoretical link between self-regulation and executive functioning. It also outlines the tasks employed and the purpose of the experiments, which is to understand what the letter-crossing task is measuring by comparing it to validated measures such as executive functioning tasks. The fourth chapter contains the first study, which identifies the best scoring methods for the letter-crossing task and investigates self-regulation performance over the letter-crossing task. The fifth chapter contains the second study, which focuses on identifying the shared ability between executive functioning and self-regulation ability employed under the letter-crossing task. The third study is presented in the sixth chapter. This study determines differences in letter-crossing ability between individuals, and if this ability translates to executive functioning abilities. The final chapter, chapter seven, reviews the findings from all studies. It concludes that the letter-crossing task employs executive functions and that self-regulation, and any depletion effects found under the letter-crossing task, is best explained by executive control. An outlook for the self-regulation theory and suggestions for future research is also included.

1.2. Overview of Chapter

This first chapter introduces the self-regulation field and provides an introduction into self-regulation and the ego-depletion effect. Theories of what powers the self-regulation domain are then considered, namely the process and strength resource models. Additional factors for the self-regulation resource are also briefly considered. The general accepted self-regulation depletion-inducing methodology is then stated, which includes a brief overview of the original letter-crossing task. The replication crisis is then described, including the reported difficulty in replicating the depletion effect under self-regulation tasks, leading to what has been termed as the selective depletion effect. The chapter concludes with the conceptual crisis.

1.3. Self-Regulation

The broad domain of self-regulation refers to the unconscious and conscious controlling of mental states, behaviours, thoughts, and responses (Baumeister, 2002). Behavioural self-regulation, or self-control, deliberately overrides immediate desirable outcomes in favour of delayed gratification for long-term benefit to the individual (Hofmann, Schmeichel, & Baddeley, 2012; Muraven & Baumeister, 2000). Simply put, self-control is the individual’s ability to override impulses (Baumeister & Heatherton, 1996). An everyday example is restricting diet and poor lifestyle choices. Applicable to real-world scenarios, self-regulation has been studied across social, personality, and cognitive psychology fields, with little communication between. This has ultimately resulted in no consensus on what the self-regulatory process measures, what the process is related to, or what resources power these processes.

Under Baumeister’s (2002) social psychology based account, self-regulation is the principal function of the self’s processing. Self-regulation can be described in terms of state or trait regulation. State is the current act of regulation, whereas trait regulation is the ongoing tendency to exert control over the self (Baumeister &

Alquist, 2009). Importantly, self-regulation is an active process, requiring mental effort and energy for goal-directed behaviour (Hofmann et al., 2012). Capacity-based theories determine self-control to be powered by a limited ‘global reservoir’ of energy, when spent, this reservoir limits the capacity to control the self (Hagger, Wood, Stiff, & Chatzisarantis, 2010; Muraven, Tice, & Baumeister, 1998; Nes, Roach, & Segerstrom, 2009). This common resource is thought to influence habit breaking, decision making, rational thinking, response inhibition, tolerating pain, mental and physical endurance (Inzlicht & Schmeichel, 2012). Complex cognition, often regarded as higher cognition, involves executive functions associated with planning, goal-attainment processing, and problem solving. As self-regulation is the capacity to control the self in regards to a goal, the energy reserve of self-regulation is then supposed to power higher cognition (Baumeister, 2002; Baumeister, Tice, & Vohs, 2018). In doing so, the theory of self-regulation accounts for the processing behind all cognition and behaviour. The more effort required by a task or process results in depletion; a decreased capacity to regulate ones’ processing and performance on such tasks (Hall & Fong, 2007; Muraven et al., 1998).

1.3.1. Depletion.

Baumeister and his colleagues termed the notion “ego-depletion” to refer to the state following exhaustion of self-regulation, which was made popular amongst social psychologists. Ego-depletion is the temporary depleted state of resources that empower the self-regulation reservoir (Baumeister & Vohs, 2007). During this short-term mental fatigue, self-regulation is considered to operate at less than full capacity (Baumeister, 2002; Inzlicht & Berkman, 2015; Muraven & Baumeister, 2000). This exhausted state results in ongoing failures of self-regulation (Muraven et al., 1998). Depletion translates across self-control domains including regulating emotions and behaviours, physical endurance, suppressing thoughts, and cognitive persistence tasks (Muraven et al., 1998). The consensus is that self-regulation is then powered by a common reserve of mental energy, which is recruited by variety of tasks (Muraven & Baumeister, 2000). This reserve appears to be somewhat small, as even short and slightly demanding tasks appear to deplete it (Muraven & Baumeister, 2000). Tasks requiring controlled thinking or withholding actions manages to negatively affect self-regulation on seemingly unrelated future tasks. The original experiment conducted by Baumeister, Bratslavsky, Muraven, and Tice (1998) required participants to resist eating cookies and chocolates, which resulted in a decrease in persistence on unsolvable puzzles that followed. This ongoing demand of self-regulation leads to the fatigue of resources and reduced capacity of self-regulation (Nes, Roach, & Segerstrom, 2009).

Within the laboratory, evidence for the self-regulation failure relies on a decrease in performance on a secondary self-regulating task following an initial task requiring self-regulation, this method is known as the sequential-task paradigm (Inzlicht & Schmeichel, 2012); outlined in section 1.6.1. Depletion effects are achieved by comparing performance differences between control (i.e., non-depleted) and experimental (i.e., depleted) groups, as individuals who have previously engaged in self-regulation and have been depleted as a result, will quit effortful processing earlier than non-depleted individuals on follow-up tasks (Tice, Baumeister, Shmueli, & Muraven, 2007). Successful depletion of self-regulatory resources has been induced across a number of simple and brief measures. The original study by Muraven, et al. (1998) requested participants regulate their emotions while watching a sad or distressing video. This was followed by a persistent hang-grip measure (Muraven et al., 1998). Under the same set of studies, depletion was also encouraged through the suppression of thoughts about a white bear, followed by measuring persistence on unsolvable anagrams (Muraven et al., 1998). Baumeister et al. (1998) have also induced depletion through the suppression of desires to eat delicious food such as cookies and chocolates after skipping a meal and instead encourage efforts to eat only radishes, which was followed by measuring persistence on tracing a geometric figure.

As self-regulation is a complex and multi-faceted processing system, failure can occur on many different levels, and because of this, it is difficult to interpret the direct cause of failure (Baumeister & Heatherton, 1996). Failure may occur due to under-regulation or misregulation. Under-regulation occurs when the strength or willpower is the primary cause of weakness in the system, eventually causing failure, whereas misregulation is the misguided efforts or maligned goals leading to self-regulatory failure (Baumeister & Heatherton, 1996). Baumeister and Heatherton (1996) propose self-regulation failure to be a combination of acting on impulse and deliberate decision-making. These conscious compromises, such as trade-offs on resource expenditure, are likely to use some resources in order to make this decision (Baumeister & Alquist, 2009). The conservation hypothesis of self-regulation suggests the ‘mental muscle’ adapts to the demand level by strategically allowing depletion to occur when continuing mental load is required (Hagger et al., 2010). This allows the occasional conservation of expending valuable resources. Occasionally decisions may result in individuals underestimating their self-regulation ability even on simple tasks. Such as a speed-accuracy trade-off, where individuals consciously decide to speed up processing on a task without recognising the negative impact this will have on their task accuracy (Baumeister & Heatherton, 1996). Additionally, self-regulatory failure may occur from insufficient capacity to monitor task goals, a lack of motivation towards the task or set-goal, or a lack of capacity to employ self-regulatory processes (Hofmann et al., 2012).

Inzlicht and Schmeichel (2012) argue the depletion effect is primarily caused from a shift in motivation and attention towards something else, which results in failed self-control on the secondary task under the process model. Motivational and attentional processes that oversee conflict or goal discrepancies are redirected towards something more rewarding (Inzlicht & Schmeichel, 2012). For example, an individual may direct their attention and motivation away from continuing to eat the radishes and attend towards the more pleasantly tasting cookies instead, thereby resulting in a poorer functioning state on the secondary task. Individuals will abandon original goals to escape certain aversive conditions. Nordgren and Chou (2012) found that individuals’ self-regulation of behaviour was more likely to fail when participants were in a high visceral state and must actively decide on an action. Visceral states were encouraged in participants by restricting cigarettes for smokers and forcing hungry dieters to make quick food decisions, as their needs (nicotine or food cravings) were not immediately satisfied. Under Nordgren and Chou’s (2012) research, self-control resources were thought to have encouraged depletion by focusing on immediately satisfying urges in order to escape the visceral states inflicted (Nordgren & Chou, 2012). However, individuals with low to no cognitive load were found to delay gratification on smoking for longer than higher cognitive load individuals, who were required to remember a number-string (Nordgren & Chou, 2012). Regardless of cognitive load, behavioural self-regulation was consistently hindered under high visceral state conditions (Nordgren & Chou, 2012). Increasing stress levels or negative affective states is therefore considered to be detrimental to self-regulatory resources. Emotions may wreak havoc on self-regulatory resources, particularly emotional distress, causing individuals to lose sight of their goals and risk self-regulation failure (Baumeister & Heatherton, 1996).

While self-regulatory depletion is a temporary failure and not a permanent deficit, there appears to be some individuals with greater levels of self-regulation who are less susceptible to self-regulatory failure. Examples of real-world self-regulatory failures have been linked to low academic performance, procrastination, illegal activities, cigarette smoking, binge eating, alcohol and drug abuse, spousal abuse and violent outbreaks, gambling, and promiscuity (Baumeister & Heatherton, 1996). Dual low self-control in domestic partnerships has been found to influence alcohol consumption and encourage aggression between partners (Quigley, et al., 2018). Whereas greater self-regulation capacity has been linked to higher grade-point averages in academia, fewer cases of eating disorders, less alcohol abuse, less psychopathology and mental health issues, stable interpersonal relationships, better anger management, fewer emotional issues, and fewer delinquent behaviours (Engels, Finkenauer, & den Exter, 2000; Tangney, Baumeister, & Boone, 2004). Even though it has been broadly accepted that there are individual differences in self-control, the resource behind this system has not been successfully identified yet and whether the theoretical capacity behind self-control is a limited resource or a limited processing system, is debatable. Namely, whether depletion is due to the exhaustion of all self-control resources within an individual, or whether depletion is due to the exhaustion of resources willing to be spent by the individual at the current time.

1.4. Self-Regulation Theoretical Models

A number of models have attempted to explain how self-control results in failures. Generally, self-regulation models have followed the blueprint of a limited capacity system to explain depletion effects, wherein depletion results from the emptying or the consumption of some self-regulatory resources. In their literature review of self-regulation, Baumeister and Heatherton (1996) suggested that the consensus across the fields was that a limited resource powered the ability to regulate the self. Limited capacity theories of self-regulation state that self-regulatory failures occur due to the exhaustion, or rather the depletion, of all potential self-control resources. In other words, the self-regulation capacity, at any one time, has absolute processing limits. Importantly, due to this resource being shared across domains, any effortful processing of the self is thought to employ this capacity’s energy. Exertion of these resources then leads to the temporary exhaustion of resources (Muraven et al., 1998). Although, this resource is understood to be renewable (Baumeister & Heatherton, 1996). Earlier work compared this resource to working memory or attentional control capacity, wherein limited processing of information or control over the self can occur at once, but with practice, the amount that can be processed can increase over time (Baumeister & Heatherton, 1996). This limited capacity approach has been applied to explain depletion effects across a variety of domains in self-control, including the consumption of glucose stores within the human body (Gailliot et al., 2007). Although, a state of depletion was not typically induced in research until the early 1990s (Karoly, 1993), when Baumeister's theoretical strength model of self-regulation was applied in experimental methods. Most theories have since moved away from a limited resource, towards a consumption-based theory.

Consumption based theories of self-regulation suggest that depletion occurs when ongoing self-control processing leads to a limitation of what a reserve can process at any one time. Baumeister et al. (1998) refer to the self-regulation resource as a finite resource, where depletion effects are considered similar to muscle fatigue. During this depleted state, ongoing self-control might then require more effort (Muraven & Slessareva, 2003). Alternatively, individuals conserve their resources during this state of fatigue and are unlikely to want to expend their valuable resources (Muraven & Slessareva, 2003). Thereby, individuals act as if their self-control is limited in resources, when in theory, their self-control processing is limited (Muraven & Slessareva, 2003).

While both limited and consumption-based theories appear to define the self-regulation reserve as limited in processing capacity, there is a distinction between them in the effect of depletion. The key difference between these two theories is that the limited capacity approach suggests depletion arises from the exertion of all self-control resources, thereby leaving nothing left, unlike consumption-based theories that suggest continued processing during a state of depletion can occur if necessary. Limited resources would therefore be accessible, consequently resulting in a lowered level of self-control. Although, there is no consensus on what powers this consumption-based self-regulation reserve. For the purpose of understanding the complexities behind self-regulatory depletion, two major consumption-based models are briefly outlined here: the process model and the strength resource model.

1.4.1. Process model.

Inzlicht, Legault, and Teper (2014) view self-regulation as a motivational and cognitive process system, rather than a finite resource system, where motivations strongly shape the outcome of resources. After individuals have committed to controlling the self for a time, a "cognitive leisure state" is perceived as a higher priority than to continue with self-control, which is seen as an aversive experience (Inzlicht et al., 2014). In other words, individuals rationalise that they owe themselves a break from self-control. Depletion is then just laziness, an unmotivated state to apply self-control resources in order to achieve the set goals (Inzlicht et al., 2014). Similar to Carver and Scheier's work (1981), self-control under the process model occurs during conflicts between the goal state and the current state. For example, the goal state of studying would be in conflict if the current state were watching videos. Self-control processing systems would need to be activated in order to achieve the goal state. The self-control process system put forth by Inzlicht and colleagues entails a self-monitoring feedback system that requires a goal setting system, a monitoring system that assesses behaviour with internal goals, and then an implementing system that endorses behaviours to match these internal goals (Inzlicht et al., 2014). Under this system, the implementing system sends feedback to the monitoring system (Inzlicht et al., 2014). The monitoring system follows orders from to a goal setting system, that is, the individual's goals. The feedback between systems continues until a respectable level of match between internal goals and outer behaviour is achieved between systems (Inzlicht et al., 2014). In the previous example, the implementing system would provide the feedback that study was not occurring, the monitoring system receives feedback from the goal-setting system and identifies this conflict between a desire (i.e., watching TV) and the set goal (i.e., studying). The monitoring system then relays this information to the implementation system, where behavioural change should then occur to meet goal standards. During self-regulatory depletion, motivations towards ongoing goal attainment are eventually shifted towards the more ideal cognitive leisure state (Inzlicht & Schmeichel, 2012). Simultaneously, attentional processes fail to attend to the conflicts arising between goal-setting and monitoring systems (Inzlicht & Schmeichel, 2012). The result is goal neglect, which in the example would be watching TV over the set goal of studying.

Inzlicht et al. (2014) have reported a neural basis for this process model system linked to the anterior cingulate cortex (ACC). Within the ACC, an evoked brain potential termed 'error-related negativity' occurs when neurons respond to errors or conflict, and the resulting affective response to such conflict (Inzlicht et al., 2014). Conflicts between goals and behaviours output by the individual, allows the individual to refocus on the goal and adjust behaviours accordingly in order to effectively achieve the desired state (Inzlicht et al., 2014). The goals set by the individual must be achievable yet challenging, and refrain from abstract ideas or timeless sets. For example, an individual dieting may set the abstract goal "to eat less" rather than "to limit their intake of sugars and fats five days a week". Self-determination theory puts forth the idea that goals that align with an individual's morals are more likely to be achieved, even if the individual is fatigued, as internally driven goals are more effective in controlling behaviour (Inzlicht et al., 2014). Thereby suggesting the influential strength of internal motivation over self-control and in turn, self-regulation depletion. Although, Wan and Sternthal (2008) found that the depletion effect was eliminated with the addition of immediate task performance feedback.

Under Wan and Sternthal's (2008) studies, participants were told they would complete a series of tasks for one hour, and were to remove all devices that told them the time. The first study required participants to complete a computerised letter-crossing task of meaningless text. Control participants were told to mark any instance of '*e*', whereas the depletion condition participants were required to mark an '*e*' when it did not neighbour a vowel or a letter next to a vowel. Participants then attempted an unsolvable puzzle that required the reorder of numbers (1-15) within a 4 x 4 matrix into numerical order, with one space at the bottom of the matrix to move numbers adjacent for reordering. A clock provided immediate feedback as to how long participants had spent on attempting to solve the puzzle under the feedback condition. It was argued that the clock provided feedback as to how much self-regulatory resources were spent (Wan & Sternthal, 2008). The same self-control letter-crossing task and unsolvable persistence puzzle task were employed in the second study. Feedback was manipulated through two clock conditions and a control condition with no clocks. One clock was accurate in the amount of time that had passed; the other was faster by four seconds. Both clocks updated every 15 seconds. The faster clock then provided an incorrect feedback condition on how much self-regulatory resources were spent by the individual (Wan & Sternthal, 2008). By using a moderator approach, Wan and Sternthal (2008) found explicit task feedback and individual differences in self-monitoring counters depletion effects. That is, participants persisted on the unsolvable puzzles when presented with feedback cues (i.e., the clock). Unless the feedback was inaccurate (i.e., the faster clock), in which depletion effects were observed (Wan & Sternthal, 2008). This suggests depletion is a monitoring issue rather than resource depletion issue. Participants are thought to monitor their self-regulatory resources spent based upon a standard (i.e., time that has passed). Quitting can then occur once participants believe this standard, or allocated period of time, has been reached (Wan & Sternthal, 2008). If this standard is false, such as the incorrect feedback clock condition, then participants will quit earlier due to their false belief on self-regulatory resources consumed (Wan & Sternthal, 2008).

From these results, Inzlicht and Schmeichel (2012) argue that performance task feedback could be used as an external conflict-monitoring system for an individual to attend to. Conversely, this was not the case under Wallace and Baumeister's (2002) feedback results, where both negative and positive feedback on participant task performance was found to not moderate any effect on Stroop task performance. Participants were allocated into one of four groups: Depletion-Positive-Feedback, Depletion-Negative Feedback, Depletion-No-Feedback, and a No-Depletion-No-Feedback condition. To induce depletion, participants were given a variant of the Stroop task (Stroop, 1935), where participants were required to name the ink the word was presented in (i.e., the colour-naming task). For the non-depletion group, all of the trials were congruent. That is, all the inks and words were matched with one another (i.e., the word 'green' was presented in a green ink). All trials of items were 100% incongruent for the depletion condition groups, thereby an overriding response was required in order to name the ink colour, rather than read the word. For example, 'green' would be presented in a blue ink. If the ink colour was red, however, participants were to read the word rather than name the ink colour. The feedback participants received was dependent on their experimental condition. Participants in the Positive-Feedback condition were told they had scored higher in accuracy than most other participants and were faster to complete the task than 95% of participants (Wallace & Baumeister, 2002). In comparison, participants under the Negative-Feedback condition received feedback that they had made more errors than most other participants had, and were within the slowest 25% of the whole participant group to complete the task (Wallace & Baumeister, 2002). Participants under the No-Feedback condition received no input in regards to their colour-naming task performance. The follow-up task was an unsolvable finger-tracing task to measure depletion in persistence, whereby participants were required to trace a figure on paper with their fingers without lifting their finger from the page or retracing lines. As participants did not know the task was unsolvable, they were told to work on the task until they completed it successfully or gave up. While depletion effects continued to occur regardless of feedback, by way of a decline in persistence on the follow-up task, the results found the Negative-Feedback condition, rather than the Positive-Feedback condition, persisted on the unsolvable task for longer (Wallace & Baumeister, 2002). Likewise, reinforcing the need that participants must perform well on the experimental tasks, which is often encouraged through cash incentives and competition, has reportedly resulted in poorer performance and has been theoretically linked to self-presentation concerns (Baumeister, 1982; Baumeister, 1984; Baumeister, Hamilton, and Tice, 1985; Baumeister, & Steinhilber, 1984).

The process model proposed by Inzlicht and colleagues (2014) strongly relies on the personal willpower and conscious processing of internal goals and external actions by the individual, and does not overtly account for the findings by Wallace and Baumeister (2002). An alternative is the strength model that Baumeister and colleagues have put forth and developed over the years, which similarly proposes a finite processing capacity. Whilst motivations remain a component within the processing ability, this reserve is not primarily powered by motivations.

1.4.2. Strength resource model and conservation hypothesis.

Unlike limited capacity models that assume depletion results from the exhaustion of all resources and results in an empty tank, the strength model proposes the reserve is more similar to a muscle (Muraven & Baumeister, 2000). Self-regulation is thought to temporarily wear down the reserve over time, which ultimately results in the observed decline in performance (Muraven & Baumeister, 2000). Notably, self-regulation does not consume all the resources, rather like the conservation hypothesis suggests, it consumes some and retains some energy for the future processing of tasks deemed to be important (Baumeister, 2002; Baumeister & Vohs, 2007). After an exertion of limited processing power, the reserve can be replenished through rest or positive emotions such as humour (Baumeister, 2002; Muraven & Baumeister, 2000; Tice et al., 2007; Tyler & Burns, 2008).

There are several key assumptions to this strength resource model, outlined by Muraven and Baumeister (2000). The self’s executive function, which powers and controls output from the reserve, is critical to everyday functioning (Murvan & Baumeister, 2000). This reserve powers all tasks, which require control over the self, but this reserve is limited and can only control a finite amount of tasks over a period (Muraven & Baumeister, 2000). If this strength resource is weakened, through natural expenditure of control and processing, failure of the conflict-monitoring system occurs and results in depletion (Muraven & Baumeister, 2000), however, this failure is not permanent (Muraven & Baumeister, 2000). Additionally, the strength resource resolves issues arising from competition with the task goal (Muraven & Baumeister, 2000). This competition resolution process would then naturally consume the limited capacity, however, ongoing goal-maintenance would also require some power to consistently meet the task's needs (Dang, 2017). When fatigued, self-regulation failure is likely to occur even on simple self-regulating tasks (Baumeister & Heatherton, 1996). For example, most naturally occurring failures of self-control, such as diets, tend to occur later in the day or evening when self-control is already exhausted (Baumeister & Heatherton, 1996). Baumeister and Heatherton (1996) suggest that after a day of controlling impulses, redirecting desires, and maintaining a good level of self-regulation, the self-control 'muscle' is weakened from ongoing exertion. Goals therefore shift towards immediate gratification and rewards, that is, the well-deserved cookie or the cigarette.

Strength of this self-regulation reserve is dependent upon the individual and can be observed through individual performance on self-control tasks, however, it is assumed strength reserve can be increased through training for long-term improvement (Muraven & Baumeister, 2000; Muraven, Baumeister, & Tice, 1999). Individuals with higher self-regulatory strength are predicted to have high self-control globally, compared to individuals with lower self-control whom are likely to be susceptible to failure across multiple domains (Baumeister & Heatherton, 1996). If self-regulation is dedicated to one sphere, other spheres are more likely to fail (Baumeister & Heatherton, 1996). Furthermore, the strength reserve is susceptible to stress even after stress has occurred. Affective states such as moods and emotions can also affect the self-regulation reserve (Muraven & Baumeister, 2000). Under stressful conditions, it is thereby difficult to determine whether stress weakened the reserve or enhanced negative impulses that changed motivations irrelevant to the self-regulation goal, thereby resulting in a self-regulatory failure (Muraven & Baumeister, 2000). A reduced ability in self-regulation may then be compensated for other factors such as motivation, cognition, and emotions (Baumeister & Vohs, 2007).

Baumeister and Vohs (2007) outlined four necessities for successful self-regulation processing: standards, monitoring, strength, and motivation. Firstly, standards or goals, are often imposed externally on the individual by society or a particular task, and give the framework for the participant to work in compliance with the goal (Baumeister & Vohs, 2007). These standards need to be clear, consistent, and without conflicting goals otherwise self-regulation will fail (Baumeister & Heatherton, 1996). Secondly, monitoring via a conflict resolution system adjusts behaviour not in accordance with the standard, and responds appropriately. Originally based on the feedback-loop model of Carver and Scheier (1981, 1998), the strength resource model of self-regulation monitors inner states and the operate phase changes them in line with external and internal standards imposed (Baumeister, 2002; Baumeister & Heatherton, 1996; Muraven & Baumeister, 2000). If monitoring stops, even briefly, failure will occur in what is commonly described as “a moment of weakness”. This loss of control is particularly evident in diet failures, where one cookie results in the failure of the entire diet due to the individual viewing one cookie as the decisive failure of the goal (Baumeister & Heatherton, 1996). Thirdly, strength or willpower, as Baumeister’s strength resource model requires a limited resource for the individual to act upon (Baumeister, 2002). The absolute limit of this strength reserve has not been investigated; it may power multiple tasks at once or once a limit has been reached, no further reserve is expended (Muraven & Baumeister, 2000). Fourthly, an individual’s motivation towards the goal and actively participating in steps to reach this goal, plays an important part in administering effort required to achieve the standard. Without motivation towards achieving the set-goal, the individual is less likely to exert valuable self-control resources or attempt to work towards this standard.

This thesis explores several of the assumptions that underpin the strength model approach. Firstly, that resource depletion occurs over time with repeated expenditure of these resources. Secondly, the assumption that a common set of resources underpins all sources of self-regulation. Thirdly, that resource reduction in one task has carryover effects to a second task. The thesis also attempts to elucidate what these resources are.

1.5. Additional Factors Impacting on Self-Regulation to Consider

When self-regulatory resources are low, secondary components like motivation, emotion, and cognition, are encouraged to compensate for lower levels of self-regulation processing (Baumeister & Vohs, 2007). Across a series of studies, Tice et al. (2007) showed positive emotions could be beneficial to self-regulation and protect from self-regulatory exhaustion. After completing a self-control thought suppression task, where participants were told to list all their thoughts for five minutes but not to think about a white bear, participants in the first study were either given a small gift (i.e., small bag of candy) or a receipt for their participation. This was followed by a persistence drinking task where participants were required to drink an unpleasant drink (i.e., orange Kool-Aid and vinegar mixture). Importantly, the tasks were conducted in separate environments to ensure task-separation and limit participant's knowledge on the relationship between the tasks (Tice et al., 2007). In a similar design the second study involved a self-regulation letter-crossing task requiring participants to cross off any '*e'* in a short text, except when it appeared next to a vowel or neighboured a letter next to a vowel. Emotion was then manipulated through a positive humorous or a neutral dolphin communication short video (Tice et al., 2007). The third study required a hand-grip task, where participants were required to grip the instrument consistently for as long as they could. The hand-grip task was measured at baseline and followed the self-control thought suppression task. Emotions were then manipulated through short videos. Positive or negative mood was induced through a humorous or sad short video respectively, whereas the neutral conditions were told to take a five-minute break (Tice et al., 2007). The fourth study required participants to not haven eaten three hours prior to the study. Self-control was induced through a radish-eating task, where participants were required to eat as many radishes as they could, while suppressing the urge to eat cookies or candy presented in front of them. Furthermore, participants were told they were not to eat any radishes, cookies, or candy for the 24 hours following the experiment in order to encourage impulses to eat these foods, which in theory, would boost any depletion effects found under the experiment (Tice et al., 2007). Mood was induced through humorous or neutral dolphin communication videos (once again in a separate room to ensure task-separation). Across the four studies, the findings consistently found when positive emotions were induced in participants, from watching funny videos or receiving a gift, they were less likely to show signs of self-regulatory depletion compared to groups under neutral or negative-emotion conditions (Tice et al., 2007). However, it was not clear whether inducing positive moods impacted on the self-regulation reserve or increased motivation on the tasks (Tice et al., 2007). Nevertheless, there have been suggestions that depletion can be explained partly through lack in motivation.

1.5.1. Motivation component.

Self-regulation can be described as an attempt to control motives and curb desires behind behaviours (Baumeister & Vohs, 2007). If there is motivational conflict, that is a desire to achieve another goal irrelevant towards the previous goal, self-regulation is then engaged to realign an individual's motives towards the original goal (Baumeister & Vohs, 2007). Motivation is then widely accepted as an influential element of self-regulation (Carver & Scheier, 1981, 1998; Inzlicht & Schmeichel, 2012; Muraven & Baumeister, 2000). It is thought to work with the strength reserve for successful self-regulation processing (Muraven & Baumeister, 2000). Motivational influence may be necessary during tasks, as commitment to the task is critical particularly under self-control demanding conditions (Baumeister & Vohs, 2007). The global resource may also use self-regulation to drive the motivation behind certain behaviours (Baumeister & Vohs, 2007). However, conflict between motivations encourages self-regulatory processing and possible failure (Baumeister & Vohs, 2007).

In one study conducted by Muraven and Slessareva (2003), increased motivation to complete the task resulted in seemingly depleted participants performing at the level of highly motivated non-depleted participants. A cognitive task was allocated to induce depletion, this was either a thought persistent task (i.e., write down all your thoughts but avoid thinking about a white bear) or a memory task (i.e., a simple recall task of word items), followed by two unsolvable puzzles to test depletion under persistence (Muraven & Slessareva, 2003). While simple memory tasks arguably require inhibition in the suppression of previous items to be remembered, it was argued that this task would employ less inhibition than thought persistence tasks and would therefore utilise fewer self-control resources (Muraven & Slessareva, 2003). Importantly, a subgroup of participants were manipulated into thinking their cognitive performance on the tasks would help advance memory research, which could benefit Alzheimer's research and treatments, thereby motivating them to perform well on the tasks (Muraven & Slessareva, 2003). A manipulation check was included to assess task difficulty and effortful self-control. Participants reportedly found the thought suppression task more difficult, than those who completed the simple memory task, perhaps as a result of this, those participants who completed the thought suppression task showed depletion effects by quitting the unsolvable puzzles sooner (Muraven & Slessareva, 2003). When given a motivation incentive, participants were found to persist even in the face of depletion through the belief the task would benefit them or others (Muraven & Slessareva, 2003). Therefore, motivation was considered to play a role in self-regulation through compensation (Muraven & Slessareva, 2003). Depletion effects are most likely to occur when participants are low in motivation and self-regulatory resources (Muraven & Slessareva, 2003). Although, the authors argue that depletion cannot be eliminated wholly by internal motivation engagement (Muraven & Slessareva, 2003). Baumeister and Vohs (2007) agree with the conclusion that motivation is compensational in nature. Motivation may compensate for a depleted self-regulatory state by inspiring the individual to continue with the task and spend the limited resources left (Baumeister & Vohs, 2007). Typically, however, motivation is not measured in experiments, rather it is inferred from participants wanting to do well on the task to enticing participants with incentives (Lee, Chatzisarantis, & Hagger, 2016). Regardless, it should be considered a potential factor of self-regulation.

1.5.2. Attention component.

Associated with higher cognitive functioning, controlled or sustained attention has been accepted as another component of self-regulation processing. Self-regulation allows the individual to consciously process environmental information and exclude undesired or irrelevant information (Schmeichel, 2007). Attentional control is deemed to be critical in the effective management of impulses (Baumeister & Heatherton, 1996). As the earlier the intervention to self-regulatory depletion, the more effective the self-regulation process is (Baumeister & Heatherton, 1996). Applying focused attention to goals or standards has been termed transcendence and this act of bringing attention towards these standards then helps individuals to reach their goals (Baumeister & Heatherton, 1996). However, Muraven and Baumeister (2000) have argued that attention cannot solely explain the self-regulation process under the strength resource model, as tasks that do not require much attention, such as persisting in drinking bad-tasting beverages, have a detrimental effect on self-regulation resources. Whereas tasks that do require sustained attention, such as solving mathematical operations, reportedly do not drain self-regulation resources (Muraven & Baumeister, 2000).

In addition to the strength resource model, attentional control has also been applied to the process model of self-regulation. Task attention is then sustained through goal monitoring systems, which find conflicts between intended responses and instead present a task-relevant response to be acted upon (Inzlicht & Schmeichel, 2012). Failure to attend to the conflict responses results in a goal neglect via the conflict-monitoring system (Inzlicht & Schmeichel, 2012). The process model suggests depletion shifts motivation and attention away from task goal orientation and towards short-term reward and gratification (Inzlicht & Schmeichel, 2012). One system, the motivation or attention system, is assumed to affect the other, thereby if attentional control is interrupted then motivation will be lost. Thereby if sustained focus on the task goal of eating radishes is lost, motivation towards eating the radishes will be dropped for a more desirable one, such as eating the cookies. If factors such as emotion and motivation can explain some effects and are described as being secondary to the reserve in the literature, a depleted reserve may then be physiologically explained by a neural energy reserve.

1.5.3. Glucose theory.

A number of studies have asked whether self-regulation can be explained physiologically through low glucose levels. Glucose, or glycogen, is stored energy in the brain required for executive functioning (Gailliot et al., 2007). It has been put forward as a resource for powering self-regulation. Thereby, low glucose levels then would result in lowered executive functioning and a depleted self-regulatory state. Miller (2012) followed this glucose theory of self-regulation, where glucose was been linked to cognitive energy, namely executive functions, which in turn were thought to execute self-regulation processing. Naturally, glucose levels fluctuate throughout the day due to food intake to which cognitive processing capacity remains unaffected by, provided the fluctuations occur within a normal range (Gailliot et al., 2007).

Gailliot and colleagues (2007) conducted a series of studies assessing this assumption between self-regulatory performance and blood glucose levels. Following a baseline blood test, a range of effortful tasks were employed to induce self-regulation depletion and encourage glucose consumption. These self-control measures included typical self-control measures such as an attention video that require the suppression of attending towards subtitles during a silent video (studies 1, 3, 4, and 7); the original Stroop task (study 5 [Stroop, 1935]); or an emotional regulation task where emotions were suppressed during an emotional video (study 6). Alternatively, self-regulation depletion was encouraged through avoiding negative racial responses in an interracial discussion for five minutes (study 2); writing down the emotions surrounding the thoughts of their own death and decay of their body (study 8); or an multiple-choice exam for a university course (study 9). Gailliot and colleagues (2007) consistently found that following these tasks, follow-up blood-tests revealed that blood glucose levels dropped and lower blood glucose levels predicted poorer performance on a follow-up persistence measure, such as the figure-tracing task (studies 4-6) or a word-fragment solving task (study 8). Alternatively, the follow-up task was an executive function-task, such as the Stroop task (studies 3, 8, and 7) or their willingness in a hypothetical situation to help a stranger (study 9). Participants' lowered performance on the post-test could be counteracted by ingesting of a glucose drink (studies 7, 8, and 9; Gailliot et al., 2007). When additional emphasis was put on time, speed on the Stroop task was affected by blood glucose levels, where lower blood glucose resulted in slower Stroop performance (study 3; Gailliot et al., 2007). Some individuals were more susceptible than others under the self-control tasks, from which Gailliot and colleagues (2007) argued this finding could be explained by individual differences in personality and values, or physiological capabilities and responses. Gailliot and colleagues (2007) concluded that reduced blood glucose levels partly explain self-regulatory depletion, and that glucose could be the fuel for this higher level processing. Although, a number of issues have arisen from this set of experiments, such as the sample sizes.

Somewhat concerning is the majority of the Gailliot et al. (2007) experiments reported small sample sizes (study 2: N = 38, study 3: N =16, study 4: N = 12, study 5: N = 23, study 6: N = 17, study 9: N = 18), with the exception of studies 1 (N = 110), 7 (N = 62), and 8 (N = 73). Moreover, glucose blood levels at the beginning of the experiments did not predict task performance, which suggests that some kind of processing (whether self-regulatory or cognitive) consumed blood glucose, however, the uptake of glucose within the brain was not measured, rather inferred. Regardless, relatively small acts of self-control were thought to lower blood glucose, and from this, the brain was assumed to consume resources. Although the results favour this glucose-consumption self-regulatory reserve, cerebral glucose metabolism within the brain has not been measured, rather the measurement of blood glucose levels, which is therefore hard to support this theory. Increases in heart rate and a reduction in blood level glucose during a Serial Sevens task, a mental function task requiring participants to count down from 100 in increments of seven, imply that the brain is not the only organ responsible for glucose drops (Scholey, Harper, & Kennedy, 2001). Furthermore, Kurzban (2010) reanalysed the Gailliot et al. (2007) dataset and concluded the data cannot support the claim that self-control tasks decrease glucose levels. This conclusion came from the finding that the observed reductions in glucose were due to an abnormally high population average that upon follow-up testing regressed to the mean, and as a result provided false support for the glucose-self-regulation-consumption hypothesis (Kurzban, 2010). Kurzban (2010) suggested that rather than an energy resource, glucose should be looked at as an influencer to such computational mechanisms.

Subsequent research has failed to successfully replicate the drop in glucose after cognitive loading which has brought into question whether self-regulation can simply be explained by glucose levels (Inzlicht & Berkman, 2015). In addition to this, cerebral glucose metabolism typically remains stable over time and throughout cognitive effort (Inzlicht & Berkman, 2015). Marcora, Statiano, and Manning (2009) found blood glucose levels did not decrease under a continuous performance task, adding further to questioning the link between self-regulation and glucose. Madsen et al. (1995) reported no difference before, during, and after a Wisconsin Cart Sorting task in blood level glucose within the artery entering and exiting the brain. The Wisconsin Cart Sorting task measures set-shifting ability through matching a series of cards. Participants are required to match the cards, based on the colours or shapes presented on the cards, to a card deck already set out, however, they are not told how to match them. The experimenter provides feedback as to whether their match was correct or incorrect. As the correct card-matching method changes frequently, participants must then have cognitive flexibility in changing their set-matching methods. While the task itself is difficult, the lack of variance in glucose resource consumption suggests that glucose is not critical for complex cognitive functioning. Brain glucose absorption during difficult tasks has reportedly been estimated to only increase by approximately 1% (Raichle & Mintun, 2006). At the same time, earlier research has linked increased glucose levels with greater speed of cognitive functioning on some tasks. Memory, information processing, and reaction times were seen to improve with higher levels of glucose recorded (Benton, Owens, & Parker, 1994; Donohoe & Benton, 1999; Owens & Benton, 2008). However, this is thought to be dependent on individual differences in glucose uptake efficiency and not due to these tasks directly forcing the uptake of glucose resources (Donohoe & Benton, 1999).

Throughout this debate, these studies have been adapted to real-word settings, where a study conducted by Danziger, Levav, and Avnaim-Pesso (2011) measured favourable rulings in judicial courts of law. Danziger et al. (2011) found that probabilities of favourable rulings decreased within the session from ≈.65 - ≈0. The morning sessions were approximately 46.8 minutes in duration with an average of 7.8 cases reviewed prior to a morning tea break of 38.48 minutes (Danziger, et al., 2011). The afternoon sessions were slightly longer with 11.4 cases reviewed in a session of approximately 68.4 minutes before a lunch break of 57.37 minutes (Danziger, et al., 2011). After each break, favourable rulings resumed at ≈.65 (Danziger, et al., 2011). While each case was different there were similar characteristics shared between them, which could be considered as a limitation for the findings. The authors also noted that they did not measure mental resources, and therefore cannot determine whether it was simply a rest from the task, or eating on their break that supposedly restored judges' mental resources or favourable moods. Nonetheless, Danziger et al., (2011) concluded that mental fatigue can then be replenished by breaking the task and increasing glucose levels in the body.

Taking all of the findings into consideration, Beedie and Lane (2012) suggest the concluding remarks should consider self-regulation depletion not to be due to limited glucose resources, but rather a lack of efficiently allocating resources for consumption. Glucose may partially explain the self-regulatory depletion effect through facilitation of self-control. Although, as previously stated, factors such as motivation and attention are thought to also influence individual differences in self-regulatory decline. These factors should be considered through manipulation checks when assessing depletion.

1.6. Testing Self-Regulation

The studies adopted above have all employed similar structure where two different self-control tasks were employed in order to observe the effect of depletion. The first task is typically employed to induce self-regulatory depletion, whilst the second task measures the effect of depletion. Although both tasks usually appear to measure different modalities, they are assumed to engage self-control and therefore both require substance from the global self-regulatory resource. Notably, in the vast majority of studies that have adopted this paradigm, direct measurement of the first task is seldom tested; the effect of depletion is only inferred from group differences on the latter task.

1.6.1. Sequential task paradigm.

Depletion is produced through an initial task that is intended to deplete the general source. This is followed by a task measuring persistence or mental ability to establish depletion has occurred, that is, the dependent measure. These two supposedly unrelated consecutive tasks form the sequential task paradigm. Both tasks completed by the experimental group require self-control, whereas only the second task requires self-regulation under the control group (Lee et al., 2016). Ego-depletion is then a drop in performance for the experimental group on the second task, relative to the control group. This paradigm, however, is no longer held as the acceptable design standard of reliably inducing depletion (Francis, Milyavskya, Lin & Inzlicht, 2018). The sequential task paradigm critically relies on a number of assumptions.

Firstly, it assumes that for participants in the experimental group, the initial task requires effort and therefore participants expend most of their self-control resources on the first task with little left for the second task (Lee et al., 2016). Secondly, if participants are not motivated in the initial task, motivation and attention may remain on the follow-up task and as a result protect participants from depletion. Conversely, poorer performance on the secondary task may reflect a decline in motivation or a self-justified break by participants (Inzlicht & Schmeichel, 2012). Thirdly, the initial task may not be long enough in duration to adequately deplete self-control resources (Lee et al., 2016). Lee et al. (2016) suggest providing incentives to increase participant motivation to achieve well and ensure the initial task is long enough in duration to provide depletion effects on the second task. Further concerning is the lack of a pre-task prior to depletion. Because depletion effects are established across groups, rather than within individuals over time, performance on the post-test is inferred from the variance between groups. There is then no initial baseline measure to compare the follow-up test results. Finally, the sequential task paradigm critically assumes the depletion-inducing task and follow-up task are related by the requirement of self-regulation (Lurquin & Miyake, 2017), such as restricting responses under the letter-crossing task and persisting in button pressing.

1.6.2. Letter-crossing task.

One of the most commonly referenced self-regulation tasks is the letter-crossing task, or the e-crossing task, which requires participants to follow a set of rules while scanning through text (Baumeister et al., 1998; Tice et al., 2007). The short texts can be anything, from complex scientific writings (i.e., a page from an advanced statistical textbook) to short stories (Arber et al. 2017; Baumeister et al., 1998). The context or genre of the text for the letter-crossing task is often unspecified in the literature. Typically, participants are required to scan these short texts and cross off any '*e*' that may appear next to any vowel. Once this habit is learned, the participant must then regulate their behaviour with the addition of new rules (Baumeister et al. 1998). For example, cross-off any '*e*' that appears next to a vowel excluding '*i*'. The process requires ongoing self-control in monitoring their correct responses in regards to the numerous rules to follow (Baumeister et al., 1998). There is no method suggested on how to score the task as performance is rarely measured on the letter-crossing task. However, the strength model does predict that letter cancelling performance should deteriorate across time. One of the critical differences between this research and previous research conducted in the field is that this research employs the letter-crossing task as a measure. Previously, the letter-crossing task was administered as a manipulation task to induce self-regulation depletion, whereby task performance was deemed irrelevant and therefore not scored. As such, in the absence of measures of letter cancelling performance, depletion is inferred due to a detriment on the follow-up task under the experimental group compared to control group. For example, under the original experiment, Baumeister et al. (1998) administered a dull and boring movie following the letter-crossing task. Participants were asked to watch for as long as they could, as they would be questioned on it later. Persistence was then measured either through actively being required to hold down a button and release to quit the movie (i.e., more effortful), or to simply press a button to quit the movie (i.e., less effortful) (Baumeister et al., 1998). Participants were found to watch more of the dull movie when active quitting by pressing the button was required compared to passive quitting by releasing the button (Baumeister et al., 1998). These results are in line with previous research on persistence tasks where previously depleted individuals will quit earlier than non-depleted individuals in ongoing and sometimes unsolvable tasks (Tice et al., 2007). Similarly, the letter-crossing task requires persistent monitoring of answers, which results in the consumption of self-regulation resources.

Numerous adaptions of the task have been used to manipulate task difficulty. Reduced legibility has been induced by having the text being written by hand or having the page lightened through photocopying (Baumeister et al., 1998). A time limit is often introduced, rather than having participants complete the task in full. The administration method is usually by paper and pencil, rather than computerised methods. The employment of computerised versions of the letter-crossing task has been somewhat controversial (Baumeister & Vohs, 2016). As criticisms have followed when incorporating a computerised version of the letter-crossing task where participants are encouraged to click an item rather than physically cross it out (Baumeister & Vohs, 2016). Although, depletion effects have been replicated through employing computerised versions of the letter-crossing task (Arber et al. 2017; Dang, 2017; Sripada, Kessler, & Jonides, 2014) among other depletion measures (Schmeichel, 2007; Schmeichel & Vohs, 2009). For example, Arber et al. (2017) compared paper-based versus computer presented performance on the letter-crossing task, and more importantly explored depletion effects over time. The letter-crossing task containing five short stories was applied across several experiments. However, modifications on the application and instructions varied across experiments to investigate if the depletion effect could be replicated within the letter-crossing task and the depletion effects, the decline in letter-crossing accuracy over time, could remain consistent across administration methods. For scoring, correct and incorrect responses were recorded across experiments. Additionally completion times were recorded on studies 1 and 3. The first study replicated the original letter-crossing method, where ten minutes were allocated for participants to attempt the 13 pages of hand-written text and were required to physically cross out letters according to the original set of rules outlined by Baumeister et al. (1998). A depletion effect was found, as the findings reported a significant deterioration in mean proportion of targets identified correctly. The time-based requirement was eliminated for the second study in order to examine whether the depletion effect over the letter-crossing task originated from the explicit time limits. Participants were required to complete the task in full instead of the ten-minute time limit. The findings replicated the depletion effect from the first study (Arber et al., 2017). In order to identify whether this depletion effect under the letter-crossing task was reliant upon paper administration methods, a computer-based administration was introduced under the third study, with participants identifying target words by responding verbally instead of the physical crossing out of letters. The depletion effect was once again replicated (Arber et al., 2017). Under the fourth study, a special computer-based program was created to further establish methodological parameters for this effect. Responses were then given with a mouse-click response; "to click every single '*e*' presented next to a vowel". Thereby, "feel" then required two clicks compared to "true" which only required one '*e*' click. Participants were once again allowed ten minutes to complete the task. Computerised responses depicted a clear depletion effect trend once more on the letter-crossing task (Arber et al., 2017). This strongly suggests that both paper-based and computer administration methods of the letter-crossing task can induce the depletion effect within the self-regulation task. It should be noted, however, that the strength model does not explicitly state that depletion effects should be found within the self-regulation task, rather it advises that depletion effects are evident between the control and experimental groups after the self-regulation task has been administered (Baumeister et al., 1998). Importantly, this research aims to track depletion in real time, which requires the letter-crossing task to be treated as a measure and scored to observe potential depletion effects over the course of the task. Going forward, the letter-crossing task will first be considered by this research project as a measure, in which self-regulation is thought to be expended during the self-regulation task and thereby observed within the measure, but also evident immediately following the self-regulation task. Even though the letter-crossing task has been frequently employed, due to the ease of replication and adaptability across laboratories, the task's efficacy has been questioned as numerous replication failures on a variety of outcome tasks have been reported in the literature.

1.6.3. Meta-analytic studies of depletion.

Until recently, the depletion effect was considered a widely accepted phenomenon with consistent published works on the effect. Hagger and colleagues (2010), conducted a meta-analysis on the self-regulation literature and reported a robust moderate depletion effect size, d = 0.62, 95% CI [.57, .67], for 198 published studies. It was only when a stricter trim and fill method was employed to achieve a more accurate effect size that Carter and McCullough (2014) found no depletion effect. This null effect was thought to be influenced by small-study effect sizes, and publication bias (Carter & McCullough, 2014).

Following this null finding, Carter, Kofler, Forster, and McCullough (2015) then conducted a meta-analysis containing published and unpublished material in experiments that consisted only of frequently used experimental and outcome tasks in the literature. Tasks included the Stroop (1935) and complex span working memory tasks, among others. The results suggested the depletion effect was not consistent across tasks in the published studies, and that when accounting for small effect sizes the depletion effect was indistinguishable from zero. Additionally, the findings did not account for the limited strength resource model, as the results were shown to improve if more than a single manipulation task was included (Carter et al., 2015).

Most recently, Dang (2017) completed an updated replication of the meta-analysis conducted by Carter and colleagues (2015). Investigating the effective measures, Dang (2017) found working memory and attention video tasks did not have reliable effects of self-control. Tasks that were found to reliably induce a depleted state were the attention video, emotion video, and Stroop task (Dang, 2017). All of the tasks require some element of suppression. The attention video requires participants to suppress irrelevant information (i.e., irrelevant subtitles) whilst directing attention towards a silent video. The emotion video encourages emotional reactions of which participants must suppress the expression of emotions, whereas the classic Stroop task requires the suppression of automatic responses (i.e., reading word items rather than naming their ink colour). Dang (2017) found that the while the well-known letter-crossing task created significant depletion effects, it reported the highest heterogeneity in results, and suggested this may be due to multiple variations used by researchers. The limited strength resource model was again not supported by the meta-analysis results (Dang, 2017).

1.6.4. Task replication difficulty.

In addition to problems associated with meta-analyses, a number of high profile failures in replication have cast doubt on the veracity of depletion as a viable explanation for self-regulation failures. Sripada et al. (2014) conducted a study, whereby under a double-blind administration, participants were administered either a placebo or a methylphenidate capsule 60 minutes prior to the self-regulation experiment. Methylphenidate has been shown to increase cognitive functioning such as alertness and arousal by increasing dopamine and norepinephrine levels (Sripada et al., 2014). For this experiment, a modified version of the letter-crossing task where words were presented one at a time, with an average completion time of 7.5 minutes. The computerised letter-crossing task required participants to press a button when a word contained an '*e*', but to withhold their response if the '*e*' was positioned directly next to, or two letters away from another vowel (Sripada et al., 2014). The control group completed the same task, except were encouraged to respond to any '*e*' within a word regardless of the letter's position near another vowel. Participants then completed a modified multisource interference task for 10 minutes, where participants must respond to a target number which is unique within the three digits are presented on a screen; two are the same digit, one is different (e.g. 2, 3, 2). The target number is reinforced by a larger font size and is responded to by pressing a button. Reaction time is then measured. Under congruent trials, the target digit is relative to their serial position, for example the target '3' is presented in the third position. The digits under incongruent trials then do not correspond to their serial position. Additionally, the font size of the digit is no longer associated with the target digit. Thereby, stimulus-related cues, such as the serial position and digit size, are to be suppressed under the incongruent condition (Sripada et al., 2014). Importantly, the study suggested to have found self-regulation depletion effects can be manipulated through methylphenidate administration (Sripada et al., 2014), that is, depletion effects were eliminated with methylphenidate.

In response, Baumeister and Vohs (2016) did not accept the depletion-reversal result. Instead, Baumeister and Vohs (2016) argued that the modified letter-crossing task employed did not enforce a habit to be broken. Whereby the first rule is typically introduced on the first story (i.e., respond to any '*e*'), and the second rule is brought in on the second story (i.e., except if neighbouring next to or near another vowel), thereby creating a habit to be broken which exerts self-regulatory energy from the finite reserve. This was argued regardless of the findings that the placebo group managed to replicate depletion effects through slower reaction times on the multisource interference task. To calm any depletion effect doubts, Hagger and Chatzisarantis (2016) then attempted to conduct a large-scale replication of the Sripada et al. (2014) study.

Hagger and Chatzisarantis (2016) conducted a large-scale (*N* = 2141) replication meta-analysis across 23 laboratories with the sequential task paradigm and failed to produce a depletion effect. Even though tasks such as the letter-crossing task were considered by participants to be demanding, effortful, and difficult; this could not explain the small effect sizes found between the experimental and control groups, resulting in a lack of effect overall, d = .04 [-.07, .14] (Hagger & Chatzisarantis, 2016). The authors then suggested that the previously published ego-depletion effect sizes were inflated (Hagger & Chatzisarantis, 2016). Due to the difficulty in replicating the depletion effect, a number of researchers have begun to criticise the depletion methodology, including the tasks used to induce depletion.

Xu et al. (2014) found no depletion effects under commonly used self-regulation tasks, the persistent hand-grip and modified classic Stroop tasks, between control and experimental groups. This further suggests that by employing widely used and supported self-regulation tasks in the sequential-task paradigm between groups does not equate to the ability to reliably induce depletion.

Conversely, Healey, Hasher, and Danilova (2011) managed to replicate the depletion effect and found reduced scores in span tasks immediately following depleting tasks, however, this depleted state was dependent upon not the task similarity but the stimuli similarity between both tasks. A task measuring complex span such as operation span (OSPAN) did not suffer as a result of a demanding antisaccade task (Healey et al., 2011). The antisaccade task requires frontal lobe capacity by requiring the suppression of saccade (i.e., automatic eye movements). Participants are required to focus on a dot in the middle of a screen before a stimulus appears next to their fixation; the reflective saccade must be suppressed in order to focus vision in the opposite direction to where the stimulus appeared. Whereas, the OSPAN task requires the maintenance of memory for a list of items in their presented order, whilst responding to the accuracy of mathematical equations presented concurrently. This led the authors to argue the underlying complex span processes (i.e., the OSPAN task) were uninterrupted by resource depletion specifically under the experimental conditions.

Similarly, Dahm et al. (2011) only found depletion evident under young adults in an autobiographical memory task after completing a Stroop task. No depletion effects were found for older participants, and therefore Dahm and colleagues (2011) concluded self-regulation had developmental limits. Contrary to their initial findings, the young adult population in the control group out-performed older in the autobiographical memory task (Dahm et al. 2011). One consideration is that the older participants may have had a lower attentional capacity and therefore experienced fewer depletion effects as attention was employed at a lower level. This is in line with previous findings, as higher span capacity individuals are thought to employ greater attentional control during interference conditions (Kane & Engle, 2000). Notably, performance for higher capacity individuals dropped under proactive interference conditions on short-term memory tasks where attention was divided; unlike the lower capacity individuals who remained consistent, albeit a low performance, across attention conditions (Kane & Engle, 2000). This null effect is partially defended by earlier work from Schmeichel, Vohs, and Baumeister (2003), which asserts that a depleted self-regulation state selectively impairs executive functioning.

1.6.5. Selective depletion effect.

The strength model of self-regulation asserts that a global resource pool underpins a wide range of behaviours. A weaker version of the strength model suggests that only some cognitive processes are affected by depletion (Schmeichel et al., 2003). According to this selective depletion effect, tasks that require complex processing such as active executive control are thought to deplete a global resource shared by self-regulation (Schmeichel et al., 2003). This resource is not shared with lower level functioning tasks that do not require executive control, for example, persisting in drinking an unpleasant drink.

This selective depletion effect was illustrated in research conducted by Schmeichel, Vohs, and Baumeister (2003). Individuals who participated in emotion or attention regulating tasks were found to perform worse on tasks assessing logical reasoning abilities, but performed at a similar level to those under a control group on a general knowledge tasks (Schmeichel et al., 2003). According to Schmeichel, et al.’s (2003) results, any task measuring executive functioning should be affected as it requires a higher level of processing. This is particularly evident for tasks that require more cognitive effort, as this would place greater strain on self-regulatory resources, thereby resulting in a greater depletion effect. Based upon their results, Schmeichel et al. (2003) suggest that executive functioning and self-regulation may share the same energy resource.

1.6.6. Conceptual crisis.

Research in this field is underpinned by the notion that the tasks employed require self-control to some extent. Given that performance on the depletion task in these studies is very rarely examined, it is therefore difficult to understand whether the task is accurately recruiting self-regulation resources, or simply just appears to be. Also important to note is that the self-regulation resource is presented in the models as a limited global reserve. This reserve powers a number of tasks which drain the reserve of its' processing power. Other complex factors that are linked to the reserve are motivational drive, attentional control, and complex cognitive functioning.

As there is no current operational definition of depletion or detailed model of self-regulation processing, this means that any task, which appears to require self-control processing can be used in the research largely without question, and provided the follow-up task also requires self-control on a face validity basis, can be used to measure depletion regardless of the domain. Even if these criteria are met, depletion tasks range from being physically demanding to mentally demanding as they are assumed to utilise one ambiguous self-regulation resource. To address these conceptual issues, Lurquin and Miyake (2017) first suggest defining depletion operationally, which will allow the research field to appropriately measure an agreed upon concept of self-regulation. Furthermore, the employment of validated tasks with invalidated self-regulation tasks through correlational designs will help to validate self-regulation measures (Lurquin & Miyake, 2017). Thereby, if a conceptual link can be identified between validated outcome measures and not validated self-regulation tasks, this could help to further our understanding of depletion and the self-regulatory resource, and in turn validate self-regulation measures. Schmeichel et al. (2003) have suggested that measures of executive function may well be a suitable candidate for such shared resources. Hypothetically, all executive functioning tasks should be affected by self-regulatory depletion, as they all require higher cognitive processing and attentional control, and should thereby be deemed as sufficiently effortful and tiring. These standards fit the requirements for the recruitment and eventual depletion of self-regulation processes as identified by Schmeichel et al. (2003).

De Houwer (2011) makes a similar point to Lurquin and Miyake (2017). He critiqued the current methods in cognitive science, but presented a dual cognitive-functional analysis in order to reduce limitations in the field arguing that one needs to separate behavioural effects from the mental constructs that supposedly underpin them. One concern is that cognitive science often relies on inferring evidence to present a latent construct, as these measures are generally unconscious and cannot be directly measured. Evidence for such constructs then rely on measurable behavioural changes, from which cognitive theories are applied to. For example, working memory capacity is inferred from measurable behavioural constructs, such as the limit of items that can be recalled. The cognitive constructs, such as binding and maintenance processes, then help to explain these effects, that is, how information can be processed and maintained in memory. This method ultimately limits the conclusions that can be confidently made on these latent constructs. De Houwer (2011) then made the following suggestions for cognitive science to follow. Firstly, the need to separate conceptual terminology between constructs (i.e., different terms for cognitive and functional variables) to limit confusion. When observing effects, behavioural or functional markers should be operationalised for what is expected to be found. That is to say, a clear behavioural definition of the expected effect should be provided. Measurable evidence should then be described as behavioural, yet behavioural effects should not be acknowledged by cognitive constructs alone. Rather, behavioural effects should accompany cognitive constructs. Conclusions should be drawn by providing parallel explanations for behavioural theories on cognitive effects and vice versa (De Houwer, 2011). These suggestions are relevant to the present change in self-regulation literature and approaching the replication of the depletion effect. In short, one must first apply a functional approach by identifying the behavioural markers of a phenomenon. It is at this stage that cognitive approaches are appropriate and have the potential to explore the causal relations between the environment and behaviour.

In adopting this approach to the strength model of self-control, three behavioural markers can be derived using the sequential-task paradigm. Firstly, if resource depletion occurs through repeated acts of self-regulation, then performance decrements should be observable over time on the depletion task. Arber et al. (2017) have shown that target identification in the letter-crossing task, exhibits this functional marker. Secondly, if a global, undifferentiated set of resources underpins the depletion effect, then there should be shared variance between the tasks with the functional maker being evident through significant correlations between performance measures on depletion and outcome tasks. The third behavioural marker is the traditional one. Those who have expended more resources on the depletion task should show greater decrements on the outcome task. This thesis argues that all three behavioural markers need to be observed for the strength theory to be viable.

The functional approach attempts to describe the relationships between variables, not to explain them. The cognitive component of the functional-cognitive approach is to interpret the observed behavioural outcomes and, if need be, discriminate between possible causal relations. For example, the global reserve of the strength model predicts that there should be correlations between any combination of self-regulation and effortful outcome tasks. Conversely, the selective depletion approach would predict that correlations would exist only if both tasks involved executive function. Some combinations of outcome and depletion tasks would then not necessarily result in the behaviours markers being observed. Thus, following from Schmeichel et al. (2003), the thesis explores the relationship between executive function and the functional markers, acknowledging the fact that other causal mechanisms, such as changes in motivation, may also provide alternative accounts of the observed outcomes.

1.7. Summary of Chapter

In summary, the field of ego-depletion research is currently wrought with replication and conceptual issues. Behind the theory, the global self-regulation source is thought to control a number of things. One domain being cognitive functioning, according to Baumeister's theory (2002). The greater control required on the task through monitoring goals, employing motivation and self-regulation processes, the more depleted the resources will be (Hofmann et al., 2012). This aligns with the muscle basis of the strength resource model by Baumeister and colleagues. While there is physical evidence for self-regulation processes, conflict-responses have been mapped onto the ACC network (Inzlicht, Legault, & Teper, 2014), the support remains based on the vague theory construct of a limited domain-free resource of power that relies on such processes and are applicable to many different functions. Consequently, the current thesis makes use of De Houwer's (2011) distinction between functional and cognitive analyses. Three functional markers of depletion have been identified in the sequential task paradigm: deteriorating performance across time on the depletion task, significant correlations between performance on depletion and outcome tasks, and group differences between control and experimental depletion groups on the outcome task.

Several methodological inadequacies have also been identified in past research. Firstly, performance on the depletion task is rarely measured and even rarer still are measures across time. In the absence of such measures, it is difficult to establish direct evidence that the depletion task is in fact depleting resources. Secondly, when inducing self-regulatory depletion, the effect is determined from the performance on the follow-up task in the sequential task paradigm between experimental and control groups. Due to this method, there is no way of establishing the performance within individuals over time, as there are no initial baseline measures. Testing the three functional markers of depletion described above requires that the current methodological issues be rectified.

At the cognitive level of analysis, with weak theoretical underpinnings, almost any task regardless of modality can be applied to a sequential task paradigm to induce and measure self-regulatory depletion. However, Schmeichel et al. (2003) concluded that depletion should only affect effortful and demanding tasks, such as executive functioning, and as a result of this discrimination, they termed the 'selective depletion effect'. One problem with such a conclusion is that it suggests executive functioning is both well understood domain and that there is a commonality among executive function measures. In fact, executive functions are quite complex and often employ lower cognitive processes which are not effortful and should therefore not be susceptible to depletion effects. In the following chapter, the complexity of executive function is addressed, and the relationship between self-regulation and executive function are explored in more detail.

Chapter 2. Unity and Diversity of Executive Functioning

2.1.Overview of Chapter

The previous chapter suggested that research into self-regulation is currently problematic both at the methodological and theoretical level. At the theoretical level, the first chapter provided evidence linking self-regulation and executive functioning. The linkage between self-regulation and executive function in many instances does not consider the complexity and definitional issues raised regarding the construct of executive function. This chapter explores the complex domain of executive functioning, and the processes involved with higher order cognition. Whilst taking into account the major role of executive attention in a number of explanations for executive functioning models. The unity and diversity model of executive functioning encompasses the multifaceted nature of executive abilities (Friedman & Miyake, 2017; Miyake et al., 2000). Some relevant executive abilities are defined. This is followed by an overview of the importance of working memory processing as a concept that is sometimes accepted as a process part of executive functioning, but is sometimes classified within a separate domain. It is not the intent of the chapter to resolve the complexities in understanding executive function, rather it is to establish a set of theoretical constructs that may be relevant to self-regulation. Finally, the chapter concludes with the proposed relationship between executive functioning and self-regulation both at the functional level and at the cognitive level (De Houwer, 2011).

2.2. Executive functioning: Unity and Diversity

Executive function is the broad domain of higher-level processes that employ lower-level processes for complex cognition (Friedman & Miyake, 2017; Miyake et al., 2000). Typically, executive functions are employed for new tasks where learning is still occurring (Miyake et al., 2000). Once these tasks become automatic in nature, they require less higher cognitive processing and lower cognitive processing can then be utilised. Higher cognitive processing, such as common executive functions include planning, problem-solving, decision-making, goal-orientation, working memory, attentional control, updating working memory, inhibiting irrelevant material, suppressing material, shifting between mental task sets, and conflict-resolution amongst others (Barkley, 2001; Friedman & Miyake, 2017; Schmeichel et al., 2003). Lower cognitive processing includes functions that do not require executive processing and can be performed without much attentional control or pre-emptive thinking. These include tasks such as congruent conditions under the Stroop colour word task (Golden & Freshwater, 1978) like word-reading colour-naming subtasks, and simple short-term memory span tasks to name but a few.

Tasks that measure executive functioning have been neurologically mapped in the prefrontal cortex (PFC) and are associated with areas such as the anterior cingulate cortex (ACC) and dorsolateral prefrontal cortex (dPFC; Conway, Kane, & Engle, 2003). Where the ACC is activated during the presentation of incongruent stimuli under Stroop (1935) tasks and the resulting error interference, the dPFC is activated with goal-maintenance (Kane & Engle, 2003). These areas are critical under the Stroop task (Stroop, 1935). For instance, active maintenance of the goal in the dPFC is required in order to respond by naming the colour and employing the ACC to suppress automatic reading of the word (Kane & Engle, 2003). The ACC is then thought to house the conflict- or competition-resolution mechanism (Kane & Engle, 2003). This is the same neural area identified by Inzlicht and colleagues (Inzlicht et al., 2014), among others (Posner, Rothbart, Sheese, & Tang, 2007), where error-related negativity potentials occur due to conflicts between goals and behaviours for self-regulation. The common neurological pathways involved in a number of cognitive tasks suggest that there may well be a common set of executive function resources. The question of a single set of executive resources or discrete executive tasks has been the subject of considerable research.

Miyake and colleagues (2000) conducted a confirmatory factor analysis (CFA) to identify whether there were common processes shared between three executive functions: shifting, updating, and inhibition. These three functions were chosen because they were frequently used in the literature, had tasks strongly associated with their processing, and were clear in their processing requirements compared to other executive functions like planning (Miyake et al., 2000). Importantly, Miyake and colleagues (2000) argued that while these functions have been the focus of study, these functions are not considered to be exhaustive of the possible executive functions. With respect to the three executive functions they examined, shifting refers to the ability to shift one's mental set to another focus (Miyake et al., 2000); updating was defined as mental information to be updated and monitored (Miyake et al., 2000); and inhibition was identified as the inhibition of prepotent responses (Miyake et al., 2000). The task battery they administered to participants involved nine different cognitive tasks with three executive factors each being seen as exemplars of the proposed EF domain (i.e., shifting: plus-minus task, number-letter task, local-global task; updating: keep track task, tone monitoring task, letter memory task; and inhibition: antisaccade task, stop-signal task, Stroop task). The outcomes confirmed the relationship between the tasks and their underlying latent constructs, indicating that distinct EF dimensions existed. However, the CFA indicated that the three target executive functions shared common variance. When the statistical model considered the three target executive functions as three separable constructs this provided a better fit for the data than when the three executive functions were considered together as one construct. In simpler terms, the three executive functions should be considered as inhibiting, updating, and shifting abilities rather than one generic executive function generated from all three. Furthermore, the relationships between the three functions were only moderate, suggesting that the three functions were measuring different concepts but these concepts did slightly overlap with each other. The authors suggested that the three executive functions all required active goal information maintenance and inhibition or suppression of irrelevant mental sets, possibly in working memory (Miyake et al., 2000). From these findings, the three executive abilities should be classified as an executive function in their own right, but they do share similar processing within each other. These results suggest evidence for both the unity and diversity of executive functions.

In addition to the CFA outcomes described above, Miyake et al. (2000) also explored the relationship between the three executive functions identified and more complex executive function tasks. They used structural equation modelling (SEM) techniques to decipher which commonly used complex executive functioning tasks tapped into the three chosen target executive functions. Complex executive functioning tasks included the Wisconsin Card Sorting Task (WCST), Tower of Hanoi (TOH), Random Number Generation (RNG), Operation Span (OSPAN; Turner & Engle, 1989), and Dual Tasking. The WCST is thought to utilise flexibility in shifting ability as participants are required to shift their categorisation of cards based on colours and shapes presented on the cards and the experimenter's feedback (i.e., correct/ incorrect). Shifting ability was confirmed by the SEM results to provide the best explanation for executive ability under the WCST (Miyake et al., 2000). The TOH task requires participants to plan actions ahead of movements and is therefore believed to measure planning ability and goal direction. Of the three functions tested, inhibition was found to provide the best fit of the executive functions incorporated on the TOH task (Miyake et al., 2000). Inhibiting ability is associated to overriding the desire to make counterintuitive moves, in order to reach the end goal (Miyake et al., 2000). The RNG task was found to require both inhibiting and updating abilities (Miyake et al., 2000). Inhibition ability was thought to be related to the suppression of systematic responses (e.g. 1, 2, 3, 4…), while updating ability was linked to remembering recent responses and ensuring randomness in number generation (Miyake et al., 2000). Operation span was found to load onto the updating factor, through the storage and refreshing of new item information (Miyake et al., 2000). While the dual task was thought to tap into executive functions external to the three executive abilities tested (Miyake et al., 2000). These results show the complex nature of executive functioning tasks, such as the concurrent nature of independence and reliance between executive functioning abilities.

In one of the most recent attempts to update the classification of executive function, Friedman and Miyake (2017) summarised the results of many previous studies that evaluated inhibiting, updating, and shifting abilities. They noted that the CFA procedure had consistently shown positive correlations between these three constructs. When these correlations between constructs were explored and eliminated through SEM techniques, they found evidence for a common executive function latent variable that underpinned a diverse range of cognitive tasks. In addition to the common EF, they were able to detect independent latent variables for updating and shifting, but not for inhibition. The absence of a specific inhibition factor has subsequently been argued to suggest that inhibition is the most central facet of EF, and yet, it has also been argued that there is nothing unique about inhibition either (Friedman & Miyake, 2017). The existence of common and unique latent variables supporting diverse cognitive tasks, led Friedman and Miyake (2017) to opt for a unity and diversity model. This model suggests that there are common underpinnings to most cognitive tasks, yet there are also unique factors that are important to performance on some tasks but not others. Friedman and Miyake (2017) then argued that the shared variance between executive functions represented the ability to maintain and manage goals, particularly in the case where prepotent responses have to be managed or competition needs to be resolved (i.e., conflict-resolution). In the next section, the three components are described in more detail and then alternative views of what constitutes the common EF are discussed.

2.2.1. Executive Functioning Components

2.2.1.1. Updating ability. Updating ability is the renewing and monitoring of working memory representations (Miyake et al., 2000). This attribute is often measured using running memory span or n-back tasks, among others. The essential component of these tasks is actively manipulating relevant information through monitoring, coding, and updating new relevant information, rather than simply passively storing material in short-term memory (Miyake et al., 2000). These updating tasks also imply the forgetting of no longer relevant material (Miyake et al., 2000). Updating ability may also employ temporal tagging measures in order to maintain relevant information for recall (Miyake et al., 2000). Updating ability is associated with the dorsolateral prefrontal cortex (DLPFC) and has also been linked to the basal ganglia (Friedman & Miyake, 2017; Miyake et al., 2000). The tasks used to establish the updating construct are also closely related to the concept of working memory. Thus running memory span and n-back tasks are seen as exemplars of working memory tests. Perhaps the most commonly used working memory task is the Operation Span task (OSPAN; Turner & Engle, 1989) which requires responding to mathematical problems, whilst holding and refreshing word items for later recall.

Under the original study, Turner and Engle (1989) investigated the relation of reading comprehension to complex and simple span task performance. Turner and Engle (experiment 1; 1989) compared four complex span tasks: sentence-word span task (i.e., reading span task [RSPAN]), sentence-digit span task, operation-word span task (i.e., operation span OSPAN task), and operation-digit span task. Each complex span task required a processing response to the immediate stimuli (i.e., verifying sense in sentences or verifying correct answers within a mathematical operation) and storing an item (i.e., word or digit) following the initial stimuli. Processing skills were then determined by the immediate responses to sentences or operations, whereas storage skills were determined by the ability to maintain word and digit items in their serial order (Turner & Engle, 1989). These tasks were compared to simple span tasks of immediate word or digit recall, requiring the short-term memory recall of a list of words, or digits, in their serial order immediately after their presentation (Turner & Engle, 1989). Both complex and simple span tasks were compared with a reading comprehension test (Nelson-Denny Reading Test; Nelson & Denny, 1960), which established rate of reading and comprehension skills. Turner and Engle (1989) reported the complex span tasks, but not simple span tasks, were related to reading comprehension ability, particularly the complex-word span tasks. Numerical complex span tasks were correlated to reading comprehension when exclusively considering reading component skills, which are the word items under the OSPAN task. The second experiment modified the difficulty in the processing component of the sentence-word (RSPAN) and operation-word (OSPAN) complex span tasks into three conditions of increasing difficulty within participants. The sentence-word span task was modified by linguistic difficulty through passive or negative sentence content, and sentence length, of which were shown to decrease sentence comprehension (Turner & Engle, 1989). The operation-word span task was modified to lengthen the mathematical equations, and occasionally include a fraction (0.5) within the operation, of which both were shown to negatively affect arithmetic ability (Turner & Engle, 1989). The OSPAN task was administered from experiment two of Turner and Engle's study (1989), without the difficulty conditions. Word items were presented one at a time. Presented between word items is a mathematical equation consisting of two components; an initial multiplication or division with another number to add or subtract. A correct or incorrect answer is provided, of which participants must identify if this answer is accurate by responding "correct" or "incorrect". For example, "(9/3) – 2 = 1". Half the answers for the trials are correct, the other half are incorrect. The incorrect answers are often one or two places away from the correct sum. At the end of each trial, participants verbally recall the word items in their serial order.

While the OSPAN task has been identified as a measure of working memory capacity (WMC), which refers to the finite amount of information that can be processed at any one time, the latent variable analysis conducted by Miyake and colleagues (2000) found that OSPAN primarily loaded onto the updating ability but not the inhibiting, and shifting factors. These results suggest that some forms of executive function are closely linked to WMC.

2.2.1.2. Inhibiting ability. Inhibition is the internal deliberate control or suppression of prepotent responses (Miyake et al., 2000). Automatic or dominant responses may also be overridden through the facilitation of another response, for example, under the original and modified Stroop tasks (Golden & Freshwater, 1978; Stroop, 1935) participants are required to override the natural response of reading the word instead of naming the colour the word is written in.

Stroop's (1935) original experiments studied interference between stimuli and automatic verbal responses. Interference effects are established through reaction times in reading colour-names or naming the colours of 100 items as accurately and quickly as possible. The classic Stroop task (Stroop, 1935) generally contains three conditions: neutral, congruent, and incongruent. Neutral trials required participants to read aloud names of colours printed in black ink, or alternatively to name the colours of the inks (red, blue, green, brown, or purple) presented in coloured boxes or swastikas. Congruent trials required participants to name the colour a word is presented in. Importantly, this word corresponds with the colour it is presented in. That is, the word "red" is presented in a red ink. Congruency encourages a faster reaction time in participants compared to incongruent trials (Stroop, 1935). Incongruent trials present a colour name that does not correspond to the ink colour it is presented in. For example, "red" is presented in a green ink. This then creates the Stroop interference effect (Stroop, 1935). Experiment one found that overall reaction time was slowed by approximately 2.3 seconds when participants were reading colour-name words presented under incongruent conditions than congruent conditions (Stroop, 1935). Likewise, experiment two found a slowed reaction time of 47 seconds on average overall reaction time when participants were to correctly name the ink colour of a colour-name presented under incongruent conditions, when compared to naming colours presented in boxes under neutral conditions (Stroop, 1935). Experiment three tested performance across eight successive days under neutral (i.e., name ink colours of swastikas), congruent (i.e., reading colour-names or naming colours of colour-names), and incongruent conditions (i.e., reading colour-names or naming colours of colour-names). Practice effects were found to increase with each session through shortened reaction times across neutral, congruent, and incongruent conditions (Stroop, 1935).

Inhibition is typically associated with the right inferior frontal gyrus of the frontal lobes, this area could be in charge of monitoring for goal-relevant information, in addition to inhibition of irrelevant material and overriding incorrect responses (Friedman & Miyake, 2017). While multiple processes can be referred to as inhibition, this ability is a deliberate action that distinguishes it from natural inhibitive responses from negative activation in connectionist models (Miyake et al., 2000). However, there may be a form of sub-relationship between automatic and controlled inhibition (Miyake et al., 2000). The Stroop effect (Stroop, 1935), where there is a slower reaction time when naming the colour of the ink that a word (i.e., an incongruent colour name) is presented in, requires the inhibition of an automatic response to read the word. Reading the word, instead of naming the colour, is then the failure of executive control over inhibition.

As previously noted inhibition may be accountable for the common executive functioning factor due to the proposed influence over goal-maintenance and processing (Friedman & Miyake, 2017). If task goals are not refreshed and attended to, the likelihood of deviations from these task-goals is high, and as a result task performance will decline. Behavioural inhibition as measured by the Stroop (1935) task, is connected to the inhibition of prepotent responses (Hofmann et al., 2012). The aim of inhibition and the proposed goal-maintenance system shares similarities with the conflict-monitoring system in Baumeister and Voh's (2007) necessities for successful self-regulation, as referred to in section 1.4.2. Due to these links, inhibiting ability is arguably the closest executive function to self-regulation based on face value. If the self-regulation tasks accurately measure what they are assumed to induce and test, then self-regulation should be closely linked to inhibition.

2.2.1.3. Shifting ability. Also known as attention- or task-switching, shifting refers to transferring attention back and forth between mental sets, multiple tasks, or operations (Miyake, 2000). For example, one commonly used switching task is the number-letter task (Rogers & Monsell, 1995). In this task, number-letter pairs (e.g. *3C*) are presented on a computer screen in one of four quadrants. If the pair appears in of the top quadrants, then the participant must either decide if the number is odd or even. If the pair appears in the bottom quadrant, the participant must decide if the letter is a vowel or consonant. Thus, the decision the participant must make can vary from trial to trial, with the same decision being made on subsequent trials, or the decision changing from one trial to another (Rogers & Monsell, 1995). The shifting ability is inferred from the time it took to accurately respond under a switch between task sets and there is usually some cost in switching from condition to condition (Hofmann et al., 2012; Monsell, 2003).

Generally, switch cost tasks require attention and classification of stimulus features, in addition to memory retrieval or computation of the stimuli under some tasks (Monsell, 2003). Rules apply as to how the participant should respond to the stimuli and are changed repeatedly, such that new rules can be given as frequently as the beginning on each new trial. Goal- or error-monitoring is then required to meet the new set task's rules. Participant responses will reduce speed on switch trials compared to non-switch trials, of which the mental changing of sets is inferred (Monsell, 2003). In order to successfully complete the trials, participants may require shifting in attention between stimuli, retrieval of task goals and rules, and adjusting responses to meet the new criteria (Monsell, 2003). Higher switch-cost is thought to occur when switching is driven by internal cues, rather than external cues (Miyake et al., 2000).

While verbal, visual, and spatial shifting have been independently researched, they are similar in processing. Although, the neural circuits employed are not shared between them (Miyake et al., 2000). Shifting ability is particularly important for attentional control under models such as the SAS (Norman & Shallice, 1986). However, it is considered an unconscious process (Hofmann et al., 2012). Task-switching may be beneficial for achieving goals by identifying and leaving inefficient means for more efficient ones, or detrimental by goal-shifting which abandons the original goal for a more desirable one (Hofmann et al., 2012). Similar to the notion that inhibition may reflect a large component of the common executive functioning factor, Friedman and Miyake (2017) suggests individual differences in the speed of goal replacement may reflect shifting ability loadings onto this common executive functioning (EF) factor as well. Shifting ability may also be beneficial in the face of proactive interference or negative priming (Miyake et al., 2000).

2.2.3. Executive attention.

The Friedman and Miyake (2017) view of executive function argues for a common executive function component and subsidiary functions. While they suggest that the inhibiting factor might be the central component of the common factor, their preferred explanation involves goal-maintenance and top-down bias. That is, the ability to retrieve and implement goals at the appropriate time is essential to most cognitive tasks and if goal implementation is successful, this biases processing to reduce the effects of conflicting information. Moreover, they examined other possible explanations by evaluating their model with similar, but not identical views of executive function, including executive attention.

Executive attention, also known as attentional control, is the overriding of a dominant response for a subdominant response (Posner et al., 2007). An example of this is the antisaccade task, where attention must be drawn away from the visual stimulus (i.e., a dot appearing beside the focal point). The natural, or dominant response, is to focus attention onto the newly appeared item, whereas the task requires executive control to inhibit this response and instead move focus in the opposite direction to where the stimulus appeared. Executive attention is generally considered to be limited in capacity, thereby sustained attention significantly reduces its' resources. From an evolutionary perspective, involuntary attention is believed to be an automatic process that does not require top-down processing structures (Kaplan & Berman, 2010). Voluntary or executive attention, unlike involuntary attention, is goal-driven and requires effortful control to employ parietal and frontal cognitive structures (Kaplan & Berman, 2010). The ACC serves as an executive attention network that distributes focused control on cognition and emotion (Posner et al., 2007). Certain pathways within this neural network are activated depending upon the input modality required of the task. For instance, a connection to the visual system is activated when visual stimulus processing is required. Likewise, a connection to the limbic system is activated when emotional processing is required (Posner et al., 2007). These same brain structure connections have been observed across the lifespan and notably were discovered in infant error detection processing (Posner et al., 2007). Furthermore, the size of the dorsal anterior cingulate has been correlated to executive control ability (Posner et al., 2007).

Norman and Shallice (1986) put forth a unitary executive function based Supervisory Attention System (SAS), which frames the possible executive processing behind controlled internal and external actions. The SAS oversees task operations, monitors responses, and importantly attends to conflicts between goals and task stimuli (Norman & Shallice, 1986). Autonomous behaviours are not dictated by the SAS. Only when task actions are vague or have complexities, is the system necessary for functioning (Norman & Shallice, 1986). For example, the system would be important for item encoding under high interference conditions under serial recall tasks, where attention must be directed to the to-be-remembered stream and away from the interference items. The executive ability to shift between task and mental sets is therefore crucial for the SAS model. Under weak goal-reinforcement conditions, goal abandonment or neglect is brought on by the lack of goal-priming cues (Kane & Engle, 2003). Attending to goal-maintenance under interference conditions is therefore vital in order to limit goal uncertainty and response bias. However, the SAS is limited in capacity so sustained attention may eventually lead to a system failure (Norman & Shallice, 1986).

This idea of a central system controlling attentional resources is similar to Baddeley's working memory account (Baddeley, 2000; Baddeley, 1996; Baddeley & Hitch, 1974), where the central executive employs attentional control over processing. While limited in processing capacity, the central executive concurrently coordinates and integrates three sub-systems (the phonological loop, visuospatial sketchpad, and the episodic buffer), updating new information in long-term memory, switching focus of attention, and suppressing or inhibiting irrelevant material (Baddeley, 2000; Baddeley, 1996). Engle and colleagues (1999) also discuss a central executive being the source of controlled attention, inhibition, maintenance, and memory retrieval. Because of this, it is thought to be responsible for the strong relationship between WMC and higher cognition, such as fluid intelligence (*Gf*).

2.3. Working Memory

As is already evident, WMC is closely related to aspects of executive function and controlled attention. Unlike the other models, working memory is thought to reflect both active maintenance of information in memory and controlled attention for additional task processing (Kane & Engle, 2003). In contrast to the traditional notion of short-term memory with its emphasis on passive storage, working memory tasks involve both storage and processing functions. For instance, under the OSPAN task described earlier, memory is required to recall a series of words in a list, whilst attending to, solving, and responding to mathematical equations of which requires substantial cognitive processing. In addition to performing these tasks, items from previous trials, or lists, must be inhibited for accurate recall of the items in the current trial. Most span tasks then require the current items to be processed and maintained whilst suppressing irrelevant items or distractors. Subsequently, the WM system must manage no-longer-relevant information, such as controlling the build-up of proactive interference (PI) from previous trials influencing the current item set (Lustig, May, & Hasher, 2001). Susceptibility to PI relies on the suppression of PI items within the WM system (Lustig et al., 2001). Suppression, or inhibition, is therefore critical to accurate WM functioning (Lustig et al., 2001). Lustig et al. (2001) suggest, and have demonstrated, that individuals who are susceptible to PI perform lower on tasks involving rapid and accurate item retrieval.

The role of WMC in both simple and complex tasks has been well established. One methodology that has been adopted is to use the OSPAN task as a basis for splitting participants into higher (i.e., those who perform the task well) and lower working memory capacity groups (i.e., those who perform the task poorly). Group differences on other cognitive tasks are then used as the basis for making inferences regarding what tasks are related to WMC and what processes are involved. For example, Kane, Bleckley, Conway, and Engle (2001) propose that WMC is the ability to actively maintain information, such as a goal state. On an anti-saccade task, where attention was to be directed away from the visual stimulus and the desire to look at the cue was to be suppressed, high working memory span individuals were faster and more accurate than individuals with a lower span (Kane et al., 2001). However, under the pro-saccade task where the stimulus did not interfere with task goals, high-spans performed similarly to low-span individuals (Kane et al., 2001). Kane and colleagues (2001) argued that pro-saccade trials measured autonomous behaviour, where the cue directed attention and active goal-maintenance was not required. Conversely, the anti-saccade task can then be said to measure controlled attentional capability. Kane and colleagues (2001) concluded then that individual differences in

WMC refers to a general ability of attentional control in the face of interference and inhibition.

Similarly, Kane and Engle (2003) found individual differences in WMC were predictive of goal neglect and maintenance under the Stroop task (Stroop, 1935). Span groups were determined by WM span under the OSPAN task before the administration of the Stroop task (Kane & Engle, 2003). A Stroop effect was established as the rate of interference, namely the failure of the goal-maintenance mechanisms, were shown to increase with the higher percentage of congruent trials (75-80%) under the Stroop task (Kane & Engle, 2003). Individuals with lower WM spans were more likely than higher WM spans to respond by automatic responses, namely reading, in trials with higher congruency (Kane & Engle, 2003). This was identified when lower span individuals reported slower reaction times under trials low in congruency (0% - 20%) yet faster reaction times under higher congruency trials (75-80%; Kane & Engle, 2003). To explain this result, trials with a higher congruency percentage were believed to result in the weakening of participants' goal activation over time. Whereas, incongruent trials were thought to encourage goal-activation through the task due to the conflict between the word item and the ink colour. This conflict was thought to externally reinforce participants of the task goal; name the ink colour rather than reading the word item. Congruency trials in an earlier prior Stroop task were found to reinforce goal-maintenance on a follow-up Stroop task. If the second Stroop task was high in congruent trials, but followed on from a Stroop task that was low in congruent trials, lower interference errors occurred for lower span individuals than if the tasks were reversed. That is, slower reaction times were observed if a 75% congruent Stroop task was performed first, followed by a 0% congruent Stroop task (experiments 3A & 3B; Kane & Engle, 2003). Although, if the proportion of congruency was changed from 0% - 20% and 75% to 80%, both greater error rates and interference effects in slower reaction times were observed for lower spans regardless of Stroop task order (experiment 4; Kane & Engle, 2003). Higher span individuals remained consistent in their accurate responses, with low interference errors in reaction times, regardless of the congruency percentage in the Stroop task order (Kane & Engle, 2003). Higher incongruent trials in the initial task were thought to encourage goal priming in lower spans (Kane & Engle, 2003). However, this initial task priming effect was dependent upon a larger difference in congruency between the Stroop tasks within the sequence to occur (i.e., 0-75%, compared to 20-

80%). From these results, Kane and Engle (2003) suggested that individuals with lower spans have a slower processing competition-resolution mechanism, exclusive of goal-maintenance mechanisms.

From using the individual differences approach described above, it has been established that differences in WMC is predictive of higher cognitive abilities in addition to simpler processes (Kane & Engle, 2003; Lustig et al., 2001). Working memory capacity is considered to be the underlying factor in individual differences in Gf, with a strong correlation (r = .80 - .90) found between WMC and reasoning ability (Ackerman, Beier, & Boyle, 2002; Conway & Engle, 1994; Kyllonen & Christal, 1990; Suß et al., 2002). Additionally, WMC has been linked to language comprehension, memory retrieval, and life span development (Lustig et al., 2001). Although this WMC is thought to be limited. Higher capacity individuals are theorised to have a greater capacity when conserving resources and are more efficient in their allocation of resources, which can then explain their higher performances (Shamosh & Gray, 2007). Conversely, when individuals were under higher-pressure conditions whilst solving mathematical problems, Beilock and DeCaro (2007) found that individuals with higher WMC performed at an accuracy level similar to that of individuals with a lower WMC. Similar results can be found when higher span individuals are placed under PI conditions (Kane & Engle, 2000; Tehan, Arber, & Tolan, 2018). This strongly suggests that individual differences in complex cognition is strongly influenced by WMC.

In short, working memory supports goal-orientation, and protects against interference from irrelevant material through suppression or through conflict-resolution (Conway et al., 2003; Hofmann et al., 2012). Active maintenance and distractor blocking are then thought to be interdependent mechanisms in WMC amongst other attentional control processes (Kane & Engle, 2003). The working memory system is also thought to initiate competition-resolution mechanisms that are important for many cognitive tasks. The failure of these mechanisms results in interference and greater rates of errors (Kane & Engle, 2003).

2.3.1. Working Memory and the Binding Approach.

The description of WMC has obvious overlaps with the executive function literature. In fact, Friedman and Miyake (2017) recognise that their work on executive function emerged from the same tradition as the working memory literature. Moreover, the distinction between the common executive function factor, executive attention, and working memory capacity are hard to distinguish. According to Wilhelm, Hildebrand, and Oberauer (2013), the primary function of WMC is for building, maintaining and updating temporary, arbitrary bindings. Bindings are formed between items and their characteristics, such as ordinal position, for their later recall in span tasks. Therefore, complex span tasks such as OSPAN could be an example of the binding approach, as the task requires continuous updating of temporary bindings between each word that has to be remembered and its position within the list (Wilhelm et al., 2013). On one trial, new temporary bindings between items and positions have to be created, and maintained throughout processing. On the next trial, those old bindings must be updated through the creation of new temporary bindings between items and position markers. From this perspective, WMC is determined through interference among bindings. However, bindings are arguably employed under simple span tasks as well.

Unlike working memory tasks, short-term memory or simple span tasks do not require participants to manipulate information. Task items are simply stored for a short amount of time before being recalled. Short-term memory is the result from synchronized flexible recruitment of language, attentional, and serial order processing systems. To explain processing under complex span tasks, Unsworth and Engle (2007) put forth the dual-component theory, which proposes an active maintenance component, which they term primary memory (PM), and a controlled cue-dependent search and retrieval component, which they refer to as secondary memory (SM; Unsworth & Engle, 2007). Under the dual-component theory, items are sent from PM to SM when new or distractor items are presented which require processing (Unsworth & Engle, 2007). Items are then pulled from SM for recall. The reliance on components is dependent upon what the task requires of processing; simple tasks typically require PM more than SM, as less processing is required for the items to be shifted into SM than complex span tasks (Unsworth & Engle, 2007). Both components measure rehearsal, maintenance, updating, and controlled search (Unsworth & Engle, 2007). Binding ability is then crucial in SM for the processing of items. Recent notions have accepted the possibility of both working memory and SM, and in doing so, have formed a blend of both theories, where a greater capacity under SM may originate from a higher WMC (Wilhelm et al., 2013). An individual's WMC is then believed to represent the finite ability to maintain bindings and the ability to resolve conflicts that arise due to interference within these bindings (Oberauer, Süβ, Wilhelm, & Sander, 2007; Wilhelm et al., 2013).

Wilhelm et al. (2013) conducted a series of Confirmatory Factor Analyses (CFA) and Structural Equation Models (SEM) in order to identify the shared variance between underlying executive and WMC factors that were related to Gf. For present purposes, the key aspects of this research was the inclusion of switching, inhibiting, and updating measures derived from the executive function literature; a series of working memory updating tasks such as the recall n-back task; a series of complex span tasks including the OSPAN task; and a series of binding tasks. Importantly, the three binding tasks involved pairing stimuli from different domains. For the letter-colour binding task, this was a small set of letters were paired with a small set of colours. The letter-colour combinations changed from trial to trial, requiring the formation of new associations both within and across trials. The word-number binding task involved the pairing and repairing of a small set of words with a small set of numbers. The location-letter binding task required letters to be bound to spatial locations within a 3x3 matrix. For all three tests, the items on each trial were presented in random orders, and participants simply had to recall the pairing of a particular letter with a particular colour, a specific word with its paired number and a letter in its correct spatial location. The first outcome of the study was the confirmation that that the executive updating, working memory updating, and complex span latent variables where highly correlated (Wilhelm et al., 2013). This replicates the Miyake et al. (2000) finding of substantial shared variance between updating and complex span variables and provides evidence for the unity of executive functions. This suggested that updating information is a key cognitive variable denoting individual capacities (Wilhelm et al., 2013). In the second analysis, the binding measures were entered into the model. Binding was found to be virtually identical to the executive updating measure and correlated highly with the working memory updating and complex span measures (Wilhelm et al., 2013). Thereby, the final statistical model consisted of a common working memory latent variable that defined the binding and executive updating constructs.

In further analyses, they explored the executive factor of inhibition and the relationship of all previously added executive variables, and their relations to Gf.

This included inhibition tasks such as the Simon task and Erikson Flanker test. Although, the findings did not account for the executive attention explanation of WMC, where WMC is believed to inhibit interference effects, because of the lack of correlation between the interference errors under the incongruence conditions of the inhibition measures (Wilhelm et al., 2013). This finding therefore accounted for the diversity of executive functions. While there was a relationship between WMC and the tasks measuring inhibition, this relation was thought to be due to individual differences in reaction times and not a reflection of cognitive control through executive attention (Wilhelm et al., 2013). Rather, the best explanation for the observed pattern of results was based on the notion that WMC is used for building, maintaining and updating arbitrary bindings in presence of sources of interference, and that this is believed to reflect Gf (Wilhelm et al., 2013).

2.4. Executive Function, Executive Attention, and Working Memory Capacity: A summary.

The above review makes it clear that there is no clear conceptual distinction between executive function, executive attention, and WMC. The updated Friedman and Miyake (2017) position identifies a common executive function factor. This common executive factor is necessary for the formation, maintenance, and retrieval of goals that influence and direct subsequent processing (Friedman & Miyake, 2017). They also establish a set of unique aspects of executive function, some of which are clearly related to WMC. Another factor to consider is conflict-resolution, which works to inhibit prepotent responses and distractors. This factor has been incorporated into many cognitive tasks and consequently the ability has been accepted as a key individual differences variable. Lastly, the ability to form, maintain and update temporary bindings is considered another important determinant of cognitive abilities.

To the extent that executive function has been linked to self-regulation abilities warrants a brief examination of the self-regulation literature with respect to the different aspects of executive function. For a start, the assumption of the self-regulation literature that an undifferentiated reserve of, assumedly executive, resources underpins the self-regulation reserve becomes problematic given the evidence in favour of unity and diversity of executive functions. The strength model of self-regulation could only be maintained if what Friedman and Miyake (2017) describe as common executive function could be implicated in both depleting and outcome tasks. Moreover, if one accepts the diversity perspective of executive functions, depleting self-regulation resources with a task that targets the inhibition factor (e.g. the anti-saccade task), should have little impact on a task that involves updating of temporary bindings (e.g. the OSPAN task), as has been observed in the literature (Healy et al., 2011; Persson, Welsh, Jonides, & Reuter-Lorenz, 2007).

2.5. Relation to Self-regulation

As described in Chapter 1, there are a number of strong theoretical links between performance on executive functioning tasks and self-regulatory tasks, suggesting that the two systems either run parallel to each other and employ a similar reserve such as WMC, or could be considered to be part of the same system. Firstly, executive functioning and self-control tasks have been used interchangeably in the literature. Executive functioning measures have been used for inducing and testing the rate of depletion (Dahm et al., 2011; Healey et al., 2011). Both the Stroop task and complex span measures are frequent outcome measures in the meta-analyses conducted by Hagger et al. (2010), Carter et al. (2015), and Dang, Björklund, and Bäckström (2017). Furthermore, executive functioning measures have also been used as depletion tasks to induce decrements in self-regulatory functioning. For example, greater accuracy on the Stroop task (Stroop, 1935) have been correlated to the amount of time participants persist working on insolvable puzzles as reported by Kaplan and Berman (2010).

Secondly, while self-regulation processing has not yet been explictly mapped onto the brain, it has been hypothesised to stem from the planning, goal-orientated, higher cognitive areas. Such areas are the ACC and the dPFC, which are explicitly linked to executive functioning (Conway et al., 2003). If self-regulation and executive functioning are two separate systems, there is reason to assume the same networks are used for the processing of both systems.

Thirdly, some research has started to emerge that attempts to unpack the different components of executive function and their relationship to depletion. Dang et al. (2017) conducted a meta-analysis containing studies that involved a range of depletion tasks but importantly, studies that employed versions of the Stroop task as their outcome measure. They argued for the importance of goal-maintenance and response conflict-resolution, and that this would vary as a function of the number of congruent trials in the experiment varied (Dang et al., 2017). Specifically, in trials with greater congruency rates (>0% congruent), thereby the word reinforced naming the ink colour, the necessity to maintain the goal of responding to the ink colour and not the written word would be a key determinant of performance (Dang et al., 2017). Under trials which were solely incongruent (0% congruent), where the word does not reinforce naming the ink colour, goal-maintenance is externally reinforced and therefore is not problematic, however, active conflict-resolution would be required to limit poorer performance. It emerged that under 100% incongruent trials depletion effects were reflected in reaction time differences, but when the task involved a mixture of congruent and incongruent trials (i.e., >0% congruent), depletion effects were observed in the number of errors made (Dang et al., 2017). It was then argued that reaction time was the appropriate measure for the strength model of self-regulation and that goal-maintenance was outside the theoretical scope of this depletion model (Dang et al., 2017). As such, they argued that there was more compelling case for failures in goal-maintenance, through poor conflict-resolution, than self-regulation resource depletion under the Stroop tasks. Although some of the outcomes could have been compromised by publication bias and small sample sizes (minimum: 28; Dang et al., 2017). This research is important for current purposes because it attempts to simplify executive function into component parts (i.e., goal-maintenance and conflict-resolution), it contained multiple measures of self-regulation tasks, and it shows that depletion of the different components may not be identical. What the Dang et al. (2017) research did not do was to investigate performance on the self-regulation tasks. The current thesis extends this work by trying to unpack both executive function and depletion tasks. This will be achieved by holding the depletion task constant and varying the outcome tasks to reflect the different aspects of executive function.

2.5.1. A Dual Cognitive-Functional Analysis Approach

In the previous chapter, De Howuer's (2011) distinction between functional and cognitive analyses was introduced. Three functional markers associated of the strength model were identified. Performance on the depletion task should deteriorate across time as self-regulation resources are consumed. If the same resources underpin both depletion and outcome tasks, significant correlations should be observed between measures on the depletion task and measures on the outcome task. Finally, group differences between those whose resources have been more depleted than others, should be observed on the outcome task.

As the self-regulation task is rarely scored, the literature has not hypothesised how depletion effects should appear on the self-regulation task. It is therefore not clear whether multiple measures of the self-regulation task measure performance from the same executive reserve, measure different abilities from the same reserve, or whether these measures will show detriments in performance over time. Research conducted by Dang et al. (2017) found that depletion effects may appear differently across various measures. For instance, reaction time was negatively impacted following self-regulation, but not error rates (Dang et al., 2017). This suggested that depletion effects from the self-regulation reserve are different abilities; one ability may show depletion effects while the other ability remains intact. If the self-regulation resource recruits different executive abilities, this may resemble similar results to Dang et al.'s (2017) findings. Furthermore, these abilities should show relationships to executive performances on the outcome task. If the measures on the depletion task do not measure the same thing, then there should be no consistent relationships with the outcome task across the measures. A common executive function component, which underpins both the self-regulation and executive functioning tasks, should portray strong correlations between the two sets of tasks. However, if there are unique aspects of executive function to each of the outcome tasks, there may well be no uniform pattern of correlations, and secondly, as Dang et al. (2017) have shown, it is possible that some measures may show significant correlations and group differences on some tasks but not others.

Nonetheless, the potential relationship between depletion and cognitive tasks should also provide information at the cognitive level. By employing the dual cognitive-functional analysis, it should be possible to address the extent to which goal-maintenance, conflict-resolution, updating, inhibition or binding are involved in both depletion and outcome tasks. Furthermore, it should provide evidence that indicates what the self-regulation resources are and what they are not. At the very least, the patterns of functional outcomes that emerge across the different measures of the depletion task and the different aspects of the outcome tasks should allow some understanding of the relationship between self-regulation and executive function.

2.6. Summary of Chapter

In brief, the thesis acknowledges the diversity of executive functions. Miyake and colleagues (Miyake et al., 2000; Friedman & Miyake, 2017) have argued for both a common executive function component plus three major executive abilities, updating, inhibiting and shifting. The common executive function factor is not easily distinguishable from attentional control or WMC. In each case, goal-maintenance and response conflict-resolution processes are key aspects of all these approaches. In contrast, Wilhelm et al. (2013) have suggested that working memory is not involved in attentional control but is used to create, maintain, and retrieve temporary associative bindings. Moreover, alternative views concerning the role of goal-maintenance and response conflict-resolution also increase the complexity of defining and describing executive function.

The diversity of executive functions has implications for the strength model of self-regulation in that this model is founded on the assumption of a global reserve of energy, which could potentially be the common executive function factor, underpinning all demanding cognitive tasks. This chapter has identified three functional markers of the strength model and the possible relationships between self-regulation and executive function. In the next chapter, the operationalisations of these elements are described in detail.

Chapter 3. The Current Research

3.1 Overview of Chapter

Firstly, the theoretical and methodological issues surrounding self-regulatory research are revisited. This then introduces the theoretical standpoint that goal-maintenance and conflict-resolution can account for self-regulation processing and depletion effects put forth by this thesis. The overall rationale for the research is then discussed, including reasoning behind the executive functioning and self-regulation tasks chosen; rationale for using computer-based methods over paper-based methods; the argument behind employing an extreme-groups split design; and researcher degrees of freedom are presented. Operational definitions are provided to help configure theoretical underpinnings of the proposed model. The overall aims and hypotheses of the current research are then outlined, followed by the purpose and proposed outcome of the dissertation. This includes a brief overview of the experiments.

3.2. Theoretical and Methodological Issues, and Rationale

The case has been made in the literature that there is a strong relationship between the executive functioning domain and the field referred to as self-regulation (Hofmann et al., 2012; Schmeichel et al., 2003). The strength model argument of self-regulation proposes that individuals who have had their executive functioning depleted show a reduced self-regulatory capacity under complex thinking tasks (Hagger et al., 2010; Schmeichel et al., 2003). Interestingly, the greater the amount of executive functioning required, the greater the depleted self-regulation state (Schmeichel et al., 2003). In other words, self-regulatory capacity and executive functioning capacity are proposed to draw from the same limited processing resource, with executive functioning thought to stem from self-regulation (Baumeister, 2002; Muraven & Baumeister, 2000; Schmeichel et al., 2003). Alternatively, executive functioning has also been thought to be the global reserve powering the self-control domain (Hofmann et al., 2012). Arndt et al. (2014) further agreed that executive functioning appears to facilitate self-regulation, but added that self-regulation may also facilitate executive functioning, thereby suggesting a complex two-way relationship.

The above global reserve assumptions suggest that executive and self-regulatory resources are based on a general pool of resources, but as indicated in the previous chapter, there are unique aspects of executive resources. If there are common executive functions as Friedman and Miyake (2017) suggest, among other executive abilities, it remains to be demonstrated if self-regulation is related to the common executive factor or the other executive processes. The relationship between self-regulation and executive function is further clouded by theoretical overlap with other constructs such as attentional control and WMC, not to mention the roles attributed to goal-maintenance and response conflict-resolution in all these perspectives. Given Lurquin and Miyake's (2017) comments that the most substantial issue is that self-regulation has not yet been placed within recognised theories; the theoretical aim of the current experiments is then to explore the relationship between a commonly used depletion task, the letter-crossing task, and a series of executive functioning tasks that cover the broad range of abilities noted above.

The review in the earlier chapters also indicated that a functional approach to the veracity of the strength model could be adopted without necessarily specifying what the underpinning cognitive processes were. It was argued that there are three functional markers of the strength model: that performance on a depleting task should decline across time as a result of resource depletion; performance on the depletion task should be correlated with performance on the outcome task due to shared resources; and carryover effects on the outcome task. It is only this last functional marker that is typically assessed in ego-depletion studies between experimental and control groups. This thesis has identified, and aims to address, three major issues with the self-regulation literature:

Problem 1: Depletion has not been conceptually defined beyond a global reserve, which when depleted, results in a detriment to performance on any effortful outcome task. Any task that is assumed to employ some amount of effort is expected to require some resources from this global self-regulation reserve. As a result, a diverse range of tasks have been applied to induce depletion and to measure the effects of depletion. As the meta-analyses confirm the variety in combinations have produced varying results (Carter et al., 2015; Carter & McCullough, 2014; Dang, 2017; Hagger et al., 2010; Hagger & Chatzisarantis, 2016). One version of the global reserve assumption identifies this reserve as executive function resources (Hofmann et al., 2012). However, the diversity of executive functions is problematic from the perspective of a global reserve. This problem can be addressed through a considered selection of executive functioning tasks.

Problem 2: The fact that a wide variety depleting tasks have been used is problematic because performance on the depletion-inducing task is rarely measured and it is not clear that all tasks tax the global resource pool (if it exists). It is therefore unknown whether any task is a valid measure of self-regulatory resources, or if it is a reliable measure.

Problem 3: Associated with problem 2 is the fact that even if a depletion test was measured, such tasks can typically be measured in multiple ways. It is unknown what the most appropriate measure of resource depletion might be.

The original and significant contribution of this thesis is then to address these three problems by focusing on one commonly used depletion task, the letter-crossing task, in the following three studies. The research attempts to address the relationship between self-regulation as measured by the letter-crossing task and a range of abilities under executive functioning measures. This research provides an initial step in identifying whether self-regulation can be determined to be part of the executive functioning domain. As per the suggestions made by Lurquin and Miyake (2017), one method of approaching the conceptual crisis is to have an agreed upon definition of what self-regulation and depletion are. For the purpose of this dissertation, the definition of self-regulation follows the strength model of self-regulation (Baumeister, 2002). Self-regulation refers to the ability for an individual to conserve resource energy successfully, which is thought to power self-control and executive functioning (Baumeister, 2002). Depletion is therefore an unsuccessful attempt, or perhaps a strategic conscious effort, at conserving such resources in the face of a task that requires effortful control. As these concepts are not directly measurable (i.e., they are cognitive markers), this thesis then relies on measuring functional markers in order to understand the potential cognitive processes behind these behaviours, as recommended by De Houwer (2011).

The current set of experiments aims to measure depletion on the depletion-inducing task, that is, the letter-crossing task, and the outcome tasks. In order to do this, multiple measures of the letter-crossing task are determined and the functional markers of depletion effects are examined for each measure. Thus, the research addresses which measures reflect declines in performance over time, which measures are correlated with performance on executive functioning tasks, and which measures show carry over effects on the outcome task. On the post-depletion outcome tasks, depletion is defined as a decrease in correct items or an increase in errors from the pre-test scores. These functional markers effects then represent cognitive markers signifying depletion transfer effects due to the depletion induced in self-regulation resources from the letter-crossing task.

3.2.1. The letter-crossing task.

In the following experiments, the letter-crossing task has been used as the depletion inducing measure. The letter-crossing task is commonly employed to induce self-regulation depletion in the literature and was originally created by Baumeister and colleagues (Baumeister et al., 1998; Tice et al., 2007). In the classic letter-crossing task, participants are presented with printed text and are asked to physically cross off any *e* that may appear next to any vowel (*a, e, o, or u*), but not *i*. Classically, the initial rule must be introduced (i.e., cross off any e-vowel combination) and once a habit has been formed, typically after one short text, the second rule is introduced (i.e., do not cross *ei* or *ie* combinations). This new rule requires self-regulation through the inhibition of a learned response (Baumeister et al., 1998; Baumeister & Vohs, 2016). Under the Hagger et al. (2010) meta-analysis, 20 of the studies incorporated the letter-crossing task and this measure was associated with the largest depletion effect size (d = .77).

For this thesis, the letter-crossing task was the same modified version employed in studies 3 and 5 within the Arber et al. (2017). This letter-crossing task consisted of five different brief texts of which the characteristics are described the next chapter. Participants were to identify words containing e-vowel combinations (Rule 1) except for *ei* and *ie* combinations (Rule 2). This task was selected because it has frequently been used in the literature as a task in which participants must regulate their behaviour by inhibiting a learned response. It was also selected because it is amenable to measurement, and performance can be tracked across time. Arber et al. (2017) examined performance on this task and were able to show that accuracy in target identification declined across the five pages of text. This target accuracy measure provides one functional marker of resource depletion over time.

While the Arber et al. (2017) research centred on target identification, a number of other measures of performance on the letter-crossing task can be derived for each story. Apart from the number of targets correctly identified (*ae, ea, ee, eo, oe, eu* or *ue* combinations), the number of distractors (*ie* or *ei*) incorrectly identified can also be measured. This distractor measure is important because, at face value, it is a direct indicator of a self-regulation failure. While the task is not time-limited in the current experiments, participants are encouraged to complete the task as quickly and accurately as possible. From this, total time to complete the task could be recorded and average item processing speeds could be calculated to observe processing speed over time. For target identification, a global score of how accurately performance was maintained over the five stories served as one measure. The regression slope associated with changes in performance across the five periods provided a measure of the rate of decline across time. Thus, in the following experiment four measures of performance on the letter-crossing task are derived: overall target identification accuracy, the regression slope denoting the decline in performance in target identification across time (i.e., slope of accuracy), completion time or item processing speed, and the number of distractor errors (i.e., distractors).

From a theoretical perspective, it is possible that the different measures might reflect different aspects of executive functioning. For example, the common executive functioning factor would probably be best measured by target accuracy, and its decline across time be reflected in the slope of accuracy measure. Processing efficiency of the common executive functioning factor could be reflected in both the accuracy and time measures, with more efficient processes being more accurate and completing the task at a faster speed. The inhibition component could be reflected in the need to supress the tendency to respond to *ei* and *ie* combinations, with distractor identification being the appropriate measure. It could be argued that the letter-crossing depletion task could be measuring a unifying aspect of executive functioning: goal-maintenance. Goal-maintenance involves keeping the rules in mind such that both target accuracy and distractor identification could both be affected. Likewise, failures in response conflict-resolution could be reflected in distractor identification errors. It is not clear how executive abilities such as updating, switching, and binding abilities are directly involved in this task. None of the measures reflect the need to update memory components, participants do not need to switch between different components of the task, and no temporary associations need to be made. If this deduction is correct, there should be some conditions where some of the measures show correlations and depletion transfer effects with some outcome tasks, but not others. Therefore, the selection of outcome tasks must include conditions where relationships between the letter-crossing task and the outcome tasks should be present (i.e., inhibition) and conditions where the relationships should, hypothetically, not be evident (i.e., updating and binding).

3.2.2. Outcome tasks.

Before identifying and describing the selected outcome tasks, some issues associated with task selection need to be addressed. Firstly, while all the tasks are selected to measure some aspect of executive functioning, there is an issue with task-impurity (Miyake et al., 2000), that is, any cognitive tasks involves multiple processes. For example, one executive function task may not just utilise the considered ability, but employ multiple latent cognitive processes, which are not openly identifiable by the task. This means that two executive function tasks may be correlated because they both use verbal materials, for example, rather that both relying on the executive function. It is for this reason that executive functions and WMC have relied on procedures where latent variables are extracted from a range of related executive functioning tasks to infer these cognitive markers. Thus, while the current research aims to investigate a number of executive abilities, using single tasks to measure different aspects of executive function limits my concluding remarks. For instance, this thesis may not be able to claim that inhibition is exclusively affected by self-regulation if performance is found to decline on the Stroop task. One of the benefits of adopting De Houwer's (2011) distinction between functional and cognitive markers is that although identifying cognitive markers of a common resource might be difficult, functional markers are silent as to what the common resources or processes might be.

The second issue in choosing outcome tasks is that while there is evidence for depletion effects in commonly used executive outcome tasks, there are many instances where depletion effects have been difficult to replicate, such as operation span, Stroop, and other working memory tasks (Dang, 2017; Healey et al., 2011; Xu et al., 2014). In addition, in the Carter et al. (2015) meta-analysis, although there were significant depletion effects found on Stroop and working memory tasks, when uncorrected for publication bias, there were high levels of heterogeneity among the studies. It is important to note that some of the outcome tasks employed in the current experiment (Stroop and OSPAN tasks) deviate from more common versions of the task. The meta-analyses of the depletion literature, refer to section 1.6.3, point out that many of the executive tasks have been employed as both depleting and outcome tasks. If these tasks are in fact depleting, then it is possible that taking a pre-test measure would lead to depletion effects. As such, the choice of outcome tasks was based on finding the version of each task that could be completed in the shortest amount of time without sacrificing reliability. In short, while the outcome tasks have been selected because they have been used in previous ego-depletion research, it remains possible that depletion effects will not be evident in one or more of the current experiments.

3.2.2.1. Stroop colour word task (Stroop; Golden & Freshwater, 1978). Inhibition is often incorporated into depletion research on the assumption that participants must regulate their behaviour by inhibiting or suppressing a prepotent response (Wallace & Baumeister, 2002), such as the conflict described above. The Stroop task has been one of the most frequently used outcome measures in depletion research. The adapted Stroop task (Golden & Freshwater, 1978) requires participants to name as quickly and accurately the colour of the ink that words representing colour names are printed in.

From the executive functioning perspective, the Stroop task has been used as a measure of inhibition. Errors, or delayed responses, under the task have been explained as issues in conflict-resolution processing. The conflict arises from the desire to name the word rather than identifying the colour. However, Dang et al. (2017) also suggested that the Stroop task also involved goal-maintenance, for instance, identify the ink colour, and attempted to distinguish between goal-maintenance and response conflict-resolution under the Stroop task. Their meta-analysis indicated that depletion effects on the Stroop task were a function of whether reaction time or errors were the dependent measure and on the percentage of congruent and incongruent trials within the task (Dang et al., 2017), as outlined in section 2.5. Dang and colleagues (2017) concluded that there was more evidence for the goal-maintenance explanation than for the response conflict-resolution account.

In the current set of studies, the Stroop colour word test (Golden & Freshwater, 1978) has been adopted. The Stroop colour word task (Golden & Freshwater, 1978) is a standardised neuropsychological test that is widely used in examining cognitive impairment. This version of the Stroop task consists of three subtasks, namely word-reading, colour-naming, and colour-word naming. The word-reading subtask measures reading speed and acts as a control condition for the later incongruent subtask. The colour-naming subtask is another control condition that requires participants to accurately name colours. Lastly, the colour-word naming subtask consists of all incongruent stimuli, where participants must name the ink colour rather than read the incongruent colour names. This subtask is assumed to reflect the inhibition factor. Neither of the word reading and colour-naming subtasks should involve executive function processes, as these subtasks do not require executive abilities. As such, there would be no expectation that performance on these subtasks should be related to performance on the letter-crossing task. In contrast, there should be a relationship between letter-crossing ability and performance on the colour-word naming subtask, if executive resources, such as inhibition or goal-maintenance, underpin both.

3.2.2.2. Operation Span task (OSPAN; Turner & Engle, 1989). Working memory tasks have also been widely used in the self-regulation research. Complex span tasks have been the standard for measures of WMC research and feature prominently in experiments linking WMC to executive function and executive attention. The operation span (OSPAN) task has been adopted in the current set of experiments. This task consists of two components: a memory component as measured by serial recall and a processing component as measured by responses to mathematical equations. The dual-tasking elements of concurrent storing in memory and mathematical processing are why it is broadly accepted as a working memory measure. However, it is also clear that executive updating is also involved in the task, in that participants have to continually update and maintain multiple items in memory. They have to shift their attention continually between processing and memory components. They also must form new temporary associations between words and serial position cues, implicating the importance of binding operations. In short, the OSPAN task measures many of the executive attributes that form the basis of the current research. Furthermore, the OSPAN task has been a successful measure in reliably measuring depletion effects (Schmeichel, 2007). Processes in working memory are also responsible in goal representation and attention, suppression of irrelevant material, and the regulation of relevant items; all of which are thought to be critical to self-regulation processing (Hofmann et al., 2012).

Much of the literature that has been cited where the OSPAN task is involved, concentrates only on the memory measure (i.e., the serial recall of word items), as it is this component of the task that is geared to measure memory capacity. In the current experiments performance on the processing components of the task are also measured (i.e., the responding to mathematical equations). While, binding, updating, and memory components are not required in this aspect of the task, a common executive functioning component may be involved in the task. It is plausible that those who are efficient executive functioning processors may also be efficient mathematic problem solvers. Thereby, there may be some similarity shown in correlations between letter-crossing task performance and the processing and memory measures, but there may also be unique correlations involving the memory component only.

3.2.2.3. Immediate Serial Recall task (ISR). While complex span tasks have frequently been used in the depletion literature, simple span tasks have been less frequently used. One of the basic assumptions of the working memory literature is that complex span tasks such as digit span, or letter span, or immediate serial recall do not rely upon executive function. Instead, it has been assumed that these tasks reflect memory rehearsal mechanisms. Much of the evidence supporting this distinction comes from the demonstration that complex span tasks are better predictors of higher order cognitive abilities than simple span tasks (Unsworth & Engle, 2006, 2007). However, this general observation does not account for when the difficulty of the simple span task is increased. Thus, Unsworth and Engle (2006) showed that with short lists in the simple span task, complex span measures were better predictors of higher cognitive abilities than simple span measures. However, with list lengths of five, six or seven items simple span tasks were just as good predictors of higher cognitive abilities as complex span tasks. This suggests that as simple span tasks become more difficult, they tend to rely on those same cognitive processes that underpin complex span tasks.

While the individual differences in memory capacity literature argue for simple rehearsal processes underpinning simple span tasks, formal and computational models of immediate serial recall tasks suggest that more complex processes occur. Most models argue that memory for serial order is established by the temporary binding of items to position cues (Burgess & Hitch, 1999, 2006; Henson, 1998; Lewandowsky & Farrell, 2008). At retrieval, the position cues are used to generate potential candidates for recall with multiple candidates being generated for consideration (Burgess & Hitch, 1999, 2006; Henson, 1998; Lewandowsky & Farrell, 2008). A second step, where response competition is resolved to generate a single item for recall, is also an essential component of the recall process. Thereby, while common executive functioning abilities (i.e., inhibition, switching and updating) do not feature highly in the discussion of immediate serial recall, binding and conflict-resolution are heavily involved in explanations of this task.

In the following experiments an immediate serial recall (ISR) task is employed. On each trial participants recall five words in the order in which they were presented. Task difficulty is manipulated by varying the length of the words that have to be remembered. A robust finding in the serial recall research is that short words are better remembered than long words. Therefore, recall memory for two-syllable words are compared to three-syllable words and four syllable words in the current research. The working assumption is that executive function markers will not be observed on the trials containing shorter words. However, it might be the case that such markers will be present on the trials involving the four syllable words.

3.2.2.4. Proactive Interference Immediate Serial Recall task (PI-ISR). While the ISR task is not typically associated with executive functioning, the task can be adapted to test some of those functions. The fourth outcome task involves the ISR task where proactive interference is manipulated (Tolan & Tehan, 2002). Proactive interference is the intrusion of prior memories of items or information with current memories. In this task, each trial consists of either one block, or list, of four words or two blocks of four words. The one-block trials correspond to the simple ISR task where participants are presented with four words and have to recall those items in their serial order of presentation. In the two-block trials the first block is presented and this is followed by a *!*, which is a cue for the participant to forget the first block and concentrate on remembering the second block for recall. On these two-block trials, two items from the same semantic category are presented in the lists, one in the first block (i.e., a foil) and one in the second block (i.e., a target). The foil item in the first block is introduced to proactively interfere with the recall of target item in the second block. Thereby two sources of proactive interference may occur: the four to-be-forgotten items in the first block can interfere with the items in the second block, and the to-be-forgotten member of a category in the first block can interfere with the related item in the second block.

All trials involve the temporary bindings of items with their block and serial position within the block. The two-block trials also require updating the to-be-remembered items, the suppression or inhibition of the items in the first block, and some means of resolving response conflict between targets and foils. Goal-maintenance is clearly required to identify and maintain relevant items. Goal-neglect would then result in the intrusion of an irrelevant item or the breaking down of bindings of the relevant to-be-remembered stream of items to their serial positions.

Given the logic involved in the predictions involved in the ISR task, executive function markers should not be involved in the one-block trials but rather, should be evident in the two-block trials. That is, there is no expectation that performance on the letter-crossing task should be related to performance on the one-block trials. There should be relationships between the letter-crossing task and the two-block trials to the extent that both tasks share executive resources.

3.2.3. Reasoning for not including shifting ability.

Shifting, or switching, ability has not been included as an executive ability within the task set due to some concern over the importance of the function, as outlined in section 2.2.1.3. While there is a conceptual link between shifting ability and speed of goal replacement, which is assumed to reflect part of the common executive functioning factor, and a link with resistance to proactive interference and negative priming (Friedman & Miyake, 2017; Miyake et al., 2000), this ability is not distinct enough in what it measures, as there is a large conceptual overlap with the abilities incorporated into the current experiments: inhibition, updating, and binding. Furthermore, shifting ability is not considered to explain a large enough component of Gf (Friedman & Miyake, 2017), compared to the other executive functions. While it remains an executive function, important for switching between sets under executive functioning tasks, Friedman and Miyake's (2017) findings do not provide sufficient support to include this ability in the current research, however, it should be noted that due to task impurity within executive functions, shifting ability may be partially recruited under the proposed executive functioning tasks.

3.2.4. Research degrees of freedom.

The replication crisis in self-regulation literature has reinstated the need for strict methods in data collection and reporting, particularly within this field. Recommendations have thereby been put forth to reduce *p*-hacking and type II errors in published journal articles. It is recommended that a pre-determined rule for sample size be followed and reported, to limit continuing data collection until a significant finding is found (Simmons, Nelson, & Simonsohn, 2011; Wicherts, et al., 2016). All variables, experimental conditions, and observations included in the data collection and analysis, even if these manipulations failed, should also be reported (Simmons et al., 2011; Wicherts, et al., 2016). Furthermore, hypotheses should be explicitly provided, including the direction of the predicted effect; participants and experimenters should be blinded and randomly allocated to conditions where possible; and appropriate methods in cleaning the data and reporting the data to limit cherry-picking significant findings, among other suggestions (Wicherts, et al., 2016).

These steps were taken in order to reduce bias in the current data-collection, analysis, and report. Undergraduates acted as recruiters and experimenters for the current participant sample. This was done to separate the principle researcher and the experimenters to reduce expectancy bias in the data collection stage. Recruitment pre-determined the expected sample sizes with researcher degrees of freedom. That is, each student recruiter was asked to recruit a minimum of six participants. The aim of the total sample size was then to be approximately 516 participants. This recruitment method is outlined in section 4.4. Four stages of data collection occurred: each stage employed a different outcome task, however, the modified letter-crossing task was always employed as the depletion inducing measure. Each undergraduate was required to collect a minimum of six individuals, which determined sample size for each experiment. Control conditions were included in data collection, but as the experiments were determined to replicate depletion effects within the depletion-inducing task and between tasks under the outcome task results, control conditions were not the primary focus of the current research. Due to the research being exploratory in nature, the four executive functioning tasks were chosen based upon face value. That is, each measure employed a different executive ability to compare against letter-crossing ability. A separate measure, an immediate cued recall task with proactive interference, was included in the data collection but was not reported in the current studies. It was determined this cued recall measure did not significantly measure a different concept from the ISR task employed. Therefore this task was dropped from the analysis. Separate researchers (Madeleine Arber and Gerald Tehan) independently conducted the data analyses. A set of hypotheses and scoring methods were developed beforehand, prior to the data collection and analyses, to limit exploratory analyses and type II errors (i.e., false negative findings). Operational definitions and directions of the predicted effects were defined prior to the analysis, refer to section 3.2. Effect sizes were calculated prior to analyses and are reported where necessary for confidence in the findings.

3.3. Overall Aims and Hypotheses

The previous chapters have identified three problems with the self-regulation research in general, and the strength model in particular. These have been identified as the lack of specificity regarding the global pool of resources that underpin self-regulatory behaviour, the lack of measurement of depletion task performance in the sequential-task paradigm, and the lack of any knowledge as to what constitutes a valid and reliable measure of self-regulation on such depletion tasks.

To address these issues, De Houwer's (2011) distinction between functional and cognitive explanations for cognitive performance was utilised. Three functional markers have been identified that test the basic assumptions of the strength model. The first is that performance decrements should emerge across time on the depletion task. The second is that if depletion and outcome tasks share the same self-regulatory resources, performance on the two tasks should be correlated. Thirdly, any resource depletion noted in the depletion task, should have carryover effects on the outcome task that have been described above. At the cognitive marker level, the notion that self-regulatory resources are related to executive function resources was tested.

To test these ideas, a modified sequential-task paradigm was employed, where the letter-crossing task was the depletion inducing measure. A series of outcome tasks were selected on the basis that there are conditions where the relationships between the letter-crossing task and outcome task are expected to be observed, such as under inhibition conditions, and conditions where they are expected to be absent (i.e., binding). If an appropriate measure of performance on the letter-crossing task can be identified, it is expected that correlations and depletion transfer effects based on that letter-crossing measure should be observed in the outcome tasks where executive function is assumed to be in operation. Conversely, there should be no correlation or depletion transfer effects in those conditions where executive function is not thought to be involved.

3.3.1. Brief overview of experiments.

The following experiments employed a modified sequential-task paradigm, where an executive task was administered directly before and after the letter-crossing task for performance comparisons. The letter-crossing task was used in all experiments, but the outcome tasks changed across the experiments. The data were collected in four waves, reflecting the different outcome tasks.

The approach to reporting the outcomes of the research does not conform to the standard procedure of analysing each wave separately. Instead, the results are organised around the functional markers that have been identified and as such, studies are referred to rather than experiments. Thus, the first study explores performance on the letter-crossing task to establish which measures show deterioration across time. The second study explores performance on the letter-crossing task with performance on the pre-test measures of each of the four outcome tasks. The third study then looks at depletion transfer effects on each of the four outcome tasks. Therefore, the dataset will remain the same across the three studies in this thesis, however, the design of the analysis will change.

The first study will be observing the full dataset (*N* = 466) under one experiment, and the analysis will be focused on performance under the letter-crossing task alone. The following two studies will have four experiments each. Each of the four experiments in each chapter will incorporate one of the four executive tasks (Stroop, OSPAN, ISR, and PI-ISR), that is, Study 2 and 3 will incorporate the Stroop task in

Experiment 1 and 5, and will therefore contain the same dataset, but this data will be analysed according to the study aims.

3.3.1.1. Study one. According to the strength model of self-regulation, a decline in performance should be observed on the depletion task, in this case the letter-crossing task. Arber et al. (2017) have shown that target accuracy is one letter-crossing task measure that conforms to expectations. The first study then aims to replicate the deterioration over time within target identification accuracy, as first reported by Arber et al. (2017). Alternative scoring methods will be examined. These measures are overall accuracy of target identification, average item processing speed, the slope of the regression line measuring decreases in target identification over time, and the frequency of false distractor identifications. These scores will be observed over the five stories and the relationships with one another to examine independence between scores.

3.3.1.2. Study two. The aim of the second study is to identify what functional measures of the letter-crossing task are related to which aspects of the outcome task. The presence, or absence, of between task correlations will be used to assess what ability the letter-crossing task is measuring. This is to be assessed by a correlational design involving the letter-crossing task and the pre-test measures of the outcome tasks, that is, the Stroop task (Experiment 1), OSPAN task (Experiment 2), ISR task (Experiment 3), and PI-ISR task (Experiment 4).

3.3.1.3. Study three. The third study aims to explore the third functional marker of the strength model, which are depletion transfer effects from the letter-crossing task to the outcome tasks. Such effects would be reflected in a decline in performance from the pre-test to the post-test as self-regulatory resources are presumed to have been taxed. This study addresses which measures of the letter-crossing task are associated with the expected depletion transfer effects, with the expectation that such changes will be observed on those conditions in the outcome tasks where executive function is assumed to be in operation, but not in those conditions in which executive function plays little role. The study can then determine whether individual self-regulation performance translates into executive abilities. Explicitly, the experimental design assesses whether target accuracy, completion time, goal-neglect errors, and overall time taken on the letter-crossing task are associated with pre-test post-test differences account for inhibiting on the Stroop task (Experiment 5), the OSPAN task (Experiment 6), the ISR task (Experiment 7), or PI-ISR task (Experiment 8).

3.4. Summary of Chapter

To summarize, this chapter reiterated some of the fundamental issues associated with the self-regulation literature in general and specifically with the strength model. The methodology for how these problems were to be addressed, the rationale for selecting depletion and outcome tasks, resulted in a set of aims and hypotheses. These aims and hypotheses are articulated both at the functional and cognitive levels. A general description of the studies followed. The next three chapters represent the outcomes of these studies.

Chapter 4. Study 1

4.1 Overview of Chapter

The first fundamental assertion of Baumeister's strength model (Muraven & Baumeister, 2000) is that repeated acts of self-regulation deplete a limited pool of resources available for sufficient self-regulation processing. Logically, this implies that depletion occurs over time. At the functional level, this means that one should be able to observe a decrement in performance over time on measures of self-regulation. The aim of this chapter is to determine if this functional marker of depletion can be observed in the different measures of the letter-crossing task.

In examining performance on the letter-crossing task, a number of different measures can be taken: the overall accuracy of target detection, the rate of decline in performance across time, how fast the task is completed, and the number of times participants do not follow instructions and falsely identify a word containing *ei* or *ie* combinations as a target item. Arber et al. (2017) provided evidence that accuracy of target detection did show deterioration across time, but there has been no similar examination of alternative measures to date. Furthermore, it is not clear if these alternative measures are correlated with each other or are measuring different components of the letter-crossing task. Finally, it is not at all clear which, if any, of the measures are operational measures of self-regulation.

4.2.Background to Study

Arber et al. (2017) outlined how to reliably administer a computer-based letter-crossing task, and from this presented an accuracy method of scoring which showed accuracy declining over the letter-crossing task. The trend lines indicated that, as per the theoretical constructs around self-regulation depletion, this was a functional marker of depletion occurring within the depletion inducing measure. Other letter-crossing scoring measures were not incorporated into this study, and as this measure has not previously been scored, accuracy may not be the most effective method or the most informative method for observing changes in performance on the letter-crossing task.

Computer administration methods were deemed to be appropriate for the current studies. Importantly, the letter-crossing task employed for the current studies, has been previously found to replicate a depletion effect trend and show potential depletion transfer effects onto a follow-up OSPAN task (Study 5; Arber et al., 2017). The literature reports the administration of the letter-crossing task from paper to computer-based methods has no apparent detrimental effect on the structure in depletion effect replication, so long as the inhibition of a response remains (Dang, 2017; Sripada, et al., 2014). Furthermore, this ensured that all the tasks are administered in the same computer-based format.

4.3.Aim of the Present Study

As outlined in section 3.2.4., the data sets were collected in a number of stages. In each stage, although the outcome task differed, the letter-crossing task was always administered in the same way. As the letter-crossing task was administered in the same way to all participants, the current experiment is based on the performance for the total sample collected. This study has a number of aims. Firstly, to replicate the Arber et al. (2017) study showing that there is a decline in target identification over time. Secondly, to explore the degree to which the other letter-crossing scoring measures also show changes over time. Thirdly, to explore the relationships between the scoring measures. If each of these scoring measures is a functional marker of self-regulation resource depletion, it is expected that accurate target identification will decline over the letter-crossing task, portraying a downward slope of accuracy over the five stories. False identification of distractors is then expected to increase over the letter-crossing task. As the measure of time has not been previously assessed, it is unknown what the appearance of this measure will be under depletion, but it can be assumed that performance might slow with the increase in resource depletion.

4.4.Recruitment and Testing.

The data for the following experiment were conducted in four waves. In each wave, recruitment of participants was a course requirement for the students enrolled in a class devoted to developing basic practical research skills. Each student was asked to recruit six participants from their social network, and where possible ensure equal numbers of male and female participants, and a large age range. As such, each sample was a sample of convenience that was derived from the wider community.

Each wave involved the administration of the letter-crossing task and one of the four executive functioning tasks described earlier. A control group was collected alongside three of the four measures (Stroop, OSPAN, and ISR tasks), whereby the control participants completed the same measures, save for the self-regulation depletion-inducing task (the letter-crossing task). Although, this control group data is not the focus of the following experiments and was therefore excluded from the total data set. The characteristics of each wave are presented in Table 4.1.

Table 4.1

Number of Participants Recruited and Tested for Each Student Researcher (SR), and Total Number of Participants Analysed for Each EF Task

EF task	Number of Student Researchers (SR)	Number of Participants (P's) Recruited by each SR	Number of P's Tested	Number of P's in Analyses
Stroop Task	17	26 SR x 6 P's 1 SR x 5 P's 1 SR x 3 P's	164	98 experimental (1 removed) 65 control
OPSAN Task	24	26 SR x 6 P's 3 SR x 5 P's 1 SR x 3 P's	174	145 experimental 29 control
ISR Task	26	1 SR x 11 P's 5 SR x 10 P's 3 SR x 9 P's 16 SR x 6 P's 2 SR x 5 P's 1 SR x 4 P's 1 SR x 3 P's 2 SR x 2 P's 1 SR x 1 P's	165	109 experimental 56 control
PI-ISR Task	19	19 SR x 6 P's	114	114

In addition to recruiting the six participants, the student researcher also was responsible for testing the participants. The cognitive tasks and the letter-crossing tasks were all presented via computer, using Microsoft PowerPoint as the presentation mechanism. Instructions to the participant were included in the presentation to ensure that all participants were tested in the same manner. The role of the student researcher was to brief each participant, complete the consent process, start the experimental procedure, and to record the participant's responses where required. The student researchers then assembled the data they had collected on a prepared Excel file. The hard (i.e., response sheets) and soft data copies (i.e., Excel spreadsheets) were then returned on a specified due date. All student researchers were given the opportunity to practice on the tasks themselves and practice at administering the tasks.

There are three aspects of this procedure that are worthy of comment. The number of student researchers in the class determined the sample size. Adopting this process has lessened issues associated with power, in that the smallest experimental sample group has 98 participants. A series of G Power analyses were conducted to determine whether the smallest collected samples had sufficient power, this is reported where needed in the respective studies (Study 2 and 3). The procedure also minimises p-hacking issues that are aimed at improving the chances of observing a significant result. The analyses conducted were completed after all data had been collected. For the total sample a post-hoc power analysis employing G*Power software (Faul, Erdfelder, Buchner, & Lang, 2009) determined that N = 466 would have significant power for the correlation analysis (α = .05, β = 0, Power = 1, with a minimum critical F value of 2.37). With the exception of one participant, who was dropped during data cleaning due to insufficient scoring by the experimenter, the data from all participants were analysed.

4.5. Study 1

4.5.1. Method.

4.5.1.1. Participants. The total sample was 466 adult volunteers in the community. In total, 46.57% reported being female, with 9.01% opting to not reveal their gender. The average age was 42.75 years (*SD* =19.37) with a range of 18 - 90 years. There were no restrictions made on age, as developmental changes have been largely disregarded in the self-regulation literature when establishing self-regulation depletion effects, save for one study conducted by Dahm et al. (2011) described in section 1.6.4. Undergraduate students conducted recruitment as part of their course requirements.

4.5.1.2. Materials and procedure. The university human research ethics committee approved ethics for this study prior to the participant recruitment stage (Approval number: H16REA031). Informed consent was obtained prior to the start of the experiment (refer to Appendix B for the participant information sheet and consent form). Participants first completed a basic demographic questionnaire prior to instructions and test trials for each task. All participants completed a pre-test executive functioning measure, detailed in Study 2 and 3, followed by the letter-crossing task and a second version of the respective executive functioning measure initially completed for the follow-up task.

The letter-crossing task was that used in Arber et al. (2017). This version of the letter-crossing task was adapted from previous research where participants were required to search through short texts and physically cross-off any *e* which neighboured a vowel (*a*, *e*, *o*, *u*), excluding *i* (Tice et al., 2007). Specifically, the rules that governed responding were that the participant first had to name any word that contained an *e* preceded or followed by vowel *a*, *e*, *o*, or *u*. However, if the combination was an *ei* or *ie*, they were not to respond. Save for when both rules are applicable, such as *canoeing*, the second rule (i.e., do not respond to *ei* or *ie*) should be abandoned as the first rule (i.e., respond to any e-vowel combination) overrides the second rule.

To create the letter-crossing task, five short texts were selected from the internet (Retrieved from: http://www.shortbreadstories.co.uk/) and displayed one at a time via PowerPoint slides in size 10.5 Times New Roman font. Each story had an online popularity rating of 50% or more (story 1: 85%, story 2: 95%, story 3: 95%, story 4: 50%, story 5: 97%). The order of the texts remained the same for all participants and experiments. The first text contained 249 words with 19 e-vowel combinations (targets) to be verbally recalled and no distractors (‘*ie*’ or ‘*ei*’ items; Cook, 2012). The second text contained 293 words with 13 targets and 6 distractors (Das, 2013). There were 306 words in the third text, with 23 targets and 4 distractors (Myers, 2013). The fourth text contained 527 words with 20 targets and 5 distractors (Palmer, 2014). The final text contained 544 words with 39 targets and 9 distractors (Sham, 2013).

Participants were told to scan each story silently and verbally read aloud each identified word with a target item for the experimenter to record. The texts were to be searched at the participants’ own pace, but participants were explicitly told to complete the task as quickly as possible without making errors. Following the administration of instructions, participants were given a practice trial to ensure the instructions were understood. This practice trial was one paragraph with 11 items to be recalled. Once participants understood what was required of the task, they commenced the first text. The number of targets correctly identified (*a*, *e*, *o*, or *u* neighbouring an *e*), the number of distractors (‘*ie*’ or ‘*ei*’) named, and the time it took to complete each page were the dependent measures. The five stories completed in full, regardless of time taken.

4.5.2. Results.

4.5.2.1. *Changes over time*. Where an effect violated Mauchley's test of sphericity, the Huynh-Feldt correction was used as $\varepsilon > 0.75$, and the degrees of freedom were adjusted. The proportion of correct responses per story is depicted by the accuracy line in Figure 4.1. It is clear that target detection decreased in a linear trend across the five stories. A one-way repeated measure ANOVA indicated that there was a significant difference between the five stories in accuracy, F (3.56, 1655.78) = 158.63, $p < .001$, $\eta^2 = .25$, with the between story differences showing a strong linear component, $\eta^2 = .48$. These strong effect sizes are consistent with those reported previously that depicted performance detriments over the letter-crossing task (Arber et al. 2017; Study 2; $\eta^2 = 0.24$). Subsequent analyses indicated that accurate performance at each story was significantly different. False identification of the distractors remained under the 10% mark of all distractors on each story, which suggested participants were not adversely prone to these self-regulatory failures, and there also does not appear to be a strong increase in these errors across stories. A one-way repeated measure ANOVA indicated that there was a significant difference between the five stories in distractors, F (2.80, 1301.05) = 5.09, $p = .002$, $\eta^2 = .01$. Subsequent analyses indicted that fewer false identifications were made in Story 5 than the earlier stories and that there was no difference between Stories 2 to 4.

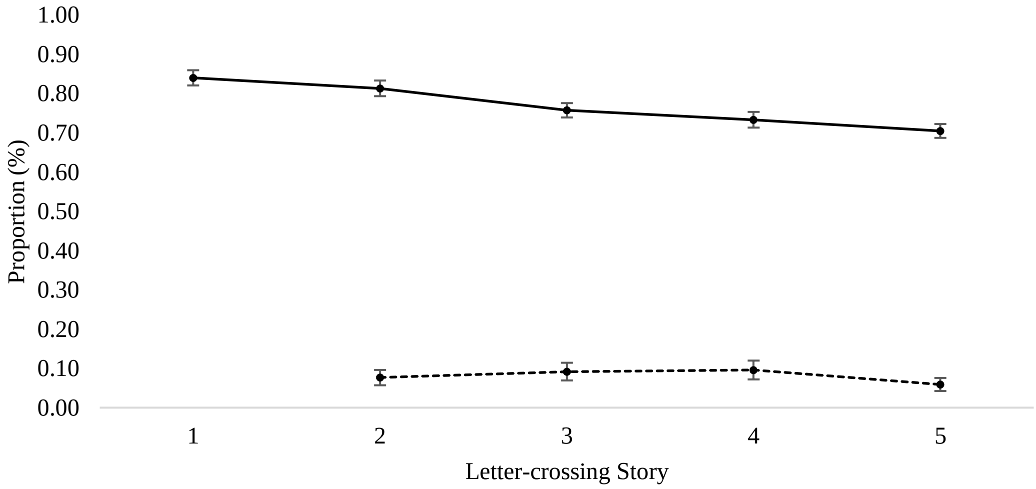


Figure 4.1. Letter-crossing task mean accuracy and distractor proportion scores across stories (1-5) for the total sample. Error bars are standard errors of the mean.

Time was investigated by the number of items analysed per second and is presented in Figure 4.2. By observing the figure, it appears participants generally showed an improvement in their average processing speed across the five stories on the letter-crossing task, although performance did slow under story three and five. A one-way repeated measure ANOVA indicated that there was a significant difference between the five stories in item processing speed, F (3.24, 1487.06) = 131.72, p < .001, η^2 = .22. Follow-up analyses confirmed that all pairwise comparisons were significant, save for the comparison between stories three and five. A bonferroni correction identified there was no significant difference in the number of items processed in stories three and five.

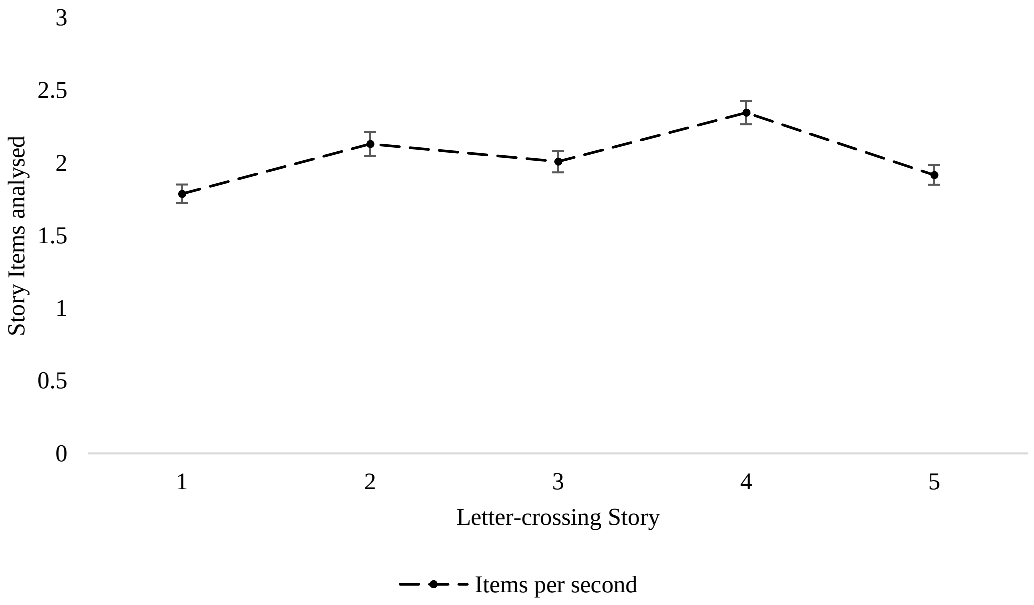


Figure 4.2. Letter-crossing items assessed per second across stories (1-5) for the total sample. Error bars are standard errors of the mean.

4.5.2.2. Relationships between measures. Pearson's *r* was calculated for each letter-crossing measure, in order to establish whether these scores were measuring related constructs. These values are presented within Table 4.2. Accuracy was found to be inversely correlated to slope of accuracy and distractors responded to. This suggested that declines in overall levels of accuracy were related to larger declines in target identification over time and to greater susceptibility to distractor responses. Item processing speed was not related to any other letter-crossing scoring method, suggesting that it was measuring a construct not accounted for by the other three measures. Slope and false identification of distractors were also related. Those individuals who showed rapid declines in performance across stories were also prone to making distractor errors. This may suggest that implicit components within the constructs, such as the slope measure, may be partially shared with other constructs, such as the distractor measure, and thereby could be considered to be partially related constructs.

Table 4.2

The Relationships Between the Four Measures of Performance on the Letter-Crossing Task

	Letter-Crossing Task			
	Accuracy	Speed	Slope	Distractor
M (SD)	.77 *(.15)*	2.01 *(.79)*	-.04 *(.04)*	.05 *(.13)*
	r	*r*	*r*	*r*
Accuracy	1.00	-.04	-.12**	-.11*
Speed	-.04	1.00	.03	.01
Slope	-.12**	.03	1.00	.11*
Distractor	-.11*	.01	.11*	1.00

Correlation is significant at the 0.01 level (**) and 0.05 level (*).

A final set of analyses that will become important in later studies, the reliability of three of the four measures were examined. These analyses confirmed a high internal consistency across the five stories for accuracy, item processing speed, and distractors (Cronbach's $\alpha = .89$, $\alpha = .93$, and $\alpha = .84$ respectively). There are no direct equivalent measures of slope, but the fact that it correlated with some of the other highly reliable measures would suggest adequate reliability.

4.6. Discussion

One fundamental proposition of the strength model is that self-regulatory resources should diminish over time, which should be reflected in behavioural changes over time (i.e., functional markers). The first assessment examined the changes in target identification, completion time, and the commission of distractor errors across the five stories. The outcomes for each of these measures were a decline in target accuracy, a general improvement in item processing speed, and fewer than 10% of distractor responses in each story, with no linear incline in the frequency of such responses. As such, the only measure that reflected the functional marker described by the strength model was target accuracy. This replicated the depletion in target accuracy over time pattern consistently found under the Arber et al. (2017) experiments. From these findings, accuracy of target identification appears to be the most appropriate measure for assessing depletion under the strength model of self-regulation.

Although, these scoring methods do not reveal whether this downward trend occurred due to naturally occurring omissions from failing to identify targets, or depletion-induced omissions from having limited self-regulation reserves.

There are other possible interpretations of the results. Given that the order of the stories remained constant for all participants, it is possible that the selection of stories involved an unintended confound of increasing difficulty. It is important to note that this specific letter-crossing task was piloted prior to being incorporated into published (Arber et al. 2017) and unpublished research. In the creation of the task, there was consideration into the letter-crossing story order. The first story contains no distractor items ('*ie*' or '*ei*') for self-regulatory failures to occur, as this method follows Baumeister's suggestion that learned responses need to be created prior to this habit being broken in an act of ego-depletion (Baumeister et al. 1998; Baumeister & Vohs, 2016; Muraven et al., 1998). The second story then introduces these distractor items for the learned habit to be broken under depleted self-regulatory conditions. Due to the standard administration of story order, whereby the task increases in difficulty from story one to five, it is possible then that any functional markers of depletion could arguably be task difficulty.

While the length of the stories increased over time (story one contained 249 words, whereas story five contained 544 words), participants were only asked to search for and respond to e-vowel combinations (excluding '*ie*' and '*ei*'). Thereby, while there was more text for participants to search through, the difference in items to identify between stories one to five was only 20 items. The average item processing speed generally increased over the five stories. This may suggest that a speed-accuracy trade-off occurred where participants improved on their item identification reaction time, but omitted accurate responses due to this faster processing speed. A speed-accuracy trade-off might better account for the current findings over task difficulty. It is fair to say that story two is harder than story one due to the introduction of distractor items. While participants were told to explicitly ignore distractor items, their presence purposefully added difficulty to the task as this is explicitly listed as a requirement for the replication of depletion effects. Although, distractor items only increased by three from story two to five. Finally, participants correctly identified above 75% of the items on average for each story and the proportion of distractors incorrectly identified remained under 10%. This suggests that participants found the task reasonably difficult, but despite this, participants remained to complete the task with high scores.

In summary, increased task difficulty or speed-accuracy trade-offs may account for the performance detriments over time instead of self-regulatory depletion. However, it is not clear in the literature if depleted self-regulation would appear as these performance effects within the self-regulation manipulation measure. As a result, these findings should be interpreted with caution and clearly future research needs to establish what the factors are that influence e-vowel detection There are other possible interpretations of the results. Given that the order of the stories remained constant for all participants, it is possible that the selection of stories involved an unintended confound of increasing difficulty, although what the determinants of e-vowel detection are have not yet been articulated.

The second aim of the study was to look at the relationship between scoring measures. Here, target identification was split into two measures, the average level of identification across the five stories (accuracy), and the rate of decline across stories (slope of accuracy). Likewise, an average score across the five stories was computed for processing speed and distractor responses. The outcomes of these analyses indicated that overall accuracy was negatively related to both slope and distractor measures. In other words, those who maintained a high degree of accuracy on the task showed lower levels of decline across time and made fewer false identification of distractors. Likewise, those who showed marked deterioration in target detection across time were also likely to make distractor identification errors. This pattern of relationships further suggests that target accuracy might well be the best measure of depletion. It is also clear that processing speed, or completion speed, is unrelated to any of the measures and is probably not a good measure of resource depletion.

4.7. Summary of chapter

In review, the accuracy scoring method provided the best functional marker for depletion effects in self-regulation. This score presented the expected observable detriments over time, and was complemented with inverse relationships with slope of accuracy over time and susceptibility to distractors. From these findings, it is safe to say that moving forward, accuracy presents the best measure for scoring the letter-crossing task and identifying whether depletion effects are present.

Chapter 5. Study 2

5.1.Overview of Chapter

The previous chapter examined performance on the letter-crossing task with the intent of determining which of the possible measures would exhibit the functional markers of self-regulation resource depletion. The outcome was that target accuracy was the only measure that exhibited the depletion trend. This chapter examines a second basic assumption of the strength model, that the self-regulation resources are domain-general. In terms of the sequential-task paradigm, this assumption would be reflected by shared variance between self-regulation and outcome tasks. The functional marker would therefore be reflected in significant correlations between measures on the letter-crossing task and relevant measures on a variety of outcome tasks. This chapter then explores the relationship between the four measures on the letter-crossing task, as determined in Chapter 4, and performance on the executive function tasks described in Chapter 3. To this end, performance on the letter-crossing task is compared to performance on the pre-test executive function tasks, when presumably self-regulation capacity is yet to be taxed.

This chapter also attempts to address the issue of what the common cognitive processes are, assuming that there are communalities among the letter-crossing task and the outcome tasks. Chapter 2 explored the cognitive construct of executive function and, while there are unique aspects of any executive function task, presented the unity and diversity account of executive functions (Friedman & Miyake, 2017; Miyake et al., 2000). Chapter 3 provided a rationale for selecting the Stroop, OSPAN, ISR, and PI-ISR tasks that measure different components of executive function. Each of four experiments will investigate one of the executive abilities. Experiment 1 investigates the relationship between inhibition and self-regulation. Experiment 2 assesses updating and self-regulation, whereas, Experiments 3 and 4 investigates the relationship between binding and self-regulation.

5.1.1. Study aims and design.

The literature has hinted at a variety of different relationships between executive functioning and self-regulation domains, where self-regulation was thought to empower the executive abilities (Baumeister, 2002; Schmeichel et al., 2003), executive functioning was thought to enable effective self-regulation processes (Hofmann et al., 2012), and a two-way relationship between these domains (Arndt et al., 2014). This thesis has identified three major problems with the self-regulation literature:

Problem 1: Depletion has not been conceptually defined beyond a global reserve, which when depleted results in a detriment to performance on any effortful task. This means any task that is assumed to employ some amount of effort is expected to require some resources from this global self-regulation reserve. Thus, the issue of the relationship between self-regulatory resources and executive function resources is still problematic.

Problem 2: It is unknown what the letter-crossing task is measuring. The tasks used to induce depletion have not been scored or measured previously (excluding the research conducted by Arber et al., 2017). Typically, depletion is only measured on the follow-up task and inferred from the difference in scores between experimental and control groups. Study 1 of this thesis assessed the different measures that can derived from the letter-crossing task, and the results suggested that accuracy in target recall exhibited the functional markers of resource depletion. Although, it did not address the issue of what resources were being depleted nor whether these resources were employed in other tasks.

Problem 3: As the letter-crossing task had not been previously scored before Arber et al. (2017) that included the accuracy measure, it is unknown whether the other letter-crossing measures (i.e., completion time, slope of accuracy, and distractor responses) are related to performance on executive function tasks. While slope, time, and distractor measures did not show the functional markers of depletion, these measures may reflect a shared domain with executive function tasks, and in fact may be better measures of executive function than self-regulation resource depletion. As this study is investigating the potential shared constructs between these two domains, it should be noted that in the following four experiments the executive functioning measures have included only the pre-test scores and not the post-test scores, so as not to compare depleted executive functioning scores with the letter-crossing task measures. Depletion carryover effects on executive functioning post-tests will be explored in Study 3. Likewise, total scores created from the five stories on the letter-crossing task have been used. For example, accuracy over the course of the five stories was averaged to present an average accuracy score.

Given the outcomes of Study 1 of this thesis, the primary expectation of the current experiments is that overall target accuracy should be significantly correlated to the primary measures of each of the executive function tasks, if the second functional marker of the strength model is to be realised. That is, a relationship should be identified between tasks if these abilities share the same domain. Currently, there is no compelling evidence to suggest that the other measures of the letter-crossing task should be correlated with the executive function tasks.

5.1.2. Power analysis.

A power analysis was conducted employing G*Power software (Faul et al., 2009) to identify whether the sample sizes collected in the following experiments were large enough to reject the null hypothesis. For two-tailed correlations, it was determined that *N* = 138 would be a large enough sample size to determine a significant effect (Power = 0.95). A further power analysis was conducted with the lowest sample size recruited (Experiment 1: *N* = 98), with the smallest significant correlation found under this experiment was *r* = .21. The G*Power software determined these values to have a power of 0.55, with a critical value of *r* = 0.19. Refer to Appendix C. Thereby, sample sizes of 98 and above were deemed large enough to have acceptable power for the correlational analyses. Nevertheless, alpha, beta, and power values were included in the analyses for clarity.

5.2.Experiment 1 – Stroop colour word task (Golden & Freshwater, 1978)

In the current experiment, the Stroop colour word task (Golden & Freshwater, 1978) was employed. This version of the Stroop task is a sensitive measure of cognitive impairment in inhibition ability, and is frequently incorporated into neuropsychological test batteries. The test manual provides reliabilities between .70 and .90, specifically the word reading subtask reported a coefficient of *r* = .86, the colour naming subtask reported a coefficient of *r* = .82, and the colour-naming subtask reported a coefficient of *r* = .73 (Golden & Freshwater, 1978). The task has three subtasks, two of which are control conditions (i.e., word-reading and colour-naming) and are unlikely to involve the use of executive function, while one subtask requires participants to override the prepotent response (i.e., colour-word naming). Under the word-reading subtask, participants are presented with a list of colour names printed in black ink and are required to read these words as quickly as possible. Under the colour-naming subtask, participants are given a list of coloured non-word stimuli (*XXXXX*) and asked to name the colour of the ink the stimuli are presented in as quickly as possible. As the colour-word naming subtask is the incongruent condition (i.e., the word *GREEN* printed in red ink), it is assumed to measure inhibition.

It is expected if self-regulation and executive function resources are related, or share the same resource, that letter-crossing task accuracy should be correlated to performance on the colour-word naming subtask. It is also anticipated that accuracy should not be related to performance on the word-reading or colour-naming subtasks as neither of these subtasks involve executive resources. Based on the Study 1 findings in this thesis, there is no expectation that any of the other letter-crossing measures will be correlated with any of the subtasks.

5.2.1. Method.

5.2.1.1. Participants. The participant sample was 99 adult volunteers, 50% reported being female. The average age was 45.5 years (*SD* = 20.5) with a range of 18 – 90 years, and 28% had studied at university. Undergraduate students at the University of Southern Queensland recruited community volunteers as part of an undergraduate course. Students were not graded on the data collection, but were awarded course requirements for participating in recruitment.

5.2.1.2. Materials and procedure. Informed consent was obtained from all participants before filling out a basic demographic questionnaire (i.e., gender, age, and education) and starting the experiment. Instructions and test trials were given to the participants prior to starting each task. Following the sequential-task paradigm, participants first completed the Stroop colour word task (Golden & Freshwater, 1978) according to the instructions provided in the manual. The instructions state participants are to read, or name the colour, for each of the items presented as quick as they can within the timeframe.

Each of the three subtasks were presented on a PowerPoint slide with 100 stimuli randomly distributed in a matrix of 5 columns and 20 rows. Participants were instructed to read down each column, rather than reading across the rows in each matrix. One test matrix (100 items in a 5 columns by 20 row matrix) was given to participants to practice on prior to each subtask to ensure they understood the instructions.

The word-reading subtask consisted of the words *red*, *green*, and *blue* presented in black ink. Participants were asked to read as many words as possible within 45 seconds, before the screen went blank. The colour-naming subtask consisted of the same matrix, with *XXXXX* as the stimulus presented in red, green, or blue ink. Participants were required to name as many colours the stimuli were presented in within the 45-second timeframe. The colour-word naming subtask consisted of the words: *red*, *green*, and *blue* presented in red, green, or blue ink. This subtask served as the incongruent condition, as the words were not presented in the respective coloured ink. For instance, blue was presented in red or green, but not blue ink. Participants were required to name as many colours of the stimuli ink within 45 seconds. If an error was made, participants were required to make the correct response before proceeding to the next item. The amount of stimuli read or named provided the dependent measure for each subtask. Participants then completed the letter-crossing task, as described in section 4.5.1.2., followed by the Stroop colour-word task (Golden & Freshwater, 1978), of which the results are reported in Chapter 6.

5.2.2. Results and discussion.

Study 1, in Chapter 4, employed a processing speed measure by converting total time to complete in order to compare time across stories. For this experiment, and all subsequent experiments, completion time measured in minutes was used as a more direct measure rather than a derived measure. Experiment 1 reported an acceptable level of power, $N = 98$, $\alpha = .05$, $\beta = .45$, Power = .55, with a minimum critical r value of 0.19. Pearson's r was calculated between the first session (pre-depletion) of the Stroop subtasks (word-reading, colour-naming, colour-word naming subtasks) and average accuracy, completion time, total slope, and average distractor recalls on the letter-crossing task ($N = 98$) are outlined in Table 5.1.

Table 5.1

Descriptive and Correlational Statistics Between Stroop (Word-Reading, Colour-Naming, and Colour-Word Naming) and Letter-Crossing Task Performance (Accuracy, Total Time, Slope, and Distractors Identified)

	Stroop Pre-Task			Letter-Crossing Task			
	Word-Reading	Colour-Naming	Colour-Word Naming	Average Accuracy	Total Time	Total Slope	Average Distractor
M (SD)	105.67 *(23.47)*	87.5 *(17.89)*	61.95 *(17.79)*	.77 *(.16)*	17.46 *(5.6)*	-.04 *(.04)*	1.86 *(4.36)*
	r	*r*	*r*	*r*	*r*	*r*	*r*
Word-Reading	1	.64**	.60**	.24*#	-.21*	-.03	-.05
Colour-Naming	.64**	1	.74**	.35**#	-.18	-.02	-.21*
Colour-Word Naming	.60**	.74**	1	.33**#	-.26**#	-.04	-.16

Note. Correlation is significant at the 0.01 level (**) and 0.05 level (*). # Variable contributes unique variance to the Stroop measure.

The correlations in Table 5.1 indicate firstly that the three Stroop subtasks were highly correlated. Secondly, those who maintained high levels of accuracy on the letter-crossing task were more accurate on all three subtasks. Likewise, those who completed the letter-crossing task quickly were more accurate on the word-reading and colour-word naming subtasks. Slope of accuracy appears to be unrelated to any of the measures, and distractor errors were related to the colour-naming subtask. Because there were multiple correlates on each subtask, which were correlated between themselves, hierarchical multiple regression was then conducted to explore the unique contribution of each letter-crossing measure. The accuracy measure was entered in the first step and any other significant correlates were entered in the second step. The outcome of this analysis is reported in Table 5.2. For the word-naming and colour-naming subtasks, only the accuracy measure contributed to performance. In the case of the colour-word naming subtask, both the accuracy and completion time measures contributed unique variance.

Table 5.2

Unique Contribution of the Different Letter-Crossing Measures to the Subtasks of the Stroop Task

Stroop Pre-Subtask	Letter-Crossing Measure	R^2 Change	*F* Change	*p* value
Word-Reading	Average Accuracy	.06	5.71	.019
	Total Time	.03	3.33	.071
Colour-Naming	Average Accuracy	.13	13.75	< .001
	Average Distractor	.03	3.80	.054
Colour-Word Naming	Average Accuracy	.11	11.87	.001
	Total Time	.05	5.29	.024

While the prediction that accuracy would be related to performance on the colour-word naming subtask was supported, based upon the assumption that both of these tasks employ the same domain, the predictions that they would not correlate with the non-executive function subtasks were not. That is, simpler subtasks such as word-reading and colour-naming were assumed to not measure inhibition and were therefore not originally predicted to correlate to the letter-crossing task. Instead, the pattern of correlations suggests that there is a common factor to all three subtasks of the Stroop task and this common factor is related to letter-crossing task accuracy. Speculation about what this common factor might be is left to the general discussion at the end of this chapter.

5.3.Experiment 2 – OSPAN task (Turner & Engle, 1989)

The OSPAN task (Turner & Engle, 1989) is a frequently used measure of WMC that has two components: the processing of mathematics problems and the simultaneous maintenance in memory of a small number of words in their serial order. Higher accuracy on the memory component of the OSPAN task is considered to reflect a higher WMC and Gf. Performance on the mathematics component of the task is not often reported, but in the current experiment both components of the task will be addressed, given the uncertainty regarding the most appropriate measures of the letter-crossing task. Given the outcomes of Experiment 1, a tentative expectation is that those participants who maintain high levels of target identification on the letter-crossing task will have scored higher in accuracy on both maths and word components of the OSPAN pre-task. Again, given the previous outcomes, the other letter-crossing task measures may not be related to the other letter-crossing measures.

5.3.1. Method.

5.3.1.1. Participants. The participant sample was 145 adult volunteers, 55.2% reported being female. The average age was 45.21 years (*SD* = 19.24) with a range of 19 – 64 years. Recruitment was conducted in the same manner as Experiment 1.

5.3.1.2. Materials and procedure. The procedure followed the same sequential-task paradigm as the first experiment. After gaining participant consent, participants filled out a basic demographic questionnaire, before receiving instructions for the first operation span task (OSPAN).

The OSPAN task consisted of eight trials made up of four simple arithmetic problems for immediate responding and four words for later recall. Sixteen trials were created by randomly pairing 64 maths operations and 64 word items. The operations consisted of two parts, the first part contained a multiplication or a division operation, followed by an addition or subtraction operation by a third number, and the total sum, for example, $4/2 + 1 = 3$. Half the mathematical operation trials were correct, the other half were incorrect. Incorrect answers were never more than plus or minus two from the correct answer. The words were one-syllable words between three to six letters in length. Half of the trials were used on the pre-depletion test; the latter was used on the post-depletion task.

Each trial began with an equation of which the participant had four seconds to decide whether this operation was correct by responding *true* or *false* before a word item was presented for one second followed by an equation item, and so on. The task was presented via PowerPoint at a programed pace in Calibri font, size 48. Each trial contained eight items presented one per slide; four simple arithmetic equations (correct or incorrect) and four word items.

After the instructions were outlined, participants were given two practice trials to ensure they understood the task instructions. For the word items, participants were encouraged to read each item aloud during presentation for easier recall when prompted by a row of question marks (*??????*). Participants wrote the four words on a participant response sheet at the end of each trial. The experimenter recorded participant responses to the mathematical equations. The study instructions stressed that the task was difficult and for participants to make the correct response to the mathematical equations and do their best on recalling the words in serial order. Participants were encouraged to leave a space blank if an item could not be remembered. The dependent measures were the number of words correctly recalled in their serial position and the number of correctly identified operations. Following the administration of the first OSPAN task session, participants completed the letter-crossing task, and then completed the remaining trials of the OSPAN task.

5.3.2. Results and discussion.

Experiment 2 reported strong power, $N = 145$, $\alpha = .05$, $\beta = .14$, Power = .86, with a minimum critical *r* value of 0.16. Pearson's *r* was calculated between OSPAN components (maths and word) and average accuracy, completion time, total slope, and average distractor (*ie* or *ei*) recalls on the letter-crossing task ($N = 145$) as outlined in Table 5.3.

Table 5.3

Descriptive and Correlational Statistics Between OSPAN (Maths and Word) and Letter-Crossing Task Performance (Accuracy, Total Time, Slope, and Distractors Identified)

		OSPAN Pre-Task		Letter-Crossing Task			
		Maths	Word	Average Accuracy	Total Time	Total Slope	Average Distractor
	M (SD)	26.81 *(5.8)*	19.92 *(7.29)*	.75 *(.16)*	17.06 *(8.23)*	-.04 *(.05)*	1 *(2.27)*
OSPAN Pre-Task		*r*	*r*	*r*	*r*	*r*	*r*
Maths		1	.28**	.46**	-.25**	-.03	-.15
Word		.28**	1	.37**	-.35**	-.02	-.15

Note. Correlation is significant at the 0.01 level (**) and 0.05 level (*).

The maths and word components of the OSPAN task were highly correlated with each other, again suggesting that there is some common cognitive process underpinning both tasks. Both components of the OSPAN task were significantly related to letter-crossing accuracy, and time measures. This suggests those who maintained high levels of accuracy and who completed the letter-crossing task quickly were more accurate on both the processing and memory components of the OSPAN task. In this experiment, neither the slope nor the distractor measures were related to either component of the OSPAN task. Again, hierarchical regressions were conducted to look at the contribution of the individual measures of the letter-crossing task. Refer to Table 5.4. In this instance, accuracy and completion times both made independent contributions to the two OSPAN task components.

Table 5.4

Unique Contribution of the Different Letter-Crossing Measures to the Components of the OSPAN Task

OSPAN Pre-Task Component	Letter-Crossing Measure	R^2 Change	*F* Change	*p* value
Maths	Average Accuracy	.22	36.34	< .001
	Total Time	.06	11.33	.001
Word	Average Accuracy	.14	22.24	< .001
	Total Time	.11	21.07	<.001

The outcomes of this study replicate a number of findings from Experiment 1. Firstly, the accuracy scoring measure is significantly correlated with all components on the outcome task. The results from Experiments 1 and 2 suggest that letter-crossing accuracy, whatever this ability is assessing, is accounting for performance on all subtasks of the Stroop colour word task and all components of the OSPAN task. Interestingly, letter-crossing completion time is also a contributing factor to both OSPAN task components and to the colour-word naming subtask of the Stroop colour word task (Golden & Freshwater, 1978). These three measures would be considered to be measures of executive function, whereas word-reading and colour-naming

Stroop subtasks would not. At this point, the data suggests that a combination of accuracy and completion time is related to executive functioning.

5.4.Experiment 3 – Immediate Serial Recall (ISR) task

In the previous experiment, WMC was assessed via the OSPAN task. The OSPAN task is an exemplar of complex span tasks which involve both processing and memory storage components. Complex span tasks are often opposed to simple span tasks, which are deemed to reflect short-term storage, rather than WMC, and because of this, simple span tasks are not considered to reflect executive functioning. Typically, simple span tasks involve presenting a short list of items that have to be recalled in serial order. While there is a memory component similar to complex span tasks, there is no processing component. The distinction between the two tasks is reduced when the simple span task is made more difficult. For example, immediate serial recall of short lists is not predictive of Gf, but recall of long lists shares as much variance with Gf as complex span (Unsworth and Engle, 2006). This experiment uses a simple span task where task difficulty is varied by the length of the words that have to be remembered.

In the ISR task, a prime determinant of performance is the length of the word as measured by the number of syllables in each word. Baddeley, Thomson, and Buchanan (1975) showed that memory lists containing one-syllable words was better than lists containing words with multiple syllables. This word length effect is a robust finding that has been since replicated (Tehan, Hendry, & Kocinski, 2001). In the current experiment, recall of 2-syllable, 3-syllable, and 4-syllable words are compared. The expectation is that recall of the shorter words (i.e., 2- and 3-syllable words) are not underpinned by executive resources and so should not correlate with the letter-crossing measures. For the 4-syllable words, where it is possible that working memory resources are involved, there should be correlations found with at least the accuracy measure of the letter-crossing task.

5.4.1. Method.

5.4.1.1. Participants. The participant sample was 109 adult volunteers, 56% reported being female. The average age was 36.5 years (*SD* = 16.65) with a range of 18 – 78 years. Five participants did not provide the gender and six participants did not provide their age. Recruitment followed the same procedure as Experiment 1.

5.4.1.2. Materials and procedure. The procedure followed the same sequential-task paradigm. Participants provided consent before filling out a basic demographic questionnaire and receiving instructions for the first task (i.e., the ISR task). For the ISR task, participants were presented with 18 trials of five items for their immediate recall. Three (2-syllable, 3-syllable, 4-syllable) 30 item word pools were created, and then randomly distributed into serial positions within trials related to their syllable length. That is to say, five 2-syllable items were placed in one trial together. On one third of the trials, the five items were two syllables in length, another third were three syllables in length, and the remaining trials contained four syllable items. The same items were used on the post-test, but the order was once again randomised to create a new set.

Participants were instructed to recall the five items in their order of presentation. Each item was presented individually in the centre of a PowerPoint slide via computer in Calibri font, size 48, at a rate of one per second. After each trial was over, signified by a row of question marks (*??????*), participants were required to write their responses in blank spaces on a response sheet. Participants controlled the onset of the next trial by pressing a space bar to begin. If an item was forgotten, participants were encouraged to leave the space blank. Following instructions, participants were given a practice trial with five items to ensure they understood the task. Participants then pressed the space bar to begin the task. The dependent measure was the proportion of items correctly recalled in their serial position. Participants then completed the letter-crossing task and the second version of the ISR task.

5.4.2. Results and discussion.

Experiment 3 reported strong power, $N = 109$, $\alpha = .05$, $\beta = .33$, Power = .67, with a minimum critical r value of 0.18. Pearson's r was calculated between pre-test ISR item lengths (2-syllable, 3-syllable, and 4-syllable) and average accuracy, total time, total slope, and average distractor (*ie* or *ei*) recalls on the letter-crossing task ($N = 109$) are outlined in Table 5.5.

Table 5.5

Descriptive and Correlational Statistics Between ISR Item Lengths (2-syllable, 3-syllable, and 4-syllable) and Letter-Crossing Task Performance (Accuracy, Total Time, Slope, and Distractors Identified)

	ISR Pre-Task			Letter-Crossing Task			
	2-Syllable	3-Syllable	4-Syllable	Average Accuracy	Total Time	Total Slope	Average Distractor
M *(SD)*	.48 *(.21)*	.39 *(.16)*	.33 *(.13)*	.79 *(.12)*	19.11 *(6.8)*	-.04 *(.04)*	1.24 *(2.78)*
	r	*r*	*r*	*r*	*r*	*r*	*r*
2-Syllable	1	.77**	.75**	.15	-.15	.01	-.14
3-Syllable	.77**	1	.79**	.17	-.09	-.04	-.12
4-Syllable	.75**	.79**	1	.23*	-.07	-.03	-.15

Note. Correlation is significant at the 0.01 level (**) and 0.05 level (*).

The standard word length effects emerged as reflected in the top row of Table 5.5. There were significant differences among the three lengths, $I\ (2,216) = 76.11$, $p < .001$, $\eta_p^2 = .41$. There were strong correlations across all item lengths within the ISR task, which suggested strong internal reliability. The only relationship found between the two tasks was 4-syllable item recall and letter-crossing accuracy as initially predicted, which suggested those individuals who recalled a greater number of 4-syllable items also performed at a higher accuracy level under the letter-crossing task.

This pattern of correlations is consistent with the distinction that Unsworth and Engle (2007) made regarding the discrimination between complex and simple span tasks, as simple span tasks increase in difficulty, they seemingly employ the same working memory processes involved in complex span tasks. The absence of significant correlations involving the easier trials, that is, the trials with 2- and 3-syllable items, is consistent with the view that under easy conditions WMC is not taxed. Working memory abilities are taxed under more difficult trials involving longer words (i.e., 4-syllable items). The significant correlation with the letter-crossing task then suggests that working memory resources are involved in both letter-crossing and complex, but not simple, processing under simple span tasks.

5.5. Experiment 4 – Immediate Serial Recall with Proactive Interference (PI-ISR) task

In the previous experiment, word length was used to manipulate task difficulty. In the current experiment, immediate serial recall is used again as the outcome task, but proactive interference is incorporated into the trials. This task is similar to the previous experiment in that immediate serial recall of a short list of words is required on each trial. It differs from the previous experiment in that the participant must continually update the most recent items studied by forming temporary bindings of the list items to context cues. Importantly, a foil item was inserted to proactively interfere with target recall. Participants need to then also override the involuntary response to recall the foil.

The task contained two sets of trials: block-1 trials consisted of only target items, and block-2 trials contained two halves or blocks, of which the first block contained the foil item, which was to be forgotten, and the second block contained the targets to be recalled. Accuracy under proactive interference conditions (i.e., the block-2 target recall) was expected to be strongly correlated to letter-crossing task accuracy. Likewise, block-1 target accuracy was assumed to be closely correlated to letter-crossing accuracy. This correlation was expected to be inverted for foil intrusions on the PI-ISR task, where foil intrusions represent susceptibility to proactive interference. As the task was once again automatic, block-1 and block-2 target recall was expected to be inversely related to total time taken to complete the letter-crossing task. Whereas, foil intrusions under the PI-ISR task, were expected to be related to slower self-regulation processing and should therefore represent a positive correlation with total time taken to complete the letter-crossing task. Accuracy on the PI-ISR task as represented by block-1 and -2 target recalls, were expected to translate onto the letter-crossing task. This was expected to be reflected by a significant correlation between block-1 and -2 targets and slope of accuracy on the letter-crossing task. Foil intrusions under the PI-ISR task were expected to be closely related to letter-crossing distractor recall, as individuals who are susceptible to proactive interference were predicted to carry over these effects and continue to be susceptible to self-regulation failures. Conversely, block-2 target recalls were expected to be inversely related to letter-crossing distractors.

5.5.1. Method.

5.5.1.1. Participants. The participant sample was 114 participants who were adult volunteers, of which 42.1% reported being female. The mean age was 43.45 years (*SD* = 20.28), with a range of 18 – 83 years. Thirty-six participants did not give their gender and age. Recruitment followed the same procedure as that described in Experiment 1.

5.5.1.2. Materials and procedure. The experiment followed the same sequential-task paradigm procedure as the prior experiments. Participants provided consent before filling out a basic demographic questionnaire and receiving instructions for the first PI-ISR task. The materials used in current experiments were those originally used in Tolan and Tehan's (2002) Experiment 6 or were items adapted from them. The participants all studied the same set of 20 trials, consisting of 15 two-block trials and 5 one-block trials. The one-block trials consisted of four words. The two-block trials consisted of eight words, four of which constituted the to-be-forgotten words and four were the to-be-remembered words. Two critical items (a target and foil) that were instances of the same taxonomic category were inserted into the two-block trials. The target item appeared within the second-block presented in each trial and was to-be-remembered. The foil was presented in the first block and was to-be-forgotten. Examples of two trials are presented below, where *swan* is the target item and *gull* is the foil item.

One-block: *bed tea white heart ??????*

Two-block: *oak bark gull ginger ! block limits swan powder 56 ?????*

Following the Tehan and Humphreys (1995; 1998) procedure, the targets and foils were created by selecting two instances from thirty different taxonomic categories from the South Florida Category Norms (McEvoy & Nelson, 1982). The block-2 targets were low dominant items within the category's hierarchy and the block-1 foils were high dominant items. The remaining items were created from a filler word-pool, which consisted of 300 words that came from the remaining categories from the South Florida Category Norms (McEvoy & Nelson, 1982) and the Shapiro and Palermo (1970) category norms. Multiple items from each category were selected as fillers. This ensured that there was no overlap in category membership between critical (i.e., target and foil) and filler items, but two items from a category could appear as filler items on a trial.

To create each two-block trial, the target item was randomly sampled without replacement from the target pool and inserted into the second position on half of the two-block trials, and in the third position on the other half of the two-block trials. The foil was placed into the first block respective to the target's ordinal position. Fillers were assigned to the remaining serial positions in each list. The one-block trials were created in the same way by assigning the target item to one of the positions and then randomly allocating filler items to the remaining positions. The 40 trials where then divided into two versions, each containing 5 randomised one-block trials and 15 randomised two-block trials. This set of trials was identical for all participants.

The first set of 20 trials were administered prior to the letter-crossing task. Each item was presented individually in PowerPoint on a computer in Calibri font, size 32, at a rate of one per second. Participants were told that each trial would consist of one or two blocks of four words and to only recall the last four items in serial order. In two-block trials, an exclamation mark (*!*) appeared after the first block that was a cue for the participant to forget the prior items and remember the next four items. Participants were unaware of whether trials were one- or two-block trials, and so were told to treat each new trial as if it were a one-block trial. On the two-block trials, after the final word a 2-digit number appeared on the screen for one second and the participant was required to read the number aloud. Reading the number served as a one second retention interval. Each trial ended with a row of question marks (*??????*) that prompted participants to write their responses in blank spaces on a response sheet. If an item was forgotten, the participant was encouraged to leave the space blank. Participants pressed the space bar to begin the next trial. Participants were given four two-block practice trials to ensure they understood the instructions before beginning the task.

Five dependent measures were derived from performance on the task. Overall recall of the four one-block items: overall recall of the four two-block items, the number of target items correctly recalled in their serial position in block-1 and block-2 trials, and any foils which were mistakenly recalled. The second set of 20 trials were administered following the letter-crossing task.

5.5.2. Results and discussion.

Experiment 4 reported an acceptable level of power, $N = 114$, $\alpha = .05$, $\beta = .43$, Power = .57, with a minimum critical r value of 0.18. Pearson's r was calculated between pre-test PI-ISR trial recalls (overall block-1 recall, overall block-2 recall, block-1 targets, block-2 targets, and foils) and average accuracy, total time, total slope, and average distractor (*ie* or *ei*) recalls on the letter-crossing task (N = 114) are outlined in Table 5.6.

Table 5.6

Descriptive and Correlational Statistics Between PI-ISR Trials and Letter-Crossing Task Performance (Accuracy, Total Time, Slope, and Distractors Identified)

	PI-ISR Pre-Task					Letter-Crossing Task			
	Block1 Overall	Block2 Overall	Block-1 Targets	Block-2 Targets	Foils	Average Accuracy	Total Time	Total Slope	Average Distractor
M (SD)	.67 (.23)	.58 (.26)	.65 (.28)	.57 (.26)	.04 (.06)	.76 (.16)	14.56 (8.42)	-.04 (.05)	.88 (1.87)
	r	*r*	*r*	*r*	*r*	*r*	*r*	*r*	*r*
Block-1 Overall	1.00	.59**	.84**	.69**	-.02	.29**	-.14	.01	-.18
Block-2 Overall	.59**	1.00	.55**	.94**	-.24**	.37**	.10	.13	-.26**
Block-1 Targets	.84**	.55**	1.00	.55**	-.03	.34**	-.07	.01	-.10
Block-2 Targets	.69**	.94**	.55**	1.00	-.28**	.36**	.14	.03	-.27**
Foils	-.02	-.24**	-.03	-.28**	1.00	-.09	-.20*	-.12	-.07

Note. Correlation is significant at the 0.01 level (**) and 0.05 level (*).

Mean recall is presented in the first row of Table 5.6. The effects of proactive interference were present in the task in that performance on the one-block trials was superior to performance on the two block trials for both overall item recall, t (113) = 4.65, $p < .001$, and target recall, t (113) = 3.61, $p < .001$. Correct recall for block-1 and block-2 items were strongly related to each other. Recall of the block-2 items were negatively correlated with foil intrusions.

Both block-1 and block-2 target recall was significantly correlated to letter-crossing accuracy. A negative relationship was found between block-2 target recall and letter-crossing distractor recall, suggesting those who recalled fewer block-2 targets were susceptible to a greater number of distractors. Foil intrusions on the PI-ISR task were negatively correlated to time taken to complete the letter-crossing task. This suggested those who were faster on the letter-crossing task made fewer block-1 intrusions on the pre-test PI-ISR task.

Again, hierarchical regressions were conducted to look at the contribution of the individual measures of the letter-crossing task. Refer to Table 5.7. In this instance, accuracy and distractor responses both made independent contributions to both overall and target recall.

Table 5.7

Unique Contribution of the Different Letter-Crossing Measures to the Block-2 Target Recalls Under the PI-ISR Task

PI-ISR Pre-Task Trials	Measure	R^2 Change	F Change	p value
Block-2 Overall	Average Accuracy	.15	18.96	< .001
	Average Distractor	.06	8.74	.004
Block-2 Targets	Average Accuracy	.13	16.77	< .001
	Average Distractor	.05	6.83	.010

5.6.Study 2 General Discussion

For the current set of experiments, the focus centred on both functional and cognitive markers of the strength model. The functional marker in this case involved identifying significant relationships between executive functioning and measures of the letter-crossing task. At the cognitive level, the experiments were designed around varied measures of executive function abilities: inhibition, updating, binding, and binding under interference.

5.6.1. Inhibition.

Experiment 1 assessed the relationship between inhibiting ability and self-regulation performance. There is a clear theoretical overlap between inhibiting ability and self-regulation, as self-regulation typically requires the restriction or inhibition of action, thoughts, or behaviours. Both the word-reading and colour-naming subtasks under the Stroop colour word task (Golden & Freshwater, 1978) served as control conditions to ensure that participants could accurately read the words on the screen and name the colours. Inhibiting ability was therefore not required until participants were asked to ignore the natural response of reading the word, and instead name the colour under the colour-word naming subtask. Thereby, a strong relationship was predicted only for the incongruent condition. Under the incongruent condition, that is the colour-word naming subtask, relationships were identified with letter-crossing accuracy and a shorter completion time on the letter-crossing task. While this aspect of performance is consistent with the strength model, this interpretation was weakened by similar significant correlations between overall accuracy on the letter-crossing task and performance on both the word-reading and colour-naming subtasks that do not require inhibition. This suggests that the shared ability between tasks were not accounted for solely by inhibition. While it is clear that the two tasks have shared processes, there is no strong evidence to support the notion that these resources are essential in inhibiting prepotent responses.

5.6.2. Updating.

Experiment 2 assessed the relationship between updating ability and self-regulation performance. The word component on the OSPAN task (Turner & Engle, 1989) required updating and holding of this information whilst processing the mathematical equations, which required a simple true or false response. Performance on the word component was therefore more critical in identifying updating ability. Both components of the OSPAN task were significantly related to letter-crossing accuracy, and time measures. This suggests those who maintained high levels of accuracy and who completed the letter-crossing task quickly, were more accurate on both the processing and memory components of the OSPAN task. Again, the lack of differential effects for processing and memory components suggests that while the letter-crossing task and the OSPAN task have some processes in common, they are not specific to updating resources, since there is no updating requirement in the processing (i.e., maths) component of the task.

5.6.3. Binding.

Experiment 3 assessed the relationship between binding ability and self-regulation performance. As binding ability is not directly acknowledged as an executive function, it was hypothesised that any relationship identified would appear different to that of the executive functions (inhibiting, updating, and competing against proactive interference). It was suggested that the 4-syllable items would require a higher binding ability as the word length effect has documented shorter words are easier to recall than longer words (Baddeley et al., 1975; Mackworth, 1964). The original prediction for binding ability in longer items was met, where only 4-syllable item accuracy showed a relationship with letter-crossing accuracy. A relationship between 2- and 3-syllable accuracy and letter-crossing ability was noticeably absent from the results as expected.

5.6.4. Binding under interference.

Experiment 4 assessed the relationship between updating and binding ability against proactive interference and self-regulation performance. Those with a greater ability to resist proactive interference were assumed to have greater self-regulation ability. Performance under the block-2 trials were therefore critical for identifying a potential inhibiting ability for proactively interfering items, which were predicted to be employed under self-regulation abilities (i.e., fewer foil intrusions and greater block-2 targets). The results suggested binding ability under proactive interference (i.e., block-2 trials) and non-proactive interference conditions (i.e., block-1 trials) were closely related to letter-crossing accuracy. The original prediction that greater target accuracy under proactive interference conditions would result in greater self-regulation ability was met, as block-2 target recalls were negatively correlated to letter-crossing distractor recalls. Furthermore, susceptibility to proactive interference (i.e., foil recalls) was found to be related to slower letter-crossing task processing. That is to say, participants who recalled foil intrusions on the PI-ISR task took longer to complete the letter-crossing task.

5.6.5. Overall.

One of the aims of this experiment was to determine which of the measures of the letter-crossing task were related to performance on the outcome tasks. The majority of relationships to executive abilities involved letter-crossing accuracy. Completion time did not consistently correlate with the outcome tasks, nor did distractor responses. The slope measure had no relationship to any of the outcome tasks. In short, neither completion time, slope, nor distractor measures show the functional markers of the strength model. It is possible that the lack of correlations between depletion and executive functioning tasks could be due to high values on the letter-crossing measures. Correlation values will be lower, and thereby non-significant, when less variance is found in the scores. This reduces the ability to establish whether there has been a relationship between the executive functioning and self-regulation measures (i.e. the restriction of range problem). This is possibly the case for the distractor measures, but is unlikely to be the case for completion time or slope measures.

In contrast, the overall accuracy measure did show significant correlations with all task measures except for immediate serial recall of 2-syllable and 3-syllable words, and foil intrusions on the PI-ISR task. Letter-crossing accuracy was found to share processes with all of the subtasks of the Stroop colour word task (Golden & Freshwater, 1978). This included the control conditions that were not expected to correlate with letter-crossing abilities. This strongly suggested that a common factor of executive functioning, and not just inhibition, was related to letter-crossing performance. Likewise, the accuracy measure shares variance with both components on the OSPAN task (Turner & Engle, 1989), the 4-syllable words on the ISR task, and with all measures except foil intrusions on the PI-ISR task. This suggests that updating, binding, and binding under proactive interference employ similar processes to the letter-crossing task. Thereby, the overall accuracy measure does show the expected functional markers associated with the strength model; that self-regulation resources are domain general.

In terms of cognitive markers, searching for targets on a letter-crossing task shares variance with a range of other cognitive tasks that measure quite different processes. Such as speed of processing on the Stroop task, arithmetic processing and memory requirements on the OSPAN task, and memory for serial order in immediate serial recall. Of which, the presence of significant correlations across this diverse range of tasks is not easily accounted for by method variance. Moreover, this letter-crossing measure does not involve any of the unique inhibition, updating, and binding characteristics of executive function tasks. Instead, this shared variance has all the characteristics of the common EF factor identified by Friedman and Miyake (2017). Given the broad conceptual overlap between the common EF factor, executive attention, and WMC, the shared variance may also reflect the potential executive attention and WMC underpinnings of all tasks. This is explored further in Chapter 7.

5.7.Summary of Chapter

The current study examined performance on the letter-crossing task with the intent of determining which of the possible measures would exhibit the functional markers of self-regulation resource depletion. The outcome was that target accuracy was the only measure that exhibited the expected characteristics. This chapter examined a second basic assumption of the strength model that the self-regulation resources are domain-general. This assumption would be reflected by shared variance between self-regulation and outcome tasks. The outcomes were consistent with such expectations only when the accuracy measure on the letter-crossing task was used. Thereby, the outcome of Study 1 in Chapter 4 was that the overall accuracy measure shows the pattern of effects predicted by the strength model. Target accuracy declines across time on the letter-crossing task and this measure is correlated with key measures on the outcome task. The results of this chapter also point to a shared domain as being the common EF factor, executive attention, or WMC. Although, the data do point towards a common resource domain affecting both letter-crossing and outcome tasks, it has yet to be demonstrated that this common EF factor is related to self-regulation. In the next chapter, the third functional marker of the strength model is addressed. That is, the expectation that transfer effects occur from the depletion task onto the outcome task should be observed.

Chapter 6. Study 3

6.1 Overview of Chapter

The current chapter contains Study 3, which includes four experiments investigating the shared processing between executive functions and self-regulation as measured by the letter-crossing task. The strength model of self-regulation suggests that performance should decline over time across tasks, resulting in a lowered performance immediately following the depletion task. Thereby this chapter will investigate the third functional marker of the transfer of depletion effects from the letter-crossing task to the post-depletion executive functioning tasks. This chapter adopts an individual differences approach, specifically the extreme-groups analysis procedure that has been used to previously explore WMC differences. Therefore, all participants undergo the same experiment procedure, but they are subsequently split into a number of subgroups on the basis of letter-crossing accuracy, speed of letter-crossing completion, slope of letter-crossing accuracy over time, and susceptibility to letter-crossing distractors. This technique allows the investigation of performance from the letter-crossing task to the executive functioning tasks, with the assumption being made that one group of participants will be more depleted by the letter-crossing task than others, and this will result in different depletion transfer effects.

Following the four experiments, post-hoc analyses were conducted to eliminate any doubt regarding the assumptions initially made in the within-subjects repeated measures analysis. That is, the assumption that all participants were depleted over the letter-crossing task, but some were more susceptible to depletion, and that without a comparison control group, it is unknown whether repetition effects concealed depletion transfer effects on the outcome tasks. By incorporating a control group to compare against the experimental group (i.e., those that completed the letter-crossing task), it can then be determined whether the absence of depletion transfer effects from pre-test to post-test on the majority of the outcome tasks were due to the methods employed or the lack of depletion effects in the data set.

6.2 Background to Study

In reviewing the self-regulation literature, a number of methodological problems were identified using the standard between-subject sequential-task paradigm. The sequential-task paradigm enlists one group of participants to complete an effortful self-regulation task, while the other participant group completes another version of a similar task that does not require self-regulation. Performance is then compared between groups on a second, unrelated self-regulation task. The replication crisis in the literature confirmed the suspicion that the sequential-task paradigm does not provide a reliable methodological design to recreate the previously published depletion effect (Francis et al., 2018).

The sequential-task procedure involves a between-subjects manipulation to examine what is arguably a within-subjects phenomenon, as detriments in performance should be assessed over time rather than between experimental groups. These between-subjects manipulations neglect the possibility that the control group depicts a greater performance on the outcome task because they are better at the task than the experimental group, instead of showing no depletion effects due to the absence of self-regulatory resources employed. Even though random assignment to experimental conditions is applied in all experiments to minimise cohort differences, it is still possible that some confounding factors exist between experimental and control groups that produce the group differences. Such confounds are more likely to be present in experiments with smaller numbers of participants. Notably, the strongest depletion effects have been reported in studies with small samples (Carter & McCullough, 2014; Carter et al., 2015). More extreme differences, positive or negative, were observed in the studies with smaller numbers of participants. Given that the cognitive changes are assumed to be operating within participants over time, a within-subject repeated measures design seems to be a more appropriate way to explore depletion effects as previously explored in the literature (Shamosh & Gray, 2007; Xu et al., 2014). By including a baseline measure, this allows for the comparison across sessions from pre- to post-test within individuals, instead of inferring depletion effects from comparing differences between groups.

A second issue arises out of the failure to measure performance on the depletion task and the reliance on the outcome task for evidence of depletion effects. There is never any consideration that the outcome task itself might be depleting, thereby showing detriments on the post-test. As all of the subtasks under the Stroop colour word task (Golden & Freshwater, 1978), including the subtasks that do not assess inhibition (i.e., word-reading and colour-naming), were found to correlate to letter-crossing accuracy in Study 2 in Chapter 5. This suggested that one executive function, such as inhibition, could not account for letter-crossing ability. Rather the results from Study 2 indicated that a common EF factor, related to all of the outcome tasks, underpinned the letter-crossing task. If self-regulatory resources are identical to, or a subset of, these common EF resources, it is possible then that outcome tasks in the sequential-task paradigm could also be depleting. If all participants within each experiment complete the same outcome task, this eliminates any comparison limitations between outcome tasks.

In order to address these issues, the following experiments explore changes in performance from pre- to post-test on cognitive performance where all participants have undergone the same depletion experience (i.e., the letter-crossing task). In doing so, the procedures widely used in exploring individual differences in WMC will be adopted. In the working memory literature, it is assumed that there are individual differences in WMC, which is operationally defined as a person's score on one or more complex span tasks, such as the OSPAN task (Conway et al., 2005). The sample is typically divided into quartiles from span measures, and the upper and lower quartile are then compared on some secondary task (Conway et al., 2005). Group differences on the secondary task are taken as evidence for the role of working memory in that task.

In adapting the individual differences approach to the current tasks, the assumption is made that there are individual differences in depletion. That is, some people will have less self-regulatory resources to draw on than others, or that they are more readily depleted than others. The strength model backs this assumption (Baumeister & Heatherton, 1996; Muraven & Baumeister, 2000; Muraven, Baumeister, & Tice, 1999). The intent is then to divide the participants into quartiles on the basis of the letter-crossing task and to then look for group differences on the outcome tasks. While this extreme-groups method is not free of criticism (Conway et al., 2005), it follows the same logic employed in the standard sequential-task analyses in that the key outcome is a difference between groups on the outcome task. An interaction between group and test session should be evident with group differences being larger at the post-test than at pre-test. Furthermore, Conway et al. (2005) and Lurquin (2017) both endorsed the use of extreme-group designs for establishing latent variables. It was therefore deemed an appropriate design for the investigation into the potential similarities between self-regulation and executive functioning abilities.

6.3 Study Aims

The following four experiments aim to assess whether self-regulatory depletion effects can transfer onto the follow-up executive function task. The following four experiments are the same as the first four experiments from Study 2, however, a different set of analyses have been conducted on the dataset. Thereby, Experiments 5 - 8 followed the exact methods and procedure of those experiments detailed in Study 2. To compare individual self-regulation functioning between executive functioning tasks, performance groups were created based on the four measures formed in Study 1. Study 2 originally found target accuracy, completion time, and distractor responses all correlated with some aspect of executive abilities and thus all may be measures of some aspect of depletion. While slope of accuracy over time was not found to be significantly correlated to executive functioning in the second study, it may remain an accurate method of scoring the letter-crossing task. It was therefore included in the current study in order to determine whether any effects could be identified under this scoring method.

For each letter-crossing measure, except distractor responses (e.g., *ie* or *ei*), quartile splits will be used to determine three groups: higher, moderate, and lower performance groups. Across the four experiments, the higher performance groups represent the top 25% of the sample, whereas the lower performance groups represent the bottom 25% of the sample and the moderate group represents the middle 50% of the sample. To be consistent with accuracy and slope measures, the completion time will be converted into a processing speed group. Under the speed scoring method, the higher performance (i.e., faster) group are the top 25% of the sample who have completed the letter-crossing task in the shortest amount of time. The lower performance (i.e., slower) groups are then the bottom 25% of the samples who were the slowest to complete the letter-crossing task. Under the distractor split however, as only a limited group of individuals responded to letter-crossing distractors, performance groups will be based on those who responded to distractors (i.e., Distractors) and those who did not respond to distractors (i.e., No-Distractors).

The introduction of a pre-test task in addition to the post-test may be problematic to the extent that performance on most cognitive tasks improves with increasing levels of practice (Morrison & Chein, 2011, Shipstead, Redick, & Engle, 2012). Depletion effects may then be difficult to detect if task repetition leads to superior post-test performance. Although, the presence, absence, or differential strength of repetition effects have the potential to provide clinically useful information in older adult populations (Darby, Maruff, Collie, & McStephen, 2002; Duff, Callister, Dennett, & Tometich, 2012). Darby et al. (2002) administered a short computerised cognitive test battery four times within three hours to a group of older patients with mild cognitive impairments and a group of healthy matched controls and found substantial improvement on the battery for the control group across the four sessions. In contrast, the clinical group showed severely attenuated learning across the four sessions (Darby et al., 2002). The authors concluded that the size of the repetition effect was a good marker of cognitive impairment (Darby et al., 2002).

In applying this logic to the depletion literature, it is possible that depletion effects might not necessarily be observed as a decrement in performance. Instead, depletion might be reflected in differential improvements from pre-test to post-test. Those who are less susceptible to depletion effects should show stronger repetition effects than those who are more susceptible to depletion effects. This repetition effect would be reflected in a group by test session interaction, with group differences having a greater magnitude at post-test than at pre-test (i.e., prior to depletion). It is expected that individuals depicting a higher performance will perform differently to lower performance individuals when depleted.

In sum, by splitting participants into groups on letter-crossing performance, equivalent group differences should also be observed on the outcome tasks, to the extent that letter-crossing and outcome tasks are correlated. As this was true for the overall accuracy measure, it is expected that group differences based on accuracy should result in group differences on all outcome measures. Likewise, since there was no correlation between slope and outcome measures, there should not be any group differences on the outcome measures. It is likely that some or all tasks will show improved performance on the post-test due to repetition effects. The functional marker for depletion transfer effects will be greater group differences at post-test than at pre-test.

6.4 Experiment 5: Stroop Colour Word Task (Golden & Freshwater, 1978)

The results of Experiment 1 indicated that overall accuracy on the letter-crossing task was significantly correlated with all three subtasks of the Stroop colour word task (Golden & Freshwater, 1978). Likewise, completion time was correlated with two of the three subtasks (word-reading and colour-word naming). On this basis, one would expect to see that the higher performing participants would show better performance on all three subtasks compared to the lower performing participants. In Experiment 1 from Study 2, slope of accuracy showed no relations with the Stroop subtasks, whereas distractor responses were correlated to only colour-naming performance. No such differences would then be expected for the slope and distractor measures. Performance on all three subtasks should improve with repetition. Depletion effects are then expected to be apparent only if group differences are magnified at post-test. A power analysis was conducted once again and determined this sample (N = 98) to be large enough to have confidence in the findings (α = .05, β = .01, Power = .99, with a critical F value of 3.09). Refer to Appendix D.

6.4.1. Methods.

As the same participant set was employed from Experiment 1 in Study 2, the methods did not change. The participants, materials, and procedures were described in Experiment 1.

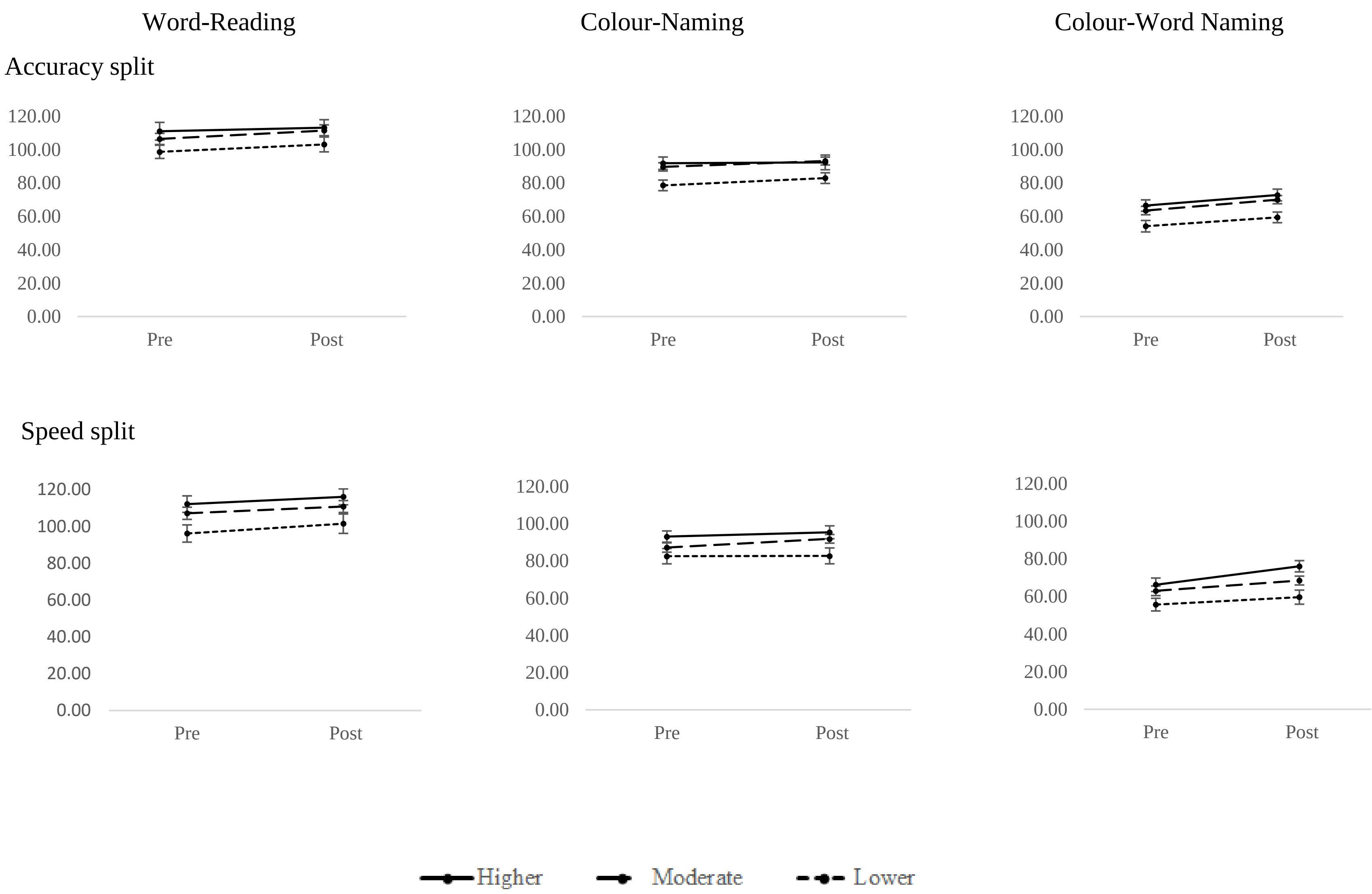
6.4.2. Results.

The data were analysed using 3 (Group: Higher, Moderate, and Lower performance) x 3 (Subtask: Word-Reading, Colour-Naming, and Colour-Word Naming) x 2 (Session: Pre-test, Post-test) mixed design ANOVAs, for accuracy, speed, and slope group splits. A 2 (Group: Distractors, No-Distractors) x 3 (Subtask: Word-Reading, Colour-Naming, and Colour-Word Naming) x 2 (Session: Pre-test, Post-test) mixed design ANOVA was conducted for the distractor group split. Where an effect violated Mauchley's test of sphericity, the Huynh-Feldt correction was used as ε > 0.75, and the degrees of freedom were adjusted. The outcomes of the study are summarised in Table 6.1 and Figure 6.1.

Table 6.1

Mean and Standard Errors for Letter-Crossing Accuracy, Speed of Completion, Slope of Accuracy Over Time, and Distractor Recalls Split By Letter-Crossing Performance Groups (Higher, Moderate, and Lower) Across Pre- and Post-Test Sessions in Word-Reading, Colour-Naming, and Colour-Word Naming Subtasks Under the Stroop Colour Word Task (Golden & Freshwater, 1978)

Letter-Crossing Task		Stroop Colour-Word Task					
		Word-Reading		Colour-Naming		Colour-Word Naming	
		Pre	Post	Pre	Post	Pre	Post
		M (*SE*)	*M* (*SE*)	*M* (*SE*)	*M* (*SE*)	*M* (*SE*)	*M* (*SE*)
Accuracy	Higher (*N* = 24)	111.04 (5.32)	113.17 (4.79)	91.88 (3.72)	92.38 (4.41)	66.5 (3.42)	72.83 (3.52)
	Moderate (*N* = 50)	106.44 (3.34)	111.44 (3.43)	89.68 (2.49)	93.2 (2.38)	63.52 (2.52)	70.02 (2.44)
	Lower (*N* = 24)	98.71 (3.96)	103.13 (4.44)	78.58 (3.18)	82.96 (3.21)	54.13 (3.45)	59.42 (3.19)
Speed	Faster (*N* = 24)	112.13 (4.49)	116.08 (4.32)	93.08 (3.03)	95.38 (3.46)	66.17 (3.59)	76 (3)
	Moderate (*N* = 50)	107.14 (3.31)	110.82 (3.22)	87.2 (2.5)	91.88 (2.33)	62.96 (2.55)	68.42 (2.35)
	Slower (*N* = 24)	96.17 (4.69)	101.5 (5.27)	82.54 (4.1)	82.71 (4.26)	55.63 (3.36)	59.58 (3.75)
Slope	Minor (*N* = 24)	105.29 (4.99)	113.21 (5.13)	85.75 (4)	88.13 (4.3)	60.42 (4.03)	67.42 (3.5)
	Moderate (*N* = 50)	105.9 (2.95)	109 (2.83)	90.22 (2.63)	92.86 (2.36)	64.7 (2.56)	70.88 (2.56)
	Major (*N* = 24)	105.58 (5.76)	108.17 (5.9)	82.83 (2.83)	108.17 (5.9)	57.75 (2.94)	63.04 (3.2)
Distractors	No-Distractors (*N* = 62)	105.42 (3.25)	108.66 (3.27)	87.26 (2.23)	89.56 (2.49)	63.21 (2.18)	28.97 (2.24)
	Distractors (*N* = 36)	106.11 (3.26)	111.83 (3.22)	87.92 (3.11)	92.08 (2.6)	59.78 (3.15)	66.64 (2.87)


Word-Reading
Colour-Naming
Colour-Word Naming
Accuracy split
Speed split
120.00
100.00
80.00
60.00
40.00
20.00
0.00
Pre
Post
Higher
Moderate
Lower

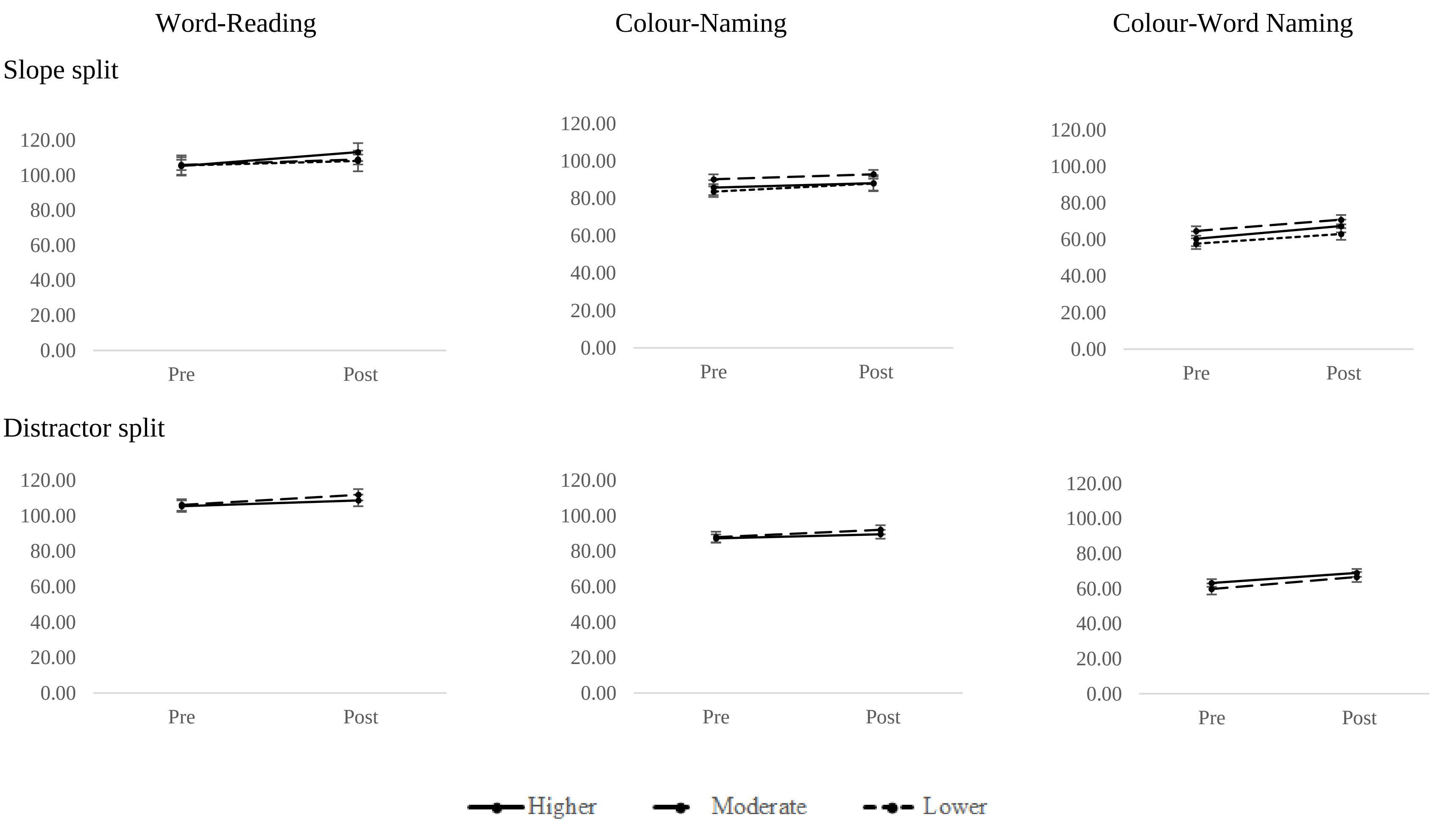


Figure 6.1. Mean accuracy (y axis) scores across pre- and post-test sessions (x axis) for Stroop colour word (Golden & Freshwater, 1978) sub-tasks: Word-Reading, Colour-Naming, and Colour-Word Naming. Error bars were created from standard error values. Whereby Higher perfor-

-mance refers to Higher Accuracy, Faster Speed, Minor Slope, and No-Distractors; Moderate performance refers to Moderate Accuracy, Moderate Speed, Moderate Slope; and Lower performance refers to Lower Accuracy, Slower Speed, Major Slope, and Distractors.

6.4.2.1. Accuracy split. There was a significant difference between the accuracy of groups, F (2, 95) = 3.62, p =.031, MSE = 1661.46, *partial* η^2 = .07, where the higher accuracy group reported greater accuracy compared to moderate and lower accuracy groups respectively. A main effect for session was found, F (1, 95) = 36.75, p <.001, MSE = 63.59, *partial* η^2 = .28, where performance improved from pre- to post-test. There was no interaction between accuracy groups and session, F (2, 95) = .81, p = .45, MSE = 63.59, *partial* η^2 = .02. There was a main effect for subtasks, F (1.68, 159.15) = 337.94, p <.001, MSE = 284.41, *partial* η^2 = .78. Performance on the word-reading subtask was better than on the colour-naming subtask, and both were better than the colour-word naming subtask. There was no significant interaction found between session and subtask, F (2, 190) = 2.52, p = .082, MSE = 46.66, *partial* η^2 =.04. No other interactions were significant.

6.4.2.2. Speed split. There was a significant difference between speed groups with the faster group reporting greater task accuracy than moderate and slower speed groups respectively, F (2, 95) = 4.08, p =.020, MSE = 1646.53, *partial* η^2 = .08. There was also a main effect for session, F (1, 95) = 39.23, p <.001, MSE = 63.69, *partial* η^2 = .29, which showed an increase in performance from pre- to post-test. There was no interaction between speed groups and session, F (2, 95) = .72, p = .487, MSE = 63.69, *partial* η^2 =.02. There was a main effect for subtask, F (2, 190) = 332.32, p <.001, MSE = 284.02, *partial* η^2 = .78. Performance was greatest under the word-reading subtask, followed by colour-naming and colour-word naming subtasks. There was a significant interaction between session and subtask, F (2, 190) = 3.88, p = .022, MSE = 45.77, *partial* η^2 = .04. Planned contrasts revealed that there was a significant quadratic interaction as accuracy increased across all three subtasks from pre-test to post-test, F (1, 96) = 4.36, p = .04. No other interactions were significant.

6.4.2.3. Slope split. There was no significant difference between slope groups, F (2, 95) = .60, p = .55, MSE = 1765.74, *partial* η^2 = .01. There was a main effect for session, F (1, 95) = 43.39, p <.001, MSE = 63.78, *partial* η^2 = .31, where performance increase from pre- to post-test. There was no interaction between session and slope groups, F (2, 95) = .66, p = .52, MSE = 63.78, *partial* η^2 = .01. There was a main effect for subtask, F (2, 190) = 360.57, p <.001, MSE = 274.84, *partial* η^2 = .79. Performance was greatest under the word-reading subtask, followed by colour-naming and colour-word naming subtasks. No other interactions were significant.

6.4.2.4. Distractor split. There was no significant difference between the distractor recall groups, *F* (1, 96) < .01, *p* = .953, *MSE* = 1769.22, *partial* η^2 = .00, but there was a main effect for session, *F* (1, 96) = 47.57, *p* <.001, *MSE* = 62.82, *partial* η^2 = .33, where performance improved from pre- to post-test. There was no significant interaction between session and distractor groups, *F* (1, 96) = 1.79, *p* = .184, *MSE* = 62.82, *partial* η^2 = .02. There was a main effect for subtask, *F* (2, 192) = 369.06, *p* <.001, *MSE* = 283.34, *partial* η^2 = .79. Significant decreases in accurate performance were found from the word-reading subtask to the colour-naming and colour-word naming subtasks. No other interactions were significant.

6.4.3. Preliminary discussion.

As expected, there were overall differences between the accuracy and speed groups on the Stroop task, but not between slope and distractor groups. The differences between the subtasks also emerged as expected with the colour-word subtask incorporating inhibition ability, therefore being substantially harder than the word-reading and colour-naming subtasks. Repetition effects were present in that performance increased from pre-test to post-test. The lack of any session by group interactions leads to the conclusion that depletion transfer effects did not occur on the post-test performance for any Stroop subtasks.

6.5. Experiment 6: OSPAN Task (Turner & Engle, 1989)

The outcomes of Experiment 2 indicated that both components of the OSPAN task were correlated with letter-crossing accuracy and completion time measures, whereas slope and distractor measures were not. Consequently, there should be group differences on accuracy and time measure splits. In addition, pre-test scores for the processing component (i.e., the maths component) indicated that on average participants correctly identified 84% of the maths problems (with 30% making zero or one errors) but only recalled 62% of the words under the memory component. Consequently, there is more room for repetition effects in the memory component than the processing component. Again, larger group differences at post-test would be a functional marker of depletion. Experiment 6 reported strong power, *N* = 145, α = .05, β = .01, Power = .99, with a critical *F* value of 3.05.

6.5.1. Methods.

The data set was the same as that employed in Experiment 2, Study 2. The participants, materials and procedures were described fully in Experiment 2.

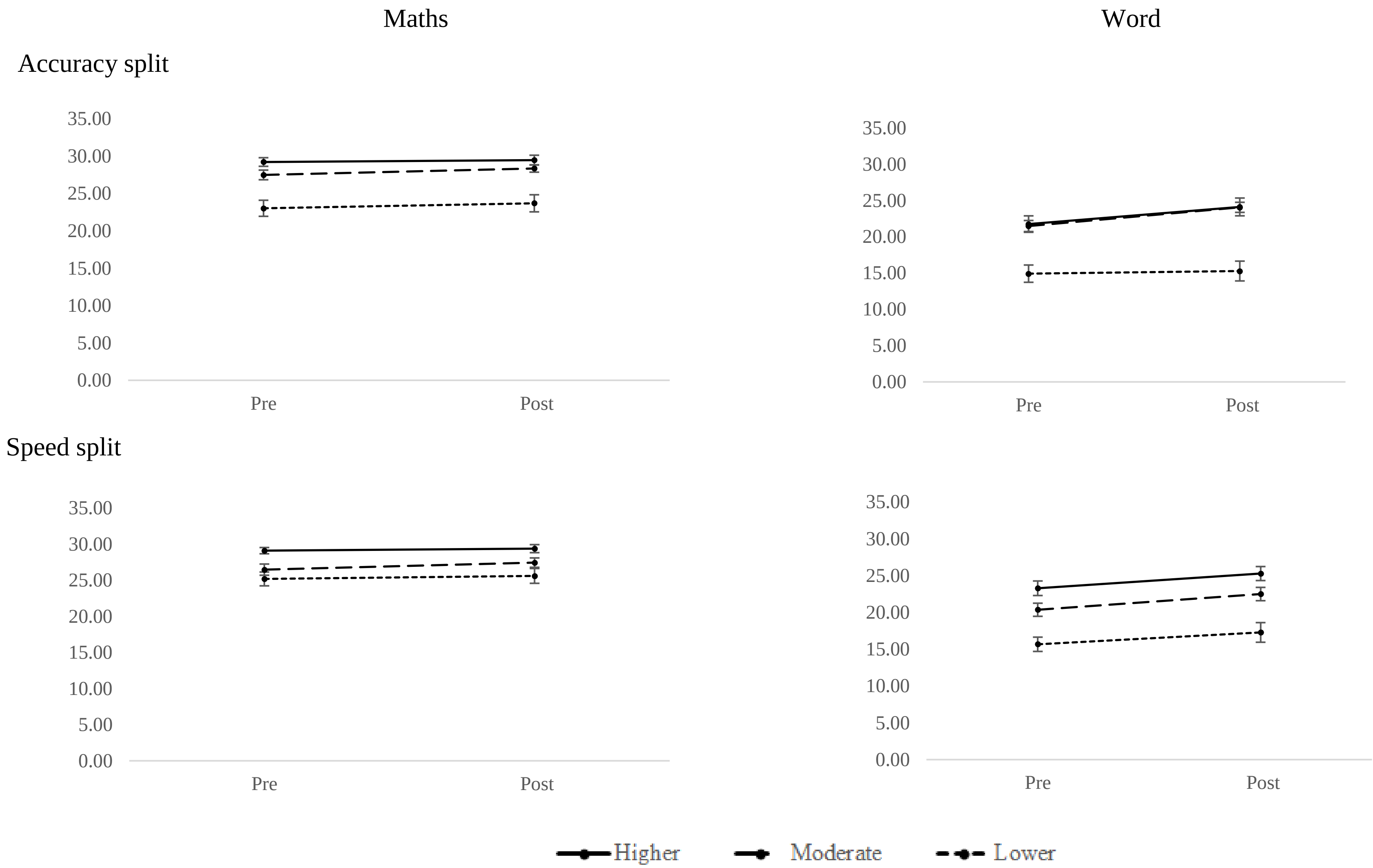
6.5.2. Results.

The analyses conducted were 3 (Group: Higher, Moderate, and Lower performance) x 2 (Component: Word, Maths) x 2 (Session: Pre-test, Post-test) mixed design ANOVAs for letter-crossing target accuracy, time taken to complete the task, and slope of accuracy over the letter-crossing task. A 2 (Group: Distractors, No-Distractors) x 2 (Component: Words, Maths) x 2 (Session: Pre-Test, Post-Test) mixed design ANOVA was conducted for the letter-crossing distractors group split. Once again, with the exception of the slope split, the higher performance group (i.e., higher accuracy, faster speed, and no distractors) generally reported higher average accuracies for both OSPAN components, compared to moderate and lower performance groups. A general increase was found in target averages from pre- to post-test. The outcomes of the experiment are summarised in Table 6.2 and Figure 6.2.

Table 6.2

Mean and Standard Errors for Letter-Crossing Accuracy, Speed of Completion in Minutes, Slope of Accuracy Over Time, and Distractor Recalls Split By Letter-Crossing Performance Groups (Higher, Moderate, and Lower) Across Pre- and Post-Test Sessions in Maths and Word Components Under the OSPAN Task (Turner & Engle, 1989)

Letter-Crossing Task		OSPAN Task			
		Maths		Word	
		Pre	Post	Pre	Post
		M (*SE*)	*M* (*SE*)	*M* (*SE*)	*M* (*SE*)
Accuracy	Higher (*N* = 36)	29.22 (.58)	29.47 (.66)	21.75 (1.13)	24.11 (1.23)
	Moderate (*N* = 73)	27.49 (.65)	28.36 (.49)	21.48 (.76)	24.05 (.7)
	Lower (*N* = 36)	23.03 (1.08)	23.69 (1.14)	14.92 (1.19)	15.28 (1.37)
Speed	Faster (*N* = 36)	29.11 (.43)	29.39 (.56)	23.28 (.99)	25.28 (.95)
	Moderate (*N* = 73)	26.48 (.77)	20.36 (.89)	27.45 (.65)	22.49 (.9)
	Slower (*N* = 36)	25.19 (.96)	25.61 (1.02)	15.67 (.97)	17.28 (1.33)
Slope	Minor (*N* = 36)	25.83 (1.2)	27.03 (.98)	19.11 (1.01)	21.69 (1.08)
	Moderate (*N* = 73)	27.58 (.59)	28.3 (.55)	20.16 (.91)	22.15 (1.02)
	Major (*N* = 36)	26.25 (.94)	26.25 (1)	20.22 (1.25)	21.56 (1.25)
Distractors	No-Distractors (*N* = 101)	27.58 (.53)	28.41 (.46)	20.62 (.7)	23.07 (.72)
	Distractors (*N* = 44)	25.05 (.97)	25.34 (.97)	18.3 (1.17)	19.18 (1.29)


Maths
Word
Accuracy split
Speed split
35.00
30.00
25.00
20.00
15.00
10.00
5.00
0.00
Pre
Post
Higher
Moderate
Lower

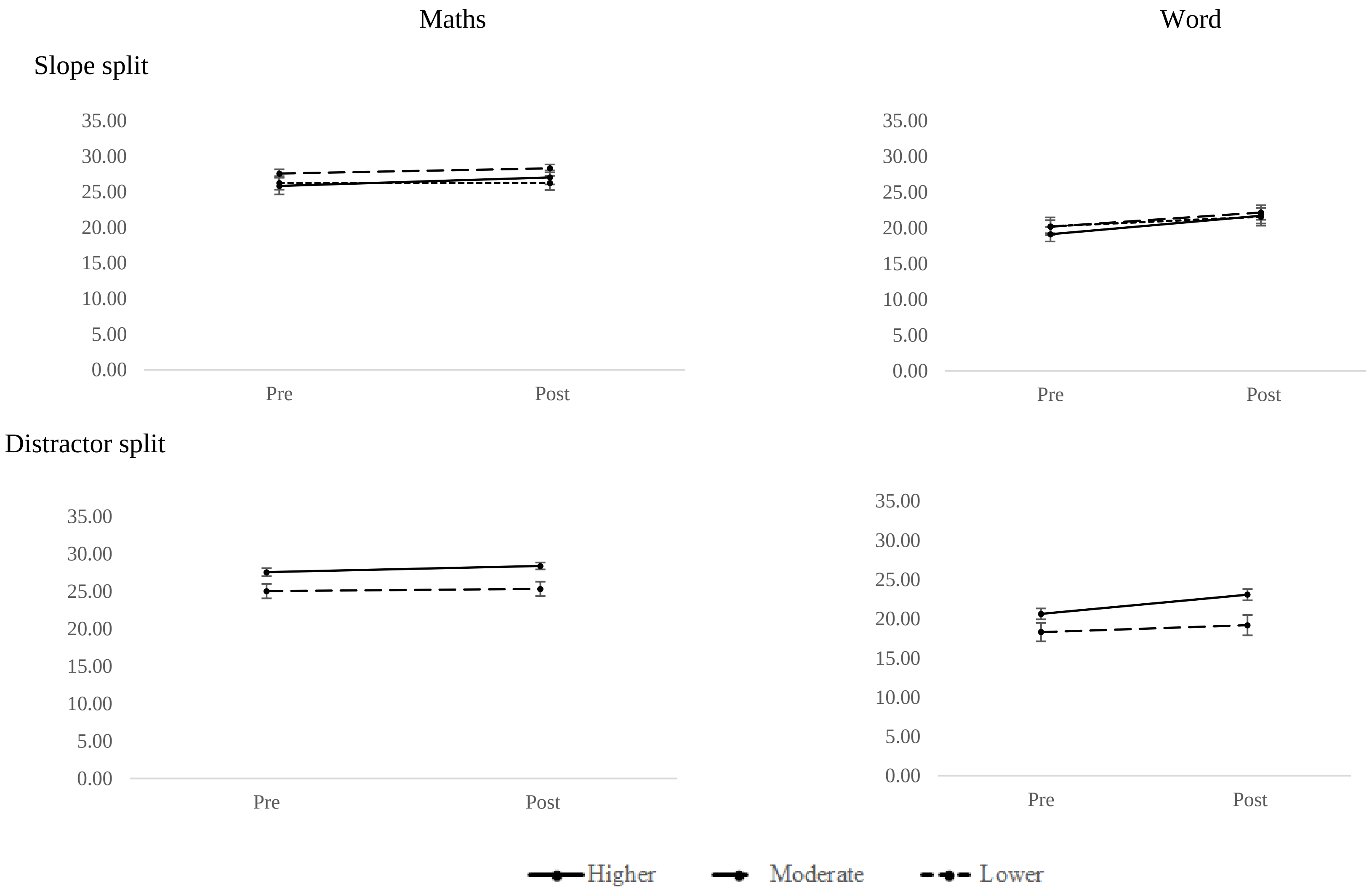


Figure 6.2. Mean accuracy (y axis) scores across pre- and post-test sessions (x axis) for OSPAN (Turner & Engle, 1989) components: Maths and Word. Error bars were created from standard error values. Whereby Higher performance refers to Higher Accuracy, Faster Speed,

Minor Slope, and No-Distractors; Moderate performance refers to Moderate Accuracy, Moderate Speed, Moderate Slope; and Lower performance refers to Lower Accuracy, Slower Speed, Major Slope, and Distractors.

6.5.2.1. Accuracy split. There was a significant difference found between accuracy groups, F (2, 142) = 27.2, p <.001, MSE = 82.09, *partial* η^2 = .28, where the higher accuracy group reported the greatest accuracy on the OSPAN task, followed by moderate and lower accuracy groups respectively. A main effect for session was found, F (1, 142) = 25.83, p <.001, MSE = 6.99, *partial* η^2 = .15. Significant improvements were observed from pre- to post-test. There was no interaction between session and accuracy groups, F (2, 142) = 2.5, p = .085, MSE = 6.99, *partial* η^2 = .03. A main effect for OSPAN component was found, F (1, 142) = 112.96, p <.001, MSE = 50.31, *partial* η^2 = .44. Higher average accuracy was reported under the maths component compared to the word component. An interaction between session and OSPAN component was found, F (1, 142) = 6.45, p = .012, MSE = 6.92, *partial* η^2 = .04. Planned contrasts revealed a significant linear interaction, F (1, 142) = 6.45, p = .012, where word component performance notably improved, with less change in the maths component. No other interactions were significant.

6.5.2.2. Speed split. A significant difference was found between speed groups, F (2, 142) = 12.79, p <.001, MSE = 96.2, *partial* η^2 = .15, where the fastest speed group generally reported the greatest OSPAN accuracy, compared to moderate and slower speed groups respectively. Save for the pre-test word component where the moderate speed group reported the greatest accuracy, followed by faster and slower speed groups respectively. There was a significant effect for session, F (1, 142) = 27.62, p <.001, MSE = 7.19, *partial* η^2 = .16, where performance increased from pre- to post-test. A main effect for component was found, F (1, 142) = 110.36, p <.001, MSE = 49.46, *partial* η^2 = .44. Higher averages were once again reported under the maths component compared to the word component. There was no interaction between session and speed groups, F (2, 142) = .6, p = .55, MSE = 4.29, *partial* η^2 =.01.There was an interaction between session and component, F (1, 142) = 8.43, p = .004, MSE = 7.14, *partial* η^2 = .06. Planned contrasts revealed there was a significant linear interaction, F (1, 142) = 8.43, p = .004, performance improved for word recalls from pre- to post-test, but there was only a minor improvement under the maths component. No other significant interactions were found.

6.5.2.3. Slope split. There was no significant difference between slope groups, F (2, 142) = .73, p = .49, MSE = 112.38, *partial* η^2 = .01. However, there was a main effect of session, *F* (1, 142) = 31.32, *p* <.001, *MSE* = 7.06, *partial* η^2 = .18, where performance increased from pre- to post-test. There was no interaction between session and slope groups, *F* (2, 142) = 1.92, *p* = .15, *MSE* = 7.06, *partial* η^2 = .03. A main effect of component was found, *F* (1, 142) = 92.38, *p* <.001, *MSE* = 51.6, *partial* η^2 = .39. Higher averages were reported under the maths component compared to the word component. An interaction between session and component was found, *F* (1, 142) = 8.01, *p* = .005, *MSE* = 7.15, *partial* η^2 = .05. A planned contrast revealed a significant linear interaction, *F* (1, 142) = 8.01, *p* = .005, where word recalls increased from pre- to post-test with less change in the maths component. No other significant interactions were found.

6.5.2.4. Distractor split. There was a significant difference found between distractor groups, *F* (1, 143) = 10.17, *p* = .002, *MSE* = 105.25, *partial* η^2 = .07, where the No-Distractors group consistently reported greater accuracy compared to the Distractors group. A main effect for session was found, *F* (1, 143) = 21.79, *p* <.001, *MSE* = 6.97, *partial* η^2 = .13, where performance increased from pre- to post-test. There was an interaction between sessions and distractor groups, *F* (1, 143) = 4.79, *p* = .03, *MSE* = 6.96, *partial* η^2 =.03. The No-Distractors group reported greater improvement across sessions compared to the Distractors group. A main effect for component was found, *F* (1, 143) = 94.39, *p* <.001, *MSE* = .51.58, *partial* η^2 = .40. The maths component reported higher averages compared to the word component. An interaction between session and component was found, *F* (1, 143) = 5.34, *p* = .022, *MSE* = 7.04, *partial* η^2 = .01. Planned contrasts revealed there was a significant linear interaction, *F* (1, 143) = 5.34, *p* = .022, where word recalls increased from pre- to post-test, with less change in the maths component. No other significant interactions were found.

6.5.3. Preliminary discussion.

Generally, there was little improvement from pre- to post-test scores on the math processing component of the task. In contrast, there was a substantial improvement with the memory component. The lack of improvement in math scores may simply reflect ceiling effects on the pre-test scores. As expected there were overall group differences for the accuracy and speed measures, and unexpectedly for the distractor measure. As was the case in Experiment 5, there were no significant group by session interactions for accuracy, time and slope measures suggesting a lack of depletion transfer effects. There was a significant group by session interaction with the distractor measure. The post-test differences between the distractor groups were magnified compared to the pre-test differences, and this is most apparent in the memory component (i.e., word component). While the group who made no distractor responses on the letter-crossing task improved on both math and word components, the group who did make errors did not improve in either component. This lack of repetition effect for the Distractors group could be interpreted as depletion of self-regulation resources, counter-balancing any repetition effects that may have be present.

6.6. Experiment 7: Immediate Serial Recall Task

The outcomes of Experiment 3 indicated that there were no correlations between letter-crossing measures and ISR task performance save for 4-syllable items and letter-crossing accuracy. As such, there was no strong expectation of group differences on any of the measures in the current experiment. Experiment 3 showed a significant word length effect in that 2-syllable words were better recalled than 3-syllable words, which were better recalled than 4-syllable words. Repetition effects are expected to be observed, with an increase in performance expected between pre- and post-test. If the results of the previous experiments are to be replicated, there should be no group by session interactions with any of the measures, save for the distractor measure. Experiment 7 reported strong power, $N = 109$, $\alpha = .05$, $\beta = .01$, Power = .99, with a critical F value of 3.08.

6.6.1. Methods.

The methods are the same as those employed in Experiment 3, Study 2. The participants, materials, and procedure were fully described in Experiment 3.

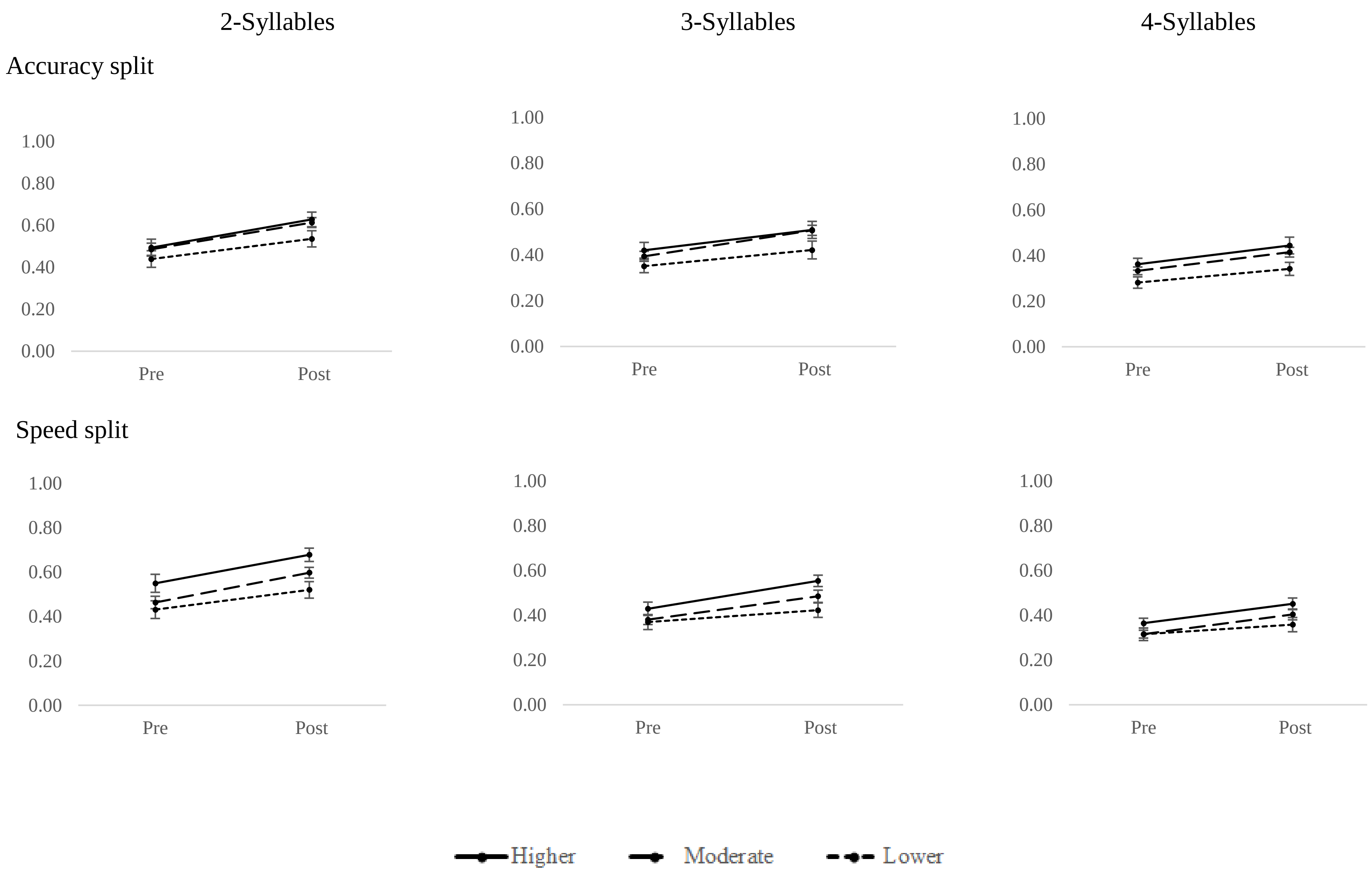
6.6.2. Results.

The analyses conducted were 3 (Group: Higher, Moderate, and Lower performance) x 3 (Syllables: 2-Syllable, 3-Syllable, and 4-Syllable items) x 2 (Session: Pre-test, Post-test) mixed design ANOVAs for letter-crossing accuracy, speed, and slope group splits. A 2 (Group: Distractors, No-Distractors) x 3 (Syllables: 2-Syllable, 3-Syllable, and 4-Syllable items) x 2 (Session: Pre-test, Post-test) mixed design ANOVA was conducted for the distractor group split. Word length effects were found with 2-syllable words recalled more often than 3-syllable words, which were better recalled than 4-syllable words. Post-test recall was also superior to pre-test performance. Group differences appeared to be weaker in this experiment. Where a variable violated Mauchley's test of sphericity, the Huynh-Feldt correction was used, as $\varepsilon > 0.75$. The outcomes of the study are summarised in Table 6.3 and Figure 6.3.

Table 6.3

Mean and Standard Errors For Letter-Crossing Accuracy, Speed of Completion in Minutes, Slope of Accuracy Over Time, and Distractor Recalls Split By Letter-Crossing Groups (Higher, Moderate, and Lower) Across Pre- and Post-Test Sessions in 2-, 3-, and 4-Syllable Items Under the ISR Task

Letter-Crossing Task		ISR Task					
		2-Syllables		3-Syllables		4-Syllables	
		Pre	Post	Pre	Post	Pre	Post
		M (*SE*)	*M* (*SE*)	*M* (*SE*)	*M* (*SE*)	*M* (*SE*)	*M* (*SE*)
Accuracy	Higher (*N* = 27)	.49 (.04)	.63 (.03)	.42 (.03)	.51 (.04)	.36 (.03)	.44 (.04)
	Moderate (*N* = 55)	.49 (.03)	.61 (.02)	.39 (.02)	.51 (.02)	.33 (.02)	.42 (.02)
	Lower (*N* = 27)	.44 (.04)	.54 (.04)	.35 (.03)	.42 (.04)	.28 (.02)	.34 (.03)
Speed	Faster (*N* = 27)	.55 (.04)	.68 (.03)	.43 (.03)	.55 (.03)	.36 (.02)	.45 (.03)
	Moderate (*N* = 55)	.46 (.03)	.60 (.02)	.38 (.02)	.48 (.03)	.32 (.02)	.40 (.02)
	Slower (*N* = 27)	.43 (.04)	.52 (.04)	.37 (.03)	.42 (.03)	.31 (.03)	.36 (.03)
Slope	Minor (*N* = 27)	.47 (.04)	.56 (.04)	.37 (.04)	.48 (.04)	.34 (.03)	.38 (.04)
	Moderate (*N* = 55)	.50 (.03)	.62 (.02)	.41 (.02)	.49 (.03)	.32 (.01)	.42 (.02)
	Major (*N* = 27)	.44 (.04)	.60 (.04)	.38 (.03)	.49 (.03)	.33 (.02)	.40 (.03)
Distractors	No-Distractors (*N* = 74)	.48 (.02)	.63 (.02)	.40 (.02)	.52 (.02)	.33 (.01)	.43 (.02)
	Distractors (*N* = 35)	.47 (.04)	.54 (.03)	.37 (.03)	.42 (.03)	.32 (.03)	.35 (.03)


2-Syllables
3-Syllables
4-Syllables
Accuracy split
Speed split
1.00
0.80
0.60
0.40
0.20
0.00
Pre
Post
Higher
Moderate
Lower

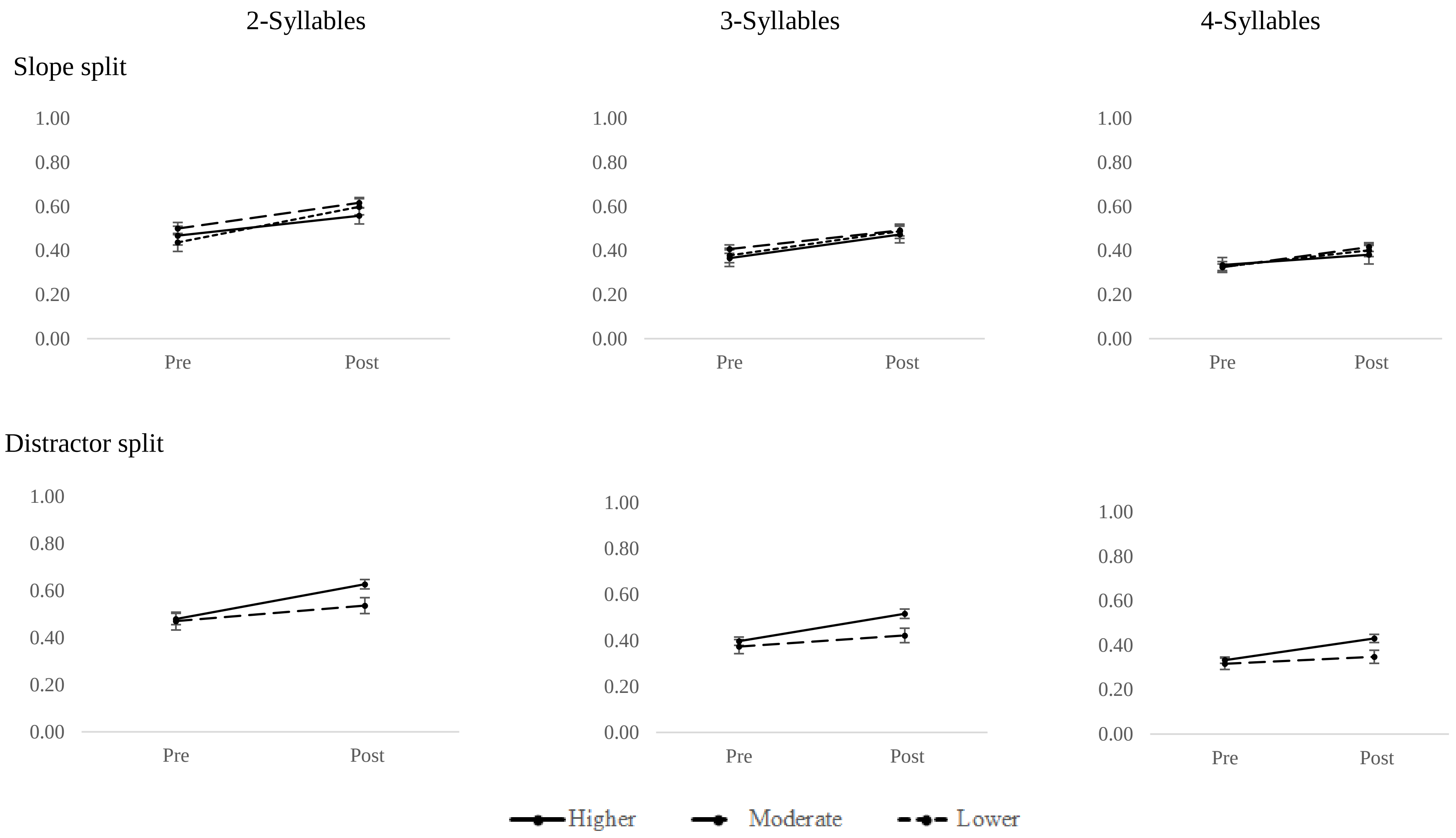


Figure 6.3. Mean accuracy (y axis) scores across pre- and post-test sessions (x axis) for ISR syllable lengths: 2-syllables, 3-syllables, and 4-syllables. Error bars were created from standard error values. Whereby Higher performance refers to Higher Accuracy, Faster Speed, Minor

Slope, and No-Distractors; Moderate performance refers to Moderate Accuracy, Moderate Speed, Moderate Slope; and Lower performance refers to Lower Accuracy, Slower Speed, Major Slope, and Distractors.

6.6.2.1. Accuracy split. There was no significant difference between groups, *F* (2, 106) = 2.44, *p* = .092, *MSE* = .13, *partial* η^2 = .04. A main effect was found for session, *F* (1, 106) = 55.32, *p* <.001, *MSE* = .02, *partial* η^2 = .34. Significant improvements were seen from pre- to post-test. There was no interaction between session and accuracy group, *F* (2, 106) = 55.32, *p* = .55, *MSE* = .02, *partial* η^2 = .01. A main effect for syllable-length was identified, *F* (2, 212) = 171.38, *p* <.001, *MSE* = .01, *partial* η^2 = .62. Where higher averages were reported under 2-syllable items, followed by 3-syllable and 4-syllable items. The session and syllable interaction was significant, *F* (1.92, 203.76) = 3.29, *p* = .041, *MSE* = .01, *partial* η^2 = .03. Planned contrasts revealed a significant interaction, *F* (1, 106) = 5.54, *p* = .020, where a significant increase was observed for all syllable lengths from pre-test to post-test. No other interactions were found.

6.6.2.2. Speed split. There was a significant difference found between groups, *F* (2, 106) = 3.49, *p* = .034, *MSE* = .12, *partial* η^2 = .06, where the faster speed group consistently reported greater accuracy compared to the moderate and slower speed groups respectively. A main effect for session was found, *F* (1, 106) = 55.78, *p* <.001, *MSE* = .02, *partial* η^2 = .35. Significant improvements in target recalls were seen from pre- to post-test. There was no interaction between session and speed groups, *F* (2, 106) = 1.55, *p* = .218, *MSE* = .02, *partial* η^2 = .03. A main effect for syllable was found, *F* (2, 212) = 180.63, *p* <.001, *MSE* = .01, *partial* η^2 = .63. Once again, the highest averages were reported under the 2-syllable items. This was followed by 3-syllable and 4-syllable item averages. The session and syllable interaction was significant, *F* (1.92, 203.77) = 3.19, *p* = .045, *MSE* = .01, *partial* η^2 = .03. Planned contrasts revealed a significant linear interaction, *F* (1, 106) = 5.50, *p* = .021, where improvements were seen across all syllable lengths from pre- to post-test. No other interactions were found.

6.6.2.3. Slope split. There was no significant difference found between groups, *F* (2, 106) = .42, *p* = .66, *MSE* = .13, *partial* η^2 =.01. A main effect for session was found, *F* (1, 106) = 59.16, *p* <.001, *MSE* = .02, *partial* η^2 = .36. Significant improvements were found from pre- to post-test. There was no interaction between session and slope groups, *F* (2, 106) = .5, *p* = .609, *MSE* = .02, *partial* η^2 = .01. Significant differences between syllables lengths were found, *F* (2, 212) = 165.07, *p* <.001, *MSE* = .01, *partial* η^2 = .61. The highest average proportions were recalled for 2-syllable items. This was followed by 3-syllable items and 4-

syllable items. The session and syllable interaction was significant, F (1.91, 202.75) = 4.40, p = .015, MSE = .01, *partial* η^2 = .04. Planned contrasts revealed a significant linear interaction, F (1, 106) = 7.53, p = .007, an increase in recall rates were observed across all syllable lengths from pre-test to post-test. No other interactions were observed.

6.6.2.4. Distractor split. There were no significant differences identified between distractor groups, F (1, 107) = 3.13, p = .08, MSE = .13, *partial* η^2 = .03. There was a main effect for session, F (1, 107) = 46.69, p <.001, MSE = 1.03, *partial* η^2 = .3. A significant improvement in target recalls was seen from pre- to post-test. An interaction between session and distractor group occurred, F (1, 107) = 8.60, p = .004, MSE = .02, *partial* η^2 = .07. Planned contrasts revealed a significant linear interaction, F (1, 107) = 8.60, p = .004, where both groups improved from pre-test to post-test. There was a main effect for items, F (2, 214) = 171.14, p <.001, MSE = .01, *partial* η^2 = .62. The highest accuracy proportion averages were consistently reported under 2-syllable items, followed by 3-syllable and 4-syllable items. There were no other interactions found.

6.6.3. Preliminary discussion.

The expected word length and repetition effects emerged in the data. As was the case in the previous experiments, there were no group by session interactions when groups were split on the basis of overall accuracy, speed, or slope values. With these measures, there is little evidence for depletion transfer effects. In contrast, post-test differences between distractor groups are much larger than pre-test differences. Thereby, the current experiment replicated a similar finding in Experiment 6, where individuals susceptible to distractor responses show limited repetition effects on the post-test, which suggests a functional marker of a depletion transfer effect was present.

6.7. Experiment 8: Immediate Serial Recall Task with Proactive Interference (Tolan & Tehan, 2002)

In Experiment 4, the relationship between the letter-crossing and target recall in a proactive interference task was examined. Overall, immediate serial recall performance and recall of a specific target item (i.e., Block-1 or Block-2 targets) in each list were highly correlated. In this experiment, the analyses have been restricted to target specific recall, as the Block-2 target measure details interference of the proactive to-be-forgotten foil, compared to Block-1 targets with no proactive interference conditions. The outcomes of Experiment 4 indicated that letter-crossing accuracy was related to PI-ISR target recall and thus it is expected that accuracy group differences should be present in the current experiment. As in the prior experiments, a significant group by session interaction serves as a functional marker of the strength model.

On the post-test measure, a manipulation was added to enhance the likelihood that the foil would be recalled instead of the target item on the Block-2 trials, as detailed in the methods section. If this manipulation is successful, then any benefits of repetition that have been observed in earlier experiments may be reduced under block-2 target recall and there should be a marked increase in foil intrusions. Experiment 8 reported strong power, $N = 114$, $\alpha = .05$, $\beta = .01$, Power = .99, with a critical F value of 3.07.

6.7.1. Method.

The participants, materials, and procedure were described in Experiment 4. There was one change to the post-test PI-ISR task that was made to increase the number of block-1 foils that were produced. Tehan and Humphreys (1998) demonstrated that if rhymes of the foil were included as filler items in the second block, target recall decreased and recall of the block-1 foil increased. Consequently, in all the post-test two-block trials, the first and fourth filler items in the second block were replaced by rhymes of the foil. Thus an example of a trial would be, *bed tea white heart ! kite cat black site 56 ?????*, where *kite* and *site* were priming the foil *white.*

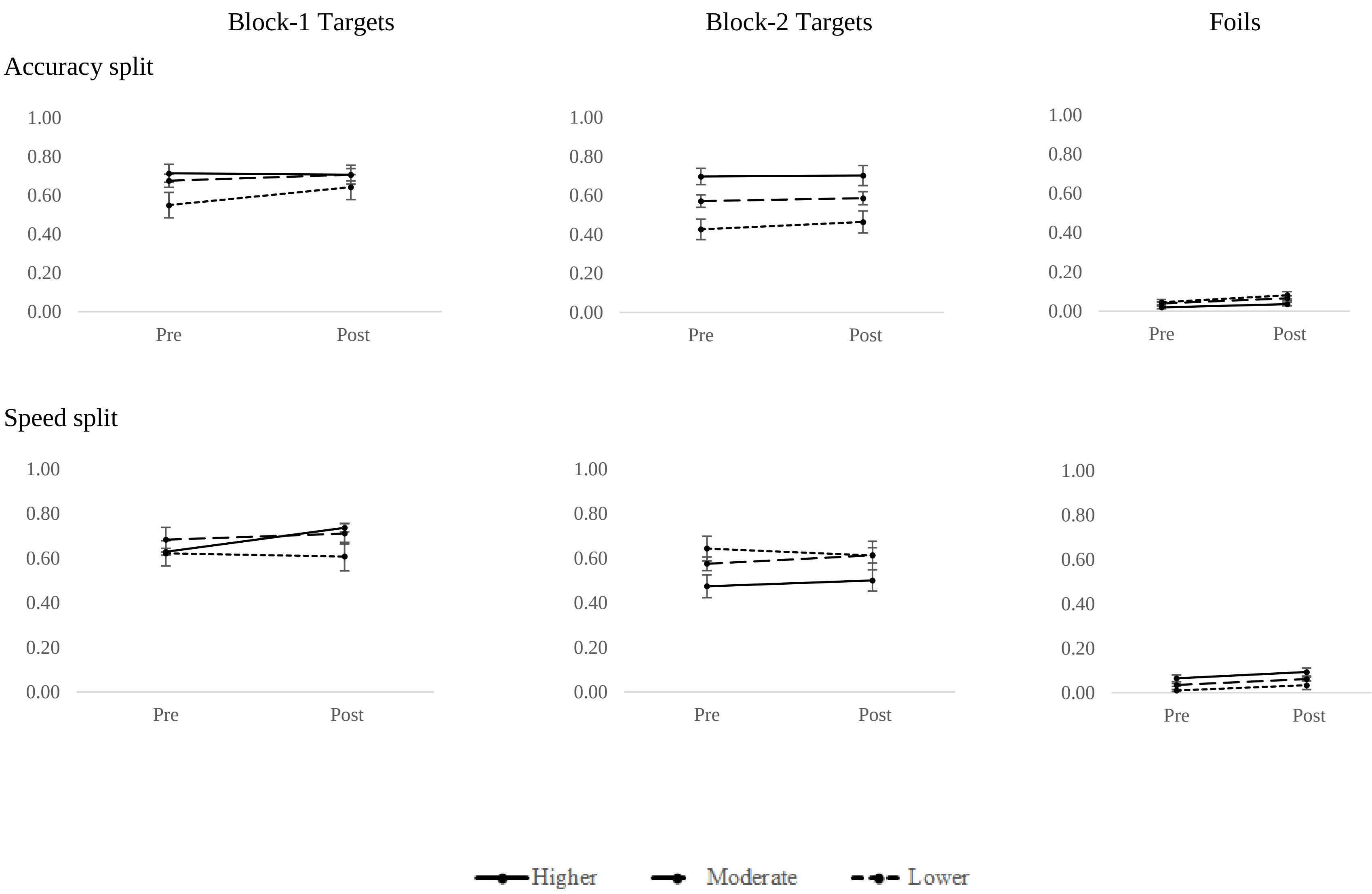
6.7.2. Results.

The analyses conducted were a 3 x (Group: Higher, Moderate, and Lower performance) x 3 (Items: Block-1 Targets, Block-2 Targets, and Foil items) x 2 (Session: Pre-test, Post-test) mixed design ANOVAs for letter-crossing accuracy, speed, and slope group splits. A 2 (Group: Distractors, No-Distractors) x 3 (Items: Block-1 Targets, Block-2 Targets, and Foil items) x 2 (Session: Pre-test, Post-test) mixed design ANOVA was conducted for the distractor group split. Generally, performance on all measures were higher on the post-test than pre-test. This means that target recall improved, but there was also an increase in foil intrusions. Where any variable violated Mauchley's test of sphericity, the Huynh-Feldt correction was used, as $\varepsilon > 0.75$. The results of the experiment are summarised in Table 6.4 and Figure 6.4.

Table 6.4

Mean and Standard Errors for Letter-Crossing Accuracy, Speed of Completion in Minutes, Slope of Accuracy Over Time, and Distractor Recalls Split By Letter-Crossing Performance Groups (Higher, Moderate, and Lower) Across Pre- and Post-Test Sessions in Block-1 Targets, Block-2 Targets, and Foil items under the PI-ISR Task (Tehan & Tolan, 2002)

Letter-Crossing Task		PI-ISR Task					
		Block-1 Targets		Block-2 Targets		Foils	
		Pre	Post	Pre	Post	Pre	Post
		M (*SE*)	*M* (*SE*)	*M* (*SE*)	*M* (*SE*)	*M* (*SE*)	*M* (*SE*)
Accuracy	Higher (*N* = 28)	.71 (.05)	.71 (.05)	.70 (.04)	.70 (.05)	.02 (.01)	.04 (.01)
	Moderate (*N* = 58)	.68 (.03)	.71 (.03)	.57 (.03)	.59 (.03)	.04 (.01)	.07 (.01)
	Lower (*N* = 28)	.55 (.07)	.64 (.06)	.43 (.05)	.46 (.06)	.05 (.01)	.08 (.02)
Speed	Faster (*N* = 28)	.63 (.06)	.74 (.05)	.47 (.05)	.50 (.05)	.06 (.02)	.09 (.02)
	Moderate (*N* = 58)	.68 (.04)	.71 (.03)	.57 (.03)	.61 (.03)	.03 (.01)	.06 (.01)
	Slower (*N* = 28)	.62 (.06)	.61 (.06)	.64 (.06)	.61 (.06)	.01 (.01)	.03 (.02)
Slope	Minor (*N* = 28)	.64 (.05)	.71 (.06)	.53 (.06)	.55 (.07)	.03 (.01)	.05 (.02)
	Moderate (*N* = 58)	.66 (.04)	.69 (.04)	.58 (.03)	.58 (.03)	.04 (.01)	.06 (.01)
	Major (*N* = 28)	.66 (.06)	.68 (.05)	.57 (.05)	.62 (.05)	.04 (.01)	.09 (.02)
Distractors	No-Distractors (*N* = 76)	.68 (.03)	.69 (.03)	.60 (.03)	.62 (.03)	.04 (.01)	.06 (.01)
	Distractors (*N* = 38)	.61 (.05)	.72 (.08)	.51 (.05)	.51 (.04)	.03 (.01)	.06 (.02)


Block-1 Targets
Block-2 Targets
Foils
Accuracy split
Speed split
1.00
0.80
0.60
0.40
0.20
0.00
Pre
Post
Higher
Moderate
Lower

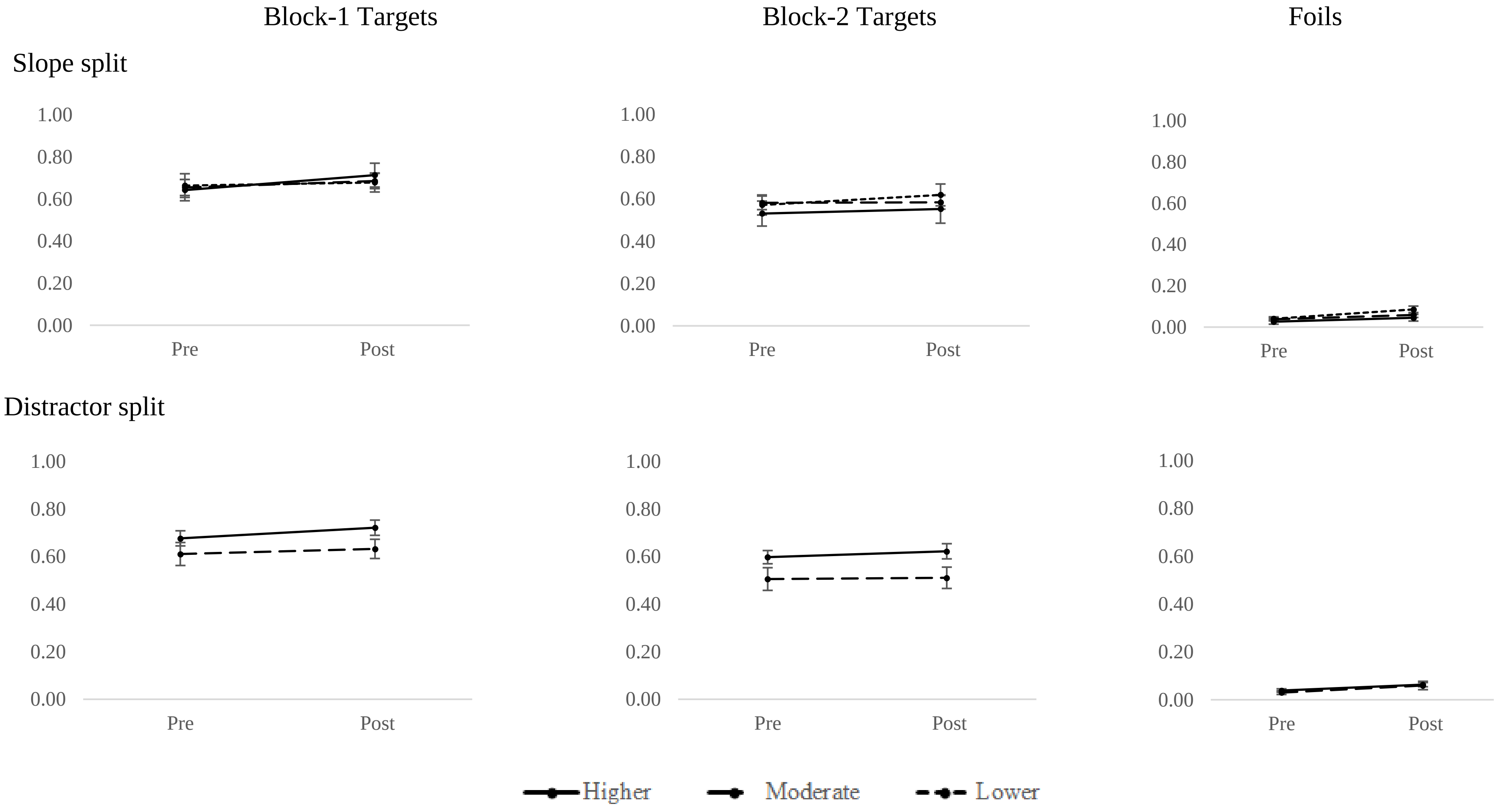


Figure 6.4. Mean accuracy (y axis) scores across pre- and post-test sessions (x axis) for PI-ISR task (Tolan & Tehan, 2002) items: Block-1 Targets, Block-2 Targets, and Foils. Error bars were created from standard error values. Whereby Higher performance refers to Higher Accuracy,

Faster Speed, Minor Slope, and No-Distractors; Moderate performance refers to Moderate Accuracy, Moderate Speed, Moderate Slope; and Lower performance refers to Lower Accuracy, Slower Speed, Major Slope, and Distractors.

6.7.2.1. Accuracy split. There were significant differences between accuracy groups, F (2, 111) = 4.62, p = .012, *MSE* = .12, *partial* η^2 = .08, where the higher accuracy group reported higher ratings of targets and lower foil intrusions, compared to moderate and lower accuracy groups respectively. A main effect for session was found, F (2, 111) = 5.35, p = .023, *MSE* = .02, *partial* η^2 = .05. Performance improved from pre- to post-test. There was no interaction between session and accuracy groups, F (2, 111) = 1.23, p = .30, *MSE* = .02, *partial* η^2 = .02. The items effect was significant, F (2, 222) = 399.90, p <.001, *MSE* = .06, *partial* η^2 = .78, with greater recall under block-1 targets, followed by block-2 targets and foil intrusions respectively. There was an interaction between items and accuracy group, F (4, 222) = 5.38, p <.001, *MSE* = .06, *partial* η^2 = .09. Planned contrasts revealed a significant quadratic interaction, F (2, 111) = 7.37, p = .001, where greater recall was evident under the higher accuracy group, compared to moderate and lower accuracy groups. Group differences were least obvious in the recall of block-1 targets and more pronounced in the block-2 measures. No other interactions were found.

6.7.2.2. Speed Split. There was no significant difference between groups, F (2, 111) = .52, p = .60, *MSE* = .13, *partial* η^2 = .01, with the moderate speed group reporting only slightly greater target rates than faster and slower speed groups respectively. The faster group reported slightly greater foil intrusions than moderate and slower speed groups respectively. There was a main effect for session, F (1, 111) = 4.55, p = .035, *MSE* = .02, *partial* η^2 = .04. An improvement in performance was observed from pre- to post-test. There was no interaction found between session and speed groups, F (2, 111) = 1.81, p = .169, *MSE* = .02, *partial* η^2 = .03. The items effect was significant, F (1.92, 212.71) = 382.17, p <.001, *MSE* = .06, *partial* η^2 = .78, with greater recall under block-1 targets, followed by block-2 targets and foil intrusions respectively. No other interactions were observed.

6.7.2.3. Slope split. There was no significant difference found between slope groups, F (2, 111) = .21, p = .814, *MSE* = .13, *partial* η^2 < .01, where the moderate slope group reported slightly greater target and foil rates than minor and major slope groups respectively. There was a main effect for session, F (1, 111) = 6.12, p = .015, *MSE* = .02, *partial* η^2 = .05, where performance increased from pre-to post-test. There was no interaction between session and slope groups, F (2, 111) = .32, p = .726, MSE = .02, *partial* η^2 = .01. The items effect was significant, F (1.92, 213.14) = 369.55, p < .001, MSE = 23.78, *partial* η^2 = .77, with greater recall under block-1 targets, followed by block-2 targets and foil intrusions respectively. No other interactions were found.

6.7.2.4. Distractor split. There was a significant difference between the distractor groups, F (1, 112) = 4.80, p = .031, MSE = .121, *partial* η^2 = .04, where the No-Distractor group reported greater target accuracy than the Distractor group, however, greater foil intrusions were also reported by the No-Distractor group. There was a main effect for session, F (1, 112) = 4.17, p = .044, MSE = .02, *partial* η^2 =.04. A significant increase in performance was observed from pre- to post test. There was no interaction between session and distractor groups, F (1, 112) = .26, p = .608, MSE < .02, *partial* η^2 = <.01. The items effect was significant, F (1.92, 214.93) = 361.22, p <.001, MSE = .06, *partial* η^2 =.76, with greater recall under block-1 targets, followed by block-2 targets and foil intrusions respectively. No other interactions were observed.

6.7.3. Preliminary discussion.

The manipulation of adding rhyming distractors to the second block had the desired effect in that there was only a modest increase in target recall at post-test and there was a significant increase in foil intrusions. Similar to the previous experiments, target recall improved from pre-test to post-test, with group differences found in the accuracy and distractor measures. Most importantly there were no group by session interactions on any of the four measures (i.e., accuracy, slope of accuracy, completion time, and distractor responses) of splitting the data. In sum, there was no evidence of depletion transfer effects on this task.

6.8. Study 3 General Discussion

The current study was designed to test the third functional marker of the strength model that depletion transfer effects would be observed on the outcome tasks using the sequential task paradigm. Given the outcomes of Studies 1 and 2 where the overall accuracy measure did show performance characteristics consistent with the strength model, it was expected that the third functional marker would be evident with this measure in the current experiments. It was therefore argued that those who maintained higher levels of accuracy across the five stories would not be depleted to the same extent as those who did not maintain high levels of accuracy (i.e., moderate and lower accuracy groups). This would be evident in greater group differences on the post-test performance than pre-test group differences. Although, repetition effects were expected to occur showing improvements in overall performance at post-test. The lower accuracy group's repetition effects were expected to be offset by depletion effects. Resulting in weaker repetition effects, no repetition effects, or a deterioration in performance. This expected pattern of performance did not emerge in any of the experiments; when group splits were based on letter-crossing accuracy, there was no evidence for differential depletion transfer effects. In contrast to the outcomes of Studies 1 and 2, there was no support for the third functional marker of the strength model. Likewise, when groups were split based upon completion time or slope of accuracy, expected individual differences between groups in depletion transfer effects were noticeably absent. Although, neither completion time nor slope of accuracy measures provided functional markers across the current studies, therefore these two measures were determined to be ineffective in studying the strength model.

Interestingly, when the participants were grouped on distraction response rates, the expected outcomes appeared in Experiments 6 and 7, but not Experiments 5 and 8. In the storage processing component of the OSPAN task (i.e., word component), and the ISR task, those who did not respond to letter-crossing distractors showed strong repetition effects. Those who responded to at least one distractor, showed less improvement from pre-test to post-test on OSPAN and ISR outcome tasks. In these experiments, there were differential depletion transfer effects that do support the third functional marker of the strength model. However, the emergence of these outcomes in two memory tasks involving updating and binding abilities, but not replicated in a third (i.e., the PI-ISR task), is problematic.

Critically, there are three assumptions that may compromise the conclusion that the accuracy split fails to show depletion transfer effects. First, the extreme groups method in the current analyses was based on the assumption that one group of participants will be more depleted by the letter-crossing task than another. As of yet, there is no evidence to support this assumption. Second, depletion transfer effects may have been comprised by the lack of a comparison control group. As all participants completed the letter-crossing task, designed to induce self-regulation depletion, any depletion transfer effects may have been equal within each group and hidden by repetition effects. Third, it was assumed that the pre-test tasks would not sufficiently deplete self-regulatory resources. As executive functioning tasks have been employed previously as depletion inducing measures, refer to section 2.5., depletion transfer effects could occur from the administration of the pre-tests onto the letter-crossing task. This would likely influence group splits and their respective outcome task performance. Post-hoc analyses were conducted to address these issues.

6.9. Post Hoc Analyses

Prior to the analyses, power analyses were conducted employing G*Power software (Faul et al., 2009). This determined that the total sample size (N = 466) with three groups would have significant power (α = .05, β = 0, Power = 1, with a critical F value of 2.37) for the mixed design ANOVA. Likewise, the smallest total sample (N = 163) for experimental (n = 98) and control groups (n = 65) under the Stroop colour word task (Golden & Freshwater, 1978) reported significant power (α = .05, β = .01, Power = .99, with a critical F value of 3.89) for the mixed design ANOVA. Refer to Appendix E and F respectively. The total sample under the OSPAN task (N = 174) for experimental (n = 145) and control groups (n = 29) reported significant power for the mixed ANOVA, α = .05, β = .01, Power = .99, with a critical F value of 3.89. As did the ISR task total sample (N = 165) for experimental (n = 109) and control groups (n = 56), α = .05, β = .01, Power = .99, with a critical F value 3.89.

6.9.1. Depletion within performance groups.

In a post-hoc examination of the data in Study 1 (Chapter 4), the overall accuracy measure was generated for all 466 participants who completed the letter-crossing task. Participants were split into three accuracy groups (higher, moderate, and lower) on the same basis as the above studies. The change in performance across the five studies was then examined. The assumption that the higher accuracy group would be less depleted than the low accuracy group would be reflected in differences in performance across the five stories. The higher accuracy group was assumed to present a more consistent performance with a minor slope of accuracy compared to the lower accuracy group, which was predicted to present larger decrements across stories in a major slope of accuracy. This would result in a group by story interaction, or alternative group differences on the slope measure. The data representing the post-hoc analysis are presented in Figure 6.5.

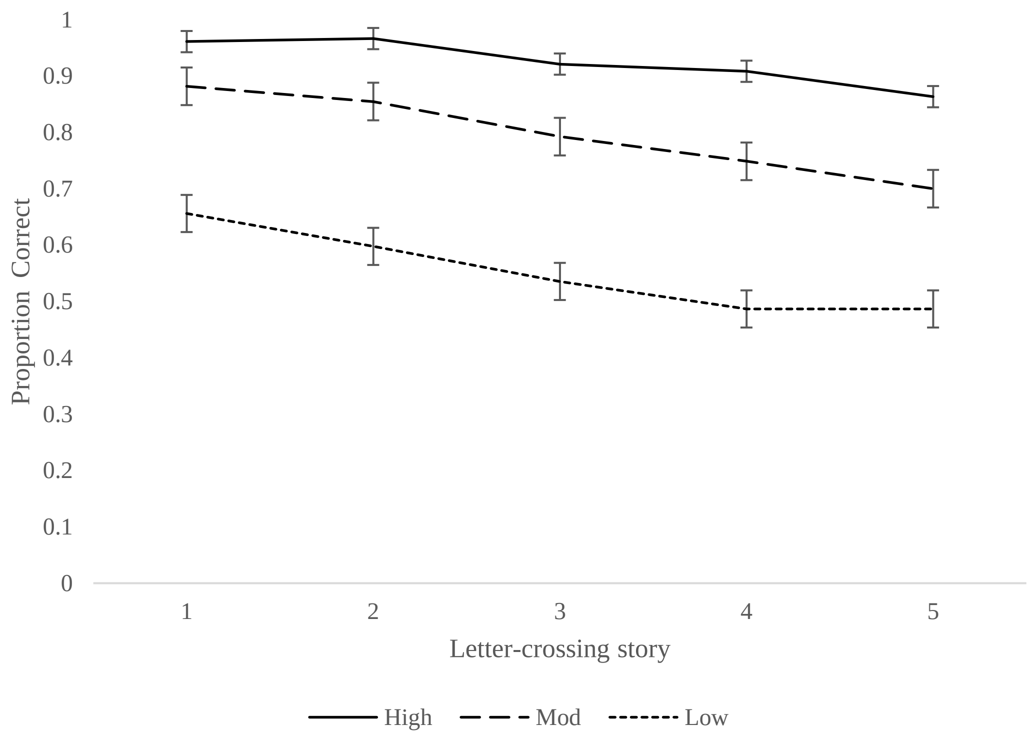


Figure 6.5. Mean accuracy scores across Letter-Crossing Task stories (1-5) for accuracy group splits (High: Higher, Mod: Moderate, Low: Lower) in the total sample. Error bars are standard errors of the mean.

A 3 (Group: Higher, Moderate, and Lower Accuracy) x 5 (Letter-Crossing Story: 1, 2, 3, 4, 5) mixed design ANOVA was conducted for the total sample. The outcomes were that the interaction between accuracy group and letter-crossing story was statistically significant, F (8, 1852) = 6.24, $p < .001$, $MSE = .01$, $\eta^2 = .03$. While performance did deteriorate across stories for the higher accuracy group, the deterioration occurred later and was less pronounced than for the other two letter-crossing accuracy groups. In short, there is support for the working assumption that as far as target accuracy goes, some participants showed more evidence of self-regulation depletion.

6.9.2. Mediation analyses.

While it is valid to assume that some individuals may be more depleted than others may, a number of concerns have been raised on individual difference methods. Notably, robust experimental tasks, such as executive functioning tasks, are known to have low between-subject variability, which ultimately results in low individual difference reliability (Hedge, Powell, & Sumner, 2018). That is, lower performance individuals may not consistently perform in the lower quartile of the performance range. Similarly, the correlation between two measures is impacted by their respective reliability coefficients. High reliability coefficients found in robust tasks will provide strong correlations, which conclude shared processing between tasks. This effect feeds back into the task impurity problem of executive functions previously raised by Miyake and colleagues (2000). This means that reliable executive measures may be reporting executive constructs that are consistent across individuals, but are not reporting the differences between individuals (Hedge, Powell, & Sumner, 2018). The results from Study 3 may then be depicting this performance variance, which resulted in the rejection of the third functional marker of the strength model. Likewise, the results from Study 2, which investigated the relationships between the letter-crossing task and executive measures, may have provided strong correlations between measures due to the robust executive measure reliabilities and provided an incorrect conclusion as a result. To determine whether letter-crossing performance had any impact on post-test performance, it would be theoretically consistent to determine whether letter-crossing performance mediated outcome task performance, exclusive of individual difference splits. This method and similar regression methods have been previously employed to assert the impact on performance following self-regulation depletion (Dang, 2016; Drummond & Philipp, 2017).

In order to test this, a post hoc analysis was conducted by a series of mediation analyses using the PROCESS macro (Hayes, 2013; model 4) on Stroop (N = 98), OSPAN (N = 145), ISR (N = 109), and PI-ISR (N = 114) datasets. The alpha coefficient is at the .05 level unless listed otherwise. To increase chances of detecting an indirect effect of letter-crossing performance on pre-test to post-test performance, the analyses were bootstrapped (5000 bootstraps). Mediators (letter-crossing accuracy, total time, total slope, and distractors recalled) were entered simultaneously to explore the direct and indirect influences of the predictor (pre-test performance) on the outcome variable (post-test performance).

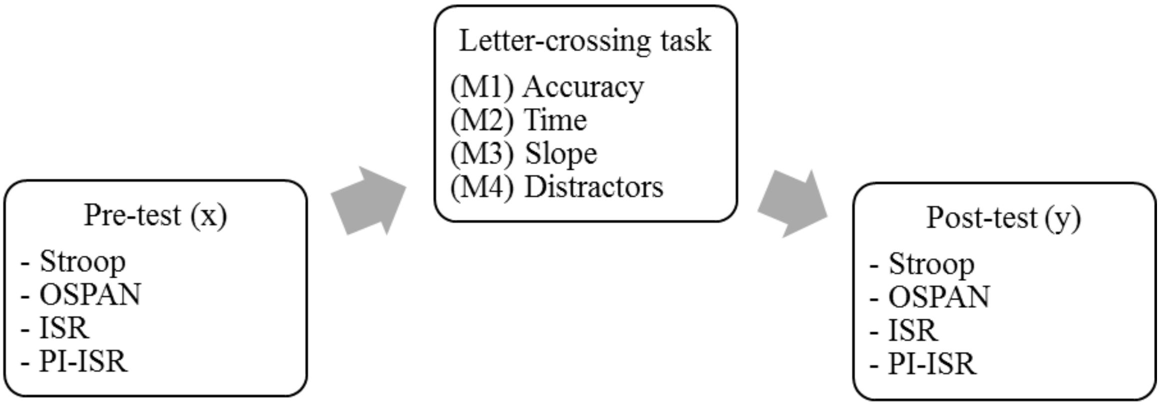


Figure 6.6. A conceptual diagram of letter-crossing performance mediating pre-test performance on post-test performance: an indirect effect.

There were no indirect effects for average Stroop pre-test performance on average Stroop post-test scores through letter-crossing measures (accuracy, total time, slope, and distractors recalled). There was an indirect effect of average OSPAN pre-test performance on average OSPAN post-test performance through letter-crossing accuracy, $b = .06$, 95% CI [.01, .13]. The indirect model explained 80% of the variance on OSPAN post-test performance, whereas the direct model (pre-test performance on post-test performance) explained 78% of the variance on the OSPAN post-test performance. Likewise, there was an indirect effect of average ISR pre-test performance on average ISR post-test performance through letter-crossing accuracy, $b = .03$, 95% CI [.01, .08]. The indirect model explained 54% of the variance on ISR post-test performance, whereas the direct model explained 48% of the variance on ISR post-test performance. There were no indirect effects for average PI-ISR pre-test performance on average PI-ISR post-test performance through letter-crossing performance.

It was therefore concluded that accurate letter-crossing performance does mediate OSPAN and ISR performance. This follows the results from Study 2, which suggests that an executive ability is shared between the letter-crossing task and OSPAN and ISR tasks. Although, this mediation effect was not found for Stroop and PI-ISR tasks which reported similar correlations in Study 2. The lack of mediation onto Stroop and PI-ISR task performance suggests that robust reliability coefficients may have influenced the results in Study 2 for these tasks (experiments 1 and 4). Alternatively, study 3 reported letter-crossing distractor resistance showed strong repetition effects on OSPAN and ISR outcome tasks. This was not evident under the mediation analysis.

6.9.3. Depletion between groups.

To test the second and third assumptions that the pre-test tasks did not sufficiently deplete self-regulation resources and that all participants may have been depleted equally on the letter-crossing task, data were collected from control groups for the Stroop colour word (*N* = 65), OSPAN (*N* = 29) and ISR (*N* = 56) tasks. These participants were administered the pre-test and post-test administrations of one of the outcome tasks and spent 15 minutes talking with the experimenter instead of completing the letter-crossing task. The assumption of the control condition is that talking is not as cognitively demanding as the letter-crossing task and therefore should not deplete self-regulation resources. Repetition effects should be greater under the control condition if depletion transfer effects occurred from the letter-crossing task to the outcome tasks under the experimental condition. Pre-existing differences between experimental and control groups may or may not exist at pre-test, any differences should be magnified at post-test if depletion effects are to be observed in the data set.

A 2 (Group: Experimental, Control) x 3 (Subtask/Items: Word-Reading/ 2-Syllable, Colour-Naming/ 3-Syllable,) x 2 (Session: Pre-test, Post-test) mixed method ANOVA was run for Stroop colour word and ISR tasks. A 2 (Group: Experimental, Control) x 2 (Component: Words, Maths) x 2 (Session: Pre-test, Post-test) mixed method ANOVA was conducted for the OSPAN task. Figure 6.7 shows the pre-test post-test differences for selected conditions in the three tasks. These three selected conditions were considered to be more critical of the task components. The colour-word naming subtask on the Stroop colour word task (Golden & Freshwater, 1978) is the only incongruent condition under the task measuring inhibition. The word component under the OSPAN task (Turner & Engle, 1989) did not achieve ceiling effects on the pre-test, unlike the maths component, therefore this component allowed space for observing repetition effects between experimental groups. The 4-syllable items under the ISR task presented as the most difficult under the ISR task from word length effects, once again allowing possible repetition effects to be observed. Other conditions in each task showed similar patterns.

The series of post-hoc ANOVAs confirmed that the interaction between group and test session was not statistically significant, F (1, 162) = 1.04, p = .309, *MSE* = 64.09, *partial* η^2 = .01, F (1, 172) = .30, p = .298, *MSE* = 6.64, *partial* η^2 < .01, and F (1, 163) = .25, p = .615, *MSE* = .02, *partial* η^2 < .01, for the Stroop, OSPAN, and ISR outcome tasks respectively. The current experimental design relies on the assumption that the control group required no self-regulation resources for conversing with the experimenter for 15 minutes. As such, there is no direct evidence that the conversation was less depleting than the letter-cancelling task and it is not at all clear how such an assumption could be empirically tested. The lack of difference between the two groups suggests that depletion transfer effects did not occur between the letter-crossing task and the pre- and post-test executive functioning tasks. In short, experimental and control groups showed the same strength of repetition effects across the three tasks. The results of these analyses are consistent with the conclusion than the accuracy measure of the letter-crossing task is not a good measure of the depletion of self-regulatory resources.

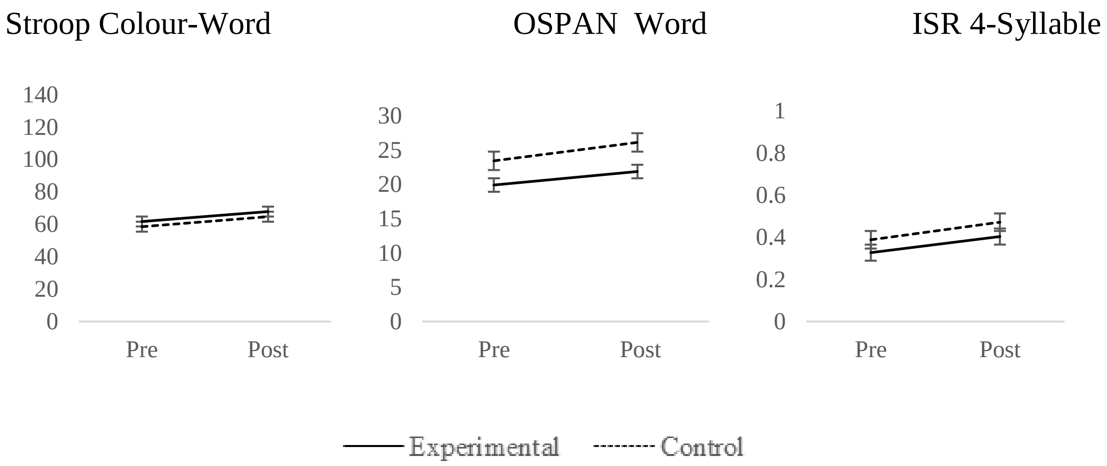


Figure 6.7. Mean proportion (y-axis) scores across pre- and post-test sessions (x-axis) for a representative subset of tasks. Error bars were created from standard error values.

Something for future consideration is the distractor response measure. The distractor group by session interaction, representing the functional marker of depletion effects, was found under the OSPAN and ISR tasks. On face value, distractor responses (i.e., responding to *ei* combinations) represent an obvious failure of the task goal in self-regulation, which suggests that this may be a more appropriate measure of depletion effects. While the distractor responses have face validity for self-regulation in the letter-crossing task, it is not clear how processes under the outcome tasks involves self-regulation, such as the ability to recall items in their correct order under the ISR task. In the PI-ISR task, which requires the regulation of responses by recalling the target instead of the foil and there is some face validity for common self-regulation across tasks, the distractor group by session interaction did not emerge. The second problem is that the interaction did not replicate consistently across the tasks used in the experiments. Distractor responses on the letter-crossing task are the most promising predictor for finding depletion transfer effects on outcome tasks, but the evidence for their utility needs to be explored further.

Therefore, in general these findings are not supportive of the strength model of self-regulation. The absences of depletion transfer effects in majority of the current results reflect the difficulty in reproducing the depletion effect that underpins the replication crisis facing the depletion literature. While Study 2 revealed that overall accuracy measures under the letter-crossing task were related to the executive functioning tasks, thereby suggesting a shared processing capacity between the two domains, the current study’s findings suggest that letter-crossing performance is not related to self-regulation. This suggests that some form of executive functioning underpins the letter-crossing task and cognitive outcome task used here, but these resources are not involved in self-regulation.

6.10. Summary of Chapter

The current chapter investigated the third functional marker of the strength model; depletion transfer effects from the self-regulation task to the outcome task. While the letter-crossing task was correlated with the executive functioning tasks, performance on the letter-crossing task did not show depletion transfer effects. The lack of transfer depletion effects does not provide strong evidence for the strength model of self-regulation. The implications of these results will be discussed further in Chapter 7.

Chapter 7. Conclusions and Implications

7.1 Overview of Chapter

The final chapter of this thesis begins with a review of the original research questions and the main aims of the research previously identified in Chapter 3. The findings from Studies 1, 2, and 3 are then reviewed in relation to the original predictions. This addresses the problems identified within the literature, including the conceptual issues surrounding self-regulation and the scoring problem associated with the letter-crossing task. Theoretical implications of the findings are then considered, and the assumptions originally made are revisited and determined to be reasonable. The methodological strengths of the research are outlined, before the methodological and conceptual limitations of the current studies are considered and addressed in order to make suggestions for future research. The conclusion determines the current findings can be best accounted for by an interaction of cognitive systems involving the common EF factor, WMC, and executive attention.

7.2 Introduction

From the conception of the letter-crossing task in the original article (Muraven et al., 1998), the task has commonly served as a method of inducing depletion without question as to what the letter-crossing task is measuring or how to score this measure. The modified version of the task employed in the current experiments requires participants to respond to any e-vowel combinations in text, however, participants must inhibit reactions to *ei* or *ie* combinations. As the task requires participants to regulate their behaviour, that is, inhibit responses to *ie* or *ei*, it was deemed to require self-regulation processing. Muraven et al. (1998) argued that in order to maintain these regulations and continue task processing energy needed to be expended. Importantly, the energy resource behind the strength model of self-regulation is not exclusively quantifiable. Instead, the energy resource powering self-regulation is inferred from a broad range of measures and can be applied to physiological, cognitive, and behavioural aspects. If one task seems to require effort, regardless of the domain, then this will impact a global energy reserve and subsequent effortful processing will then be negatively impacted (i.e., depletion transfer effects). The sequential-task paradigm then allows the assessment of performance during this depleted period. Although, depletion effects should be appearing as continuing self-regulation processing involve resources from the energy reserve. This thesis attempted to address three problems identified in the self-regulation literature and in doing so, provide empirical data that address the core assumptions of the strength model. These assumptions are that a general pool of resources underpins all self-regulatory tasks, this resource is depleted with repeated acts of self-regulation, and depletion transfer effects onto other tasks will be observed.

The first problem pertains to the conceptual issue of self-regulation, as it is considered a global energy reserve. Identifying this reserve beyond the vague description of a global energy resource is difficult as self-regulation widely entails any regulation of the self in mental or physical efforts. As a result, any task requiring some control or editing of the self can be applied to induce depletion and arguably measures self-regulation. Moreover, as self-regulation is only ever measured after the reserve has been weakened, self-regulation tasks are rarely scored and are consequently are not well understood conceptually. The letter-crossing task was therefore chosen as a self-regulation measure to induce depletion transfer effects, but also to investigate depletion effects within the task itself. The second problem this thesis then attempted to address was to understand what abilities were employed under the letter-crossing task and how these related to performance on a range of outcome tasks. The third problem was then how to best score individual performance on the letter-crossing task.

The thesis followed recommendations set by Lurquin and Miyake (2017) and De Houwer (2011) on how to approach the methodological issues in the field and appropriately establish depletion effects in cognitive science. Functional markers were employed to identify behavioural phenomenon that would be indicative of depletion effects irrespective of what resources were being applied. Cognitive markers were employed to explore the conceptual basis behind self-regulation. Under the letter-crossing task, four scoring methods were created: target accuracy, time taken to complete the task, slope of accuracy over time, and distractor responses. Three functional markers of self-regulation depletion were determined to be behavioural changes over time in the letter-crossing measures; significant correlations between executive functioning tasks and the letter-crossing task thought to employ self-regulation; and pre-test/ post-test differences signifying depletion transfer effects. Of the scoring methods, target accuracy was originally thought to establish how well self-regulatory processing was occurring, completion time provided a measure of processing efficiency, slope of accuracy over time tracked reduced processing due to the weakened energy reserve, and distractor responses allow for identification in self-regulatory failures.

To understand what processes, or cognitive markers, may be employed when engaging in self-control under the letter-crossing task, performance measures on the letter-crossing task were compared to validated executive functioning tasks, as suggested by Lurquin and Miyake (2017). These established correlations shaped the expectations of shared processing between tasks thought to show depletion transfer effects. Conclusions could then be drawn about the cognitive markers of self-regulation and the origins of the self-regulation reserve.

7.3 Conclusions about Research Questions and Hypotheses

The strength model of self-regulation asserts that the self-regulation reserve is domain general, thereby any effortful task requiring self-regulatory processing will use this reserve, and there are individual differences in the reserve with some individuals showing greater depletion effects than others (Baumeister, 2002; Baumeister & Heatherton, 1996; Muraven et al., 1998). The current research was then founded on three assumptions of the strength model. Accurate performance should deplete over time if self-regulation resources are consumed, significant correlations should be identified between letter-crossing and outcome tasks if the resource is shared between these tasks, and significant differences between individuals should be observed on the outcome tasks due to differences in resource capacities. As a result of these assumptions, the thesis set out to achieve three things: a method of scoring the letter-crossing task, finding out what the letter-crossing task was measuring, and replicating depletion transfer effects on outcome tasks.

The first study attempted to identify functional depletion markers within letter-crossing measures over time. The letter-crossing task was scored on four measures: overall accuracy, slope of accuracy, time, and distractors. Following the strength model of self-regulatory depletion, it was assumed that depletion would appear as a downward trend in accuracy over time, with increases in errors such as omissions and failures (i.e., letter-crossing distractors), and slowed processing speeds as the self-regulation resource is expended with ongoing exertion. The accuracy measure provided the best functional marker of depletion effects, and this was observed over the five letter-crossing stories for the total sample in Chapter 4. The slope of accuracy measure provided a numerical format for the linear trend. There was an improvement in item processing speed and no real change in distractor responses over time, which suggested that these measures were not adequate functional markers of depletion. Reliability of the letter-crossing task was assessed by internal correlations between the measures. The overall accuracy score reported significant correlations with slope of accuracy and distractor measures. Item processing speed did not correlate to any of the letter-crossing measures. Thus, of the four measures of the letter-crossing task, only the target accuracy score exhibited the decline with use that is a functional marker of the strength model. While the first study provided functional markers for depletion effects, it did not provide insight into the cognitive markers employed, that is, what the letter-crossing task was measuring.

The second study then attempted to establish whether shared abilities between the letter-crossing and outcome tasks were present by investigating correlations between task scores. It was originally assumed that if the self-regulation resource was domain general, then the overall accuracy measure identified in the first study would correlate to outcome tasks measuring various executive functions thought to be closely related and possibly impacted by a lower self-regulation reserve, as originally outlined in Chapter 3. Letter-crossing task accuracy was found to be strongly correlated to some aspects of all four tasks used as outcome tasks. The implication of this is that overall accuracy was related to all of the executive abilities tested: inhibition, updating, binding, and binding under proactive interference. These results suggest similar processing is occurring on both letter-crossing and outcome tasks when overall accuracy was the measure of letter-crossing performance. Therefore, the correlations between depletion and outcome tasks provide support for the second functional marker of the strength model.

The results of Study 1 were also reflected in the Study 2 findings in that the slope of accuracy measure did not provide any relationships with executive function task scores and the completion time and distractor measures were not consistent in correlations with executive function task scores. For instance, distractors were correlated with the colour-naming subtask under the Stroop colour word task (Golden & Freshwater, 1978), but not the incongruent condition of the colour-word naming subtask as originally predicted. Thus, in contrast to the accuracy measure, the alternative measures of the letter-crossing task were determined to not be reliable measures of the strength model.

As the letter-crossing accuracy measure was correlated with elements across all of the executive functioning tasks, but did not correlate to simpler processing under the ISR task (i.e., 2- and 3-syllable item binding), it could therefore be argued that the task recruits a common EF factor alike the one identified by Friedman and

Miyake (2017). Executive functions are cognitive abilities that direct cognitive processes in goal-directed behaviour (Friedman & Miyake, 2017). Friedman and Miyake (2017) consider the common EF factor, shared by most cognitive tasks, to be separate to other executive components like inhibition, updating, and shifting that are unique to particular cognitive tasks. They argue that the primary function of the common EF factor is goal-orientation and top-down bias-processing. The unity and diversity account of executive functions allows for this shared factor, and a potential general domain resource, while executive functions can maintain an independent function in their own right (Friedman & Miyake, 2017; Miyake, et al., 2000). Importantly, this common EF factor is separate from the active maintenance of goals in memory storage. Rather the common EF factor involves the selection of goals to follow, the ongoing maintenance and retrieval of relevant goals, the suppression of irrelevant or no-longer-relevant goals, and influencing over processing (Friedman & Miyake, 2017).

The third study in Chapter 6 then assessed whether letter-crossing performance would translate to outcome task performance. It was originally assumed that there would be observable differences between individuals in self-regulation depletion, as there are for self-control in social behaviour, dieting, risk-taking, among others (Baumeister & Heatherton, 1996; Engels, Finkenauer, & den Exter, 2000; Quigley, et al., 2018; Tangney, Baumeister, & Boone, 2004). Individuals were compared through letter-crossing performance group splits in accuracy, slope of accuracy, completion time, and distractor responses. It was expected that performance would deteriorate at a greater rate on the outcome tasks for individuals lower in accuracy, as the previous experiments had indicated that overall accuracy was a good functional marker of the strength model. Because slope, time, and distractor errors were not seen to be good markers of the strength model, there was no expectation that depletion transfer effects would emerge with these letter-crossing measures.

When groups were split on the basis of letter-crossing accuracy, there was no decline in performance from pre-test to post-test for any of the groups. Instead, general improvements were observed from pre-test to post-test across outcome tasks due to repetition effects. Recent evidence has suggested that repetition effects do not interact with individual differences variables like WMC (Delaney, Godbole, Holden, & Chang, 2018), so while performance might improve overall for all participants, any effects of depletion should still be evident via greater post-test differences between groups than pre-test differences. For overall accuracy, pre-test/ post-test differences were the same for all groups for all outcome tasks. In short, there was no evidence for depletion transfer effects.

As expected, there were no signs of the predicted transfer effects with slope of accuracy and time measures. The expected pattern of greater post-test group differences did emerge on the memory component of the OSPAN task (Turner & Engle, 1989) in Experiment 6 and the ISR task in Experiment 7 when the groups were determined by whether or not they made distractor responses on the letter-crossing task. However, this outcome was not present on the Stroop colour word task (Golden & Freshwater, 1978) and not on the PI-ISR task. The outcomes in Experiments 6 and 7 represent the only results that are consistent with the third functional marker of the strength model.

7.4 Theoretical Implications

The evidence obtained from the current results do not support the strength model of self-regulation as outlined by Baumeister and colleagues (Baumeister, 2002; Baumeister & Vohs, 2007; Muraven & Baumeister, 2000). The strength model posits a global energy reserve powers all self-regulatory behaviour. Evidence for the reserve is derived from lower performance following on from an initial effortful task (i.e., the sequential-task paradigm). There are believed to be differences in self-regulation capacities, which can explain why some individuals appear to have greater self-control over their behaviour in real-world settings (Baumeister & Heatherton, 1996).

The findings from the current experiments found that accurate performance did decline over the course of a task thought to require self-regulation processing, the letter-crossing task, and that performance on this task was strongly related to executive functioning tasks. On face value, this suggested that the letter-crossing task was employing similar processing to that under the executive functioning tasks, and this processing weakened with time. The letter-crossing accuracy scoring measure did produce the functional markers of depletion. Depletion effects should be observed in other scoring measures where, arguably, failures should occur more frequently if the self-regulatory resource is weakening. For which, no differences were observed in self-regulation failures (i.e., distractors) over the course of the self-regulation task. This result suggests that accurate processing is affected over the task, but not goal-failures. In other words, if the self-regulation capacity is weakened, failures of self-regulation should not change, but accurate processing should.

Despite the significant correlations between tasks, individual performance on the letter-crossing task did not predict performance on the outcome tasks as it was expected to, and importantly, depletion transfer effects were not apparent in the data set unless self-regulatory failures (i.e., distractor responses) were considered. Wherein behavioural differences on the post-test through weaker repetition effects presented potential depletion transfer effects, but only on memory tasks where the involvement of self-regulation was not obvious (i.e., OSPAN and ISR) and not on tasks requiring some form of inhibition or suppression of a foil (i.e., Stroop, PI-ISR). This suggested that while the letter-crossing accuracy measure presented as the best functional marker of depletion on the letter-crossing task and depicted the expected downward trend over time, the finding that this measure produced no depletion transfer effects lead to question whether this score was measuring depletion at all. Overall, the current results strongly suggest that the letter-crossing task, whilst it has been argued to require self-regulatory processing, does not. An alternative account is offered, whereby cognitive processes can account for processing under the letter-crossing task, and possibly self-regulation processing capacity.

Friedman and Miyake (2017) put forth the suggestion of a common EF factor, which is effectively the summation of an individual's executive functioning resources in that numerous executive functions (inhibition, updating, shifting, and so on) form this common EF factor. This is in line with the unity and diversity account of executive functions (Miyake et al., 2000), whereby each executive function overlaps with another and may employ lower cognitive processes, such as searching texts for e-vowel combinations. While letter-crossing accuracy was correlated across all of the executive functioning tasks, and was therefore considered to be closely related to the common EF factor, letter-crossing distractors were found to explain differences in performance on some memory outcome tasks (i.e., OSPAN and ISR, but not PI-ISR).

The result that letter-crossing accuracy was strongly correlated to the common EF factor, but distractor susceptibility accounted for differences in updating and binding ability, led to the conclusion that the different letter-crossing measures were tapping into different components of higher order cognition. These findings do not necessarily account for a difference between the common EF factor and WMC. The acceptance of the potential for a system containing both the common EF factor and WMC system, or the possibility that these two systems communicate with each other, or may be identical, may be controversial to some. Importantly, this thesis does not distinguish between the cognitive markers in these schools of thought. For simplicity in this thesis' argument, the two systems will be considered as more or less the same system. Whereby, letter-crossing accuracy is shared with the common EF factor, and distractor processing is accounted for by memory processing in WMC. The simplest model that can account for these results is goal-maintenance and conflict-resolution processing.

Friedman and Miyake (2007; 2008; 2017) refer to the common EF factor as the ability to maintain goals and bias-processing favouring goal-relevant behaviour. This common EF factor, also known as goal-maintenance, is generated from all executive functions, which are also arguably independent abilities in their own right. Where individuals will have different capacities in the common EF factor, but also variance in capacities within the independent abilities forming this factor. While all executive functioning tasks require effective goal-maintenance, including simpler processing such as binding under short-term memory tasks, this function is more critical under conditions of interference or distractors. Where insufficient processing towards goal-maintenance will result in goal-neglect (Friedman & Miyake, 2017). This is where conflict-resolution processes should alert cognitive systems about the malalignment of task goals. As a result, inhibition is commonly linked with goal-maintenance and ongoing processing of the task goal (Friedman et al., 2008; Friedman & Miyake, 2017). This should have been reflected in the current results. Instead, only accurate letter-crossing processing was found to be related to tasks requiring inhibition or suppression (i.e., Stroop and PI-ISR tasks), rather than the failures of conflict-resolution processing. This suggested then that another system was responsible for conflict-resolution. While the results suggested a link between distractor susceptibility and memory processing on the outcome tasks, conflict-resolution is not thought to be undertaken by WM (Unsworth, Redick, Spillers, & Brewer, 2012). It is possible then that executive attention is also critical to this high level of cognitive processing. Where failures in self-regulation, or momentary lapses in goal-maintenance and conflict-resolution processing, could be explained through diverted or insufficient attention.

Engle (2018) argues that controlled attention is required for the maintenance and disengagement of information, if required. The updated account follows that WMC forms the maintenance of information such as goals under divergent thinking, whereas the old executive attention account followed that WMC was the ability to control attention in the face of distraction and interference conditions (Engle, 2018). What has been previously termed by Muraven et al. (1998) as mental energy, crucial for sufficient self-regulation, may then refer to this control in effortful cognition.

Executive attention may then be involved in the common EF factor of goal-maintenance by directing an individuals' attention away from goal-irrelevant stimuli and towards goal-relevant information and behaviour, which can then allow an individual to achieve short- or long-term goals. Furthermore, executive attention may direct controlled processing towards critical cognitive features such as conflict-resolution mechanisms that override incorrect responses in accordance with the standards being maintained. An individual's WMC, as explained partly through updating ability, could then dictate an individual's susceptibility to disruptions in goal-maintenance. Updating ability, but not necessarily WMC, remains the best method in deducing susceptibility to conflict-resolution processing failures. This was evident in memory processing in the link found between the letter-crossing task and updating ability, as measured by the OSPAN task, and binding ability, as measured by the ISR task.

Although not cognitively demanding, the letter-crossing task arguably requires some component of executive cognition. The decline in performance over the letter-crossing task could be explained through a decline in executive attention towards effective goal-maintenance, and reduction in biased-processing towards the goal, which may include conflict-resolution processing. Of which, differences in self-regulation failures could be assumed to reflect differences in these abilities. The findings can then be accounted for by a combination of executive attention, the common EF factor, and WMC. At this point in time, these systems are too difficult to theoretically split. The conclusion that self-regulation may be a summation of cognitive abilities working as one system or multiple systems communicating with one another, then provides the simplest answer for the current findings.

7.5 Assumptions

The current research was reliant of a number of assumptions. First, that depletion effects would be apparent over time, as the self-regulation energy reserve would be weakened with ongoing processing. This was a fair assumption to make, given that the finite self-regulation resource weakens with repeated use, or in the current terminology, weakens over time (Muraven & Baumeister, 2000). The presence of the predicted behavioural changes in self-regulation processing capacity were identified within the letter-crossing task. These behavioural changes may be due to effortful tasks prior to the letter-crossing task (i.e. the pre-test executive functioning measures). Alternatively, increased effortful processing or speed-accuracy trade-offs may account for letter-crossing performance detriments over time instead of self-regulation depletion. It is possible though, participants could expend greater resources over time to maintain high levels of accuracy. That is, there could be depletion of resources across time without there being a corresponding behavioural manifestation of such depletion.

Second, correlations between the letter-crossing task and the outcome tasks should be apparent if both tasks employ the same domain. While some of the outcome tasks were assumed to require self-regulation and display significant correlations with the letter-crossing task measures, for instance, the colour-word naming subtask in the Stroop colour word task (Golden & Freshwater, 1978), some of the other measures were assumed to not require self-regulation. These measures were then expected to not correlate with the letter-crossing task, such as the shorter syllables on the ISR task, the word-reading and colour-naming subtasks of the Stroop task (Golden & Freshwater, 1978). The between-task correlations that failed to reach significance may have been due to the high performance on the letter-crossing task, which obscures the potential significant relationships between tasks. In general, however, the assumption was met, as all of the critical components within the outcome tasks showed correlations with letter-crossing task accuracy suggesting a shared domain between all tasks.

Third, the choice of outcome and self-regulation tasks would measure the concepts this thesis set out to measure. It was assumed that the letter-crossing task was a self-regulation measure, as this task employs self-regulation processing by way of inhibiting learned responses (Muraven et al., 1998). It was therefore assumed this task would measure self-regulation ability. This task does have the advantage that several aspects of performance can be derived, with distractor responses having high face validity as a measure of self-regulation failures. Likewise, it was assumed that the Stroop colour word task (Golden & Freshwater, 1978) was a good measure of inhibition, as this task contained an incongruent subtask that required the override of an automatic response. The PI-ISR task was assumed to measure the ability to suppress a proactive foil, where any foil intrusions on the task were due to a failure in adequate bindings. The OSPAN task is a well validated measure of WMC and by varying task difficulty on the ISR task, the difference between short-term memory and working memory is reduced. In short, the measures that have been used are appropriate for exploring the assumptions that underpin the strength model.

Last, the assumption has been made that depletion transfer effects should occur on the outcome task using an individual differences approach to determining group membership (i.e., higher, moderate, and lower accuracy). It is possible that the absence of depletion transfer effects by using the accuracy measure in the current studies is because both groups (i.e., all participants) have been depleted. This issue was addressed in a post-hoc manner by the standard means of comparing performance of the experimental group to the control group in Chapter 6. While, there was no evidence for depletion transfer effects in either experimental or control groups, it cannot be assumed that the control group did not require self-regulation processing in the 15 minutes of conversing with the experimenter. Regardless, the lack of depletion transfer effects between groups suggests that the individual differences approach we have used is a valid means of testing the assumptions of the strength model. Post hoc mediation analyses confirmed that letter-crossing accuracy mediated on OSPAN and ISR pre to post-test performance, suggesting a shared processing factor between tasks.

7.6 Alternative Accounts of Depletion: Consideration of Motivation

The current research has focused on one specific account of self-regulation, the strength model proposed by Baumeister and colleagues (1998), and has considered the outcomes as though depletion of resources were the only factor that determined self-regulation performance. Other accounts of depletion effects have been proposed. It is highly likely that a complete understanding of self-regulation will be a multi-dimensional capacity.

According to the process model of self-regulation, depletion effects occur due to a lack of motivation, whether internal or external, to continue with ongoing exertion in order to meet set goals (Inzlicht, Legault, & Teper, 2014). Ongoing processing is aversive to the individual. Attention is then redirected to a more pleasant goal where less effortful control is required, which has been termed a cognitive leisure state (Inzlicht et al.,2014). This cognitive leisure state then appears as a depleted resource of self-regulation, when instead an individual's motivations are weakened. The results from the process model rely on the assumption that participants were motivated to employ effortful processing on the task. This is typically artificially encouraged through participant incentives in laboratory experiments. Nonetheless, motivational drive is considered essential for individuals to commit to effortful processing in real-world scenarios.

Recently, Francis, Milyavskya, Lin, and Inzlicht (2018) published a sum of their recent experimental studies. Each of the measures employed some form of executive functioning. The cognitive estimation task requires a quick quantitative estimation. For example, *how much does an average hammer weigh?* The task does not involve crystallised intelligence (*cF*), pertaining to accessing knowledge from long-term memory, rather it measures problem-solving abilities. Similarly, the anchor effect task requires estimations on quantitative questions, however, participants are cognitively biased to estimate larger or smaller estimates than the average, based on previous question. For example, *does the president make over $500,000?,* would influence the estimate for *how much does the president earn?* The Anchor effect then primes participants' average estimations in problem solving. The Flanker task requires the inhibition of automatic responses to visual arrows under incongruent conditions. The add-3 task employs working memory as information maintenance is required (a row of digits presented for one second), whilst participants manipulate the information for recall (add three to each digit). While effortful cognitive tasks (i.e., cognitive estimation task, Flanker task, Anchor effect) were employed as self-control measures in a repeated measures within-subjects design with an add-3 task as the depletion inducing measure, there was no behavioural performance change (Francis et al., 2018). That is to say, depletion effects were not present under the cognitive tasks, but were present under the self-report measures. Within depletion-experimental blocks, participants reported feeling fatigued, however, self-reported depletion effects were absent following recovery blocks in which participants watched a fun video, where participant mood was shown to improve (Francis et al., 2018). A meta-analysis on these studies revealed the depletion effect size was small, though statistically significant (Francis et al., 2018). As these self-reported fatigue effects faded with time within the experiment, the authors suggested that any depletion effects might be too weak when competing with experimental blocks and practice effects in the experimental design (Francis et al., 2018).

This may help to partially explain the current results, as participants may have become less motivated on the letter-crossing task as time went on, depicting a downward trend of what was assumed to self-regulation resources. Participants also showed repetition effects from pre- to post-test and where repletion effects were thought to cover-up any potential depletion transfer effects in the accuracy measures on the outcome tasks. While the distractor measure did show weaker repetition effects, depletion effects may have been better captured by self-report measures, as suggested by Francis et al. (2018). Additionally, there was no interaction between control and experimental groups over the sessions, which may suggest that neither of the groups were motivated enough to invest critical resources in the tasks. This could have been assessed through self-report measures throughout the experiment, and is discussed in the limitations section.

Although, if there was limited, if any, motivation recruited under the data set in the current studies, then this explanation does not address the conceptual issue of identifying the processes employed by the letter-crossing task. A more comprehensive conceptual explanation for the current results can then be accounted for by cognitive processing relating to goal-maintenance.

7.7 Strengths and Limitations

The following section considers the strengths and limitations of the current research. Potential improvements are suggested for future research.

7.7.1. Methodological strengths.

The introduction to the thesis outlined several problems associated with the traditional use of the sequential-task paradigm. First, performance on the depletion task is rarely measured directly. The critical assumptions that are made about processes underpinning the depletion task have never been directly tested through the examination of behavioural changes in the depletion task. There is then little evidence, beyond that demonstrated by Arber et al. (2017) that assumed resource depletion is reflected in behavioural markers. Second, the use of between-groups designs to measure self-regulation presents an issue as this method relies on group differences when administering effortful experimental tasks and easier control tasks. The issue is that these designs have not determined whether either task produces a decrement in performance, or if the experimental task produces substantially greater decrements than the control task. Third, the standard practice of not administering baseline measures of the outcome task cannot eliminate cohort differences for any emergent between-group differences. Last, at the conceptual level using a between-subjects design to measure within-subject processes is less than optimal. Friese et al. (2018) also point to additional methodological issues that compromise the depletion research, citing issues such as sample size affecting statistical power, minimising error variance, controlling for confounding variables, and using within-subject pre-test/post-test designs.

The current methodology has addressed many of these issues. There was a direct examination of the behavioural changes over the depletion task. The outcomes of the current research replicate the findings of Arber et al. (2017) in that target accuracy does deteriorate across time. The current research addresses one of the limitations of the Arber et al. (2017) research by including the examination of other measures of letter-crossing performance besides target accuracy. By adopting an extreme groups approach with all participants undergoing the same self-regulation task, the problems associated with using different tasks between groups are eliminated with variance attributable to differences between self-regulation and control tasks being minimised. Using a repeated measure within-groups design allows for cohort differences to be identified and evaluated and does respond to the Friese et al. (2018), recommendation. By creating groups on the basis of their performance on the depletion task, within-subjects processes are more easily identified. Finally, the use of sample sizes in excess of 100 participants for the most part, results in better statistical power than has been the case in many previous studies.

7.7.2. Methodological limitations.

First, the conclusions are limited by the use of a single depletion-inducing task. Additional depletion-inducing measures should also be employed in order to identify a common self-regulation domain, however, these tasks should also be amenable to measurement throughout the course of depletion, such as the letter-crossing task. Writing essays using the restriction of particular letters would also be a suitable task to track performance over time, as this task has the ability to score prohibited letters as self-regulation failures. Tasks such as not thinking of white bears would be less optimal as a depletion task, due to the difficulties in task measurement and the reliance on accurate self-report for self-regulation failures (i.e., white bear thoughts).

Second, there was no effort manipulation checks made throughout the depletion task. That is to say, the studies did not assess whether participants found the letter-crossing task to be effortful and demanding enough to theoretically induce depletion. While there is a decline in accurate performance, it is unknown whether this effect is this due to the task demands or if participants invested less effort over time. While the utility of manipulation checks have been questioned (Fayant, Sigall, Lemonnier, Retsin, & Alexopoulos, 2017), if they are to be used self-report effort manipulation checks should be presented throughout the task to allow for the assessment of the individual effort invested for each story.

Third, no incentives were employed as recommended by Lee et al. (2016), in order to encourage effort and motivation in participants in self-regulation research. The decline in accuracy over the letter-crossing task may have been due to a decline in motivation to engage in identifying letter-crossing targets, instead of a functional marker for self-regulatory depletion. Lee et al. (2016) also recommended increasing the length of the first task in the sequential-task paradigm. Under the current studies, the letter-crossing task took an average of 17.01 minutes (*SD* = 7.62) with a range of 63.97 minutes (minimum: 3.48, maximum: 67.45). In addition, a pre- and post-test were included in a sequential-task paradigm in order to establish a baseline prior to assessing performance on the post-test. This allowed for the finding that depletion transfer effects were generally not found. If participants had a lack of motivation to commit to effortful processing on the tasks, then it is safe to assume that this should not account for the increase in performance on the post-test as participants would not be motivated to increase their performance on the follow-up task. Regardless, while the current tasks were arguably effortful and lengthy, future research should include an incentive to encourage participants, in addition to self-reported measures throughout the task, to assess task engagement to eliminate any motivational doubts.

Fourth, the decision to administer the letter-crossing task based on task completion, rather than a time-based method of completion, may present another limitation. As large differences in completion time were reported, with some individuals taking over an hour to complete the letter-crossing task. These individuals with slower letter-crossing processing may have been more at risk of depletion effects, however, an indefinite period allowed for more accurate processing to occur with potentially fewer self-regulatory failures (i.e., distractors) as a result. While the time measure of item-processing speed did not present as a valid measure for self-regulation under the letter-crossing task, a cut-off time of 10 minutes may have been efficient enough to establish more accurate depletion readings. That is to say, slower letter-crossing processing individuals may have then made more errors. Although, this would have presented issues with comparing scores between individuals due to the potential variance in task-completion.

Last, the outcome tasks selected may have also limited the results. The Stroop colour word task (Golden & Freshwater, 1978), while used in cognitive batteries as it is sensitive to neuropsychological impairments, did not include a reaction time score, the subtasks were only 45 seconds in total, and there was only one incongruent condition (i.e., the colour-word naming subtask) compared to other modified Stroop tasks. Depletion transfer effects may then not have been identified due to the nature of the task compared to other modified Stroop tasks that incorporate mixed trials (i.e., both congruent and incongruent trials), do not impose 45-second time-limits, and measure reaction time. Likewise, the post-test session of the PI-ISR task manipulated items by rhyming filler items to prime recall of the foil item in the first block. This likely influenced the results in foil intrusion rates and may have eliminated potential evidence for depletion transfer effects, although, this manipulation was added to reduce repetition effects.

7.7.3. Conceptual Strengths.

A number of authors have indicated that a major reason for the problems associated with the depletion research is that there is a lack of care in theorising and formalising of assumptions that allow for unambiguous confirming or disconfirming evidence (Friese et al., 2018; Lurquin & Miyake, 2017). One key issue is the operationalisation of the theoretical constructs. More specifically, how self-regulation is operationalised in depletion and outcome tasks. The current experiments address these issues to some extent. The letter-crossing task is useful in that multiple scoring measures can be derived from the task. What has been labelled as a distractor response, responding to an *ei/ie* combination when instructed not to, represents an operational definition of failure in self-regulation on this task. Accuracy in target identification, rate of deterioration of accuracy, and completion time are other aspects of performance on the task that may be important but it is much more difficult to see how these are direct measures of self-regulation.

In a similar vein, the outcome tasks have been selected on the basis of being the subject of many experiments that have validated the underlying processes in each task. In the current version of the tasks used, the foil intrusion in the PI-ISR task is the only direct measure of a goal-neglect occurrence in the tasks employed. Indirect measures of self-regulation could be hypothesised in each of the outcome tasks, but there is no obvious requirement in the Stroop, OSPAN, or ISR tasks employed that require the participant to refrain from a particular response, save for the incongruent Stroop condition (i.e., the colour-word naming subtask).

The current research contributes at the conceptual level. Accuracy on the letter-crossing task is related to performance on the outcome tasks. Given the established role of executive function and WMC in these outcome tasks, the current research suggests that the same cognitive resources support accurate target detection under the letter-crossing task. Moreover, the research suggests that these common resources are expended over time in the letter-crossing task. The absence of depletion transfer effects also suggests that the common EF resources are not necessarily involved in self-regulation.

7.7.4. Conceptual limitations.

Due to the vague and undefined nature of the self-regulation reserve, this thesis has encountered a number of operational issues, and as a result of this, cannot completely dismiss the energy resource described by Baumeister and colleagues (Muraven et al., 1998). Firstly, the self-regulation resource has not been effectively operationalised, which proves difficult when attempting to measure activity from the reserve. Regardless of any findings that may suggest self-regulation is accounted for by cognition or motivation, it could continue to be argued that an abstract energy resource influences these capacities in some way, whether self-regulation employs executive functioning or vice versa is hotly debated (Arndt, et al., 2014; Baumeister, 2002; Hofmann, Schmeichel, & Baddeley, 2012; Muraven & Baumeister, 2000; Schmeichel, Vohs, & Baumeister, 2003). Any self-control effort uses reserves within an individual’s self-regulation capacity, resulting in what is referred to as depletion effects. The conclusions made by this thesis are then limited to what the results suggest the letter-crossing task is employing and thereby measuring.

These findings propose that the letter-crossing task largely employs higher order cognition. Although, a major conceptual limitation is the possibility that the letter-crossing task may not be measuring, or employing, self-regulation. Even though this is a widely-used self-regulation task, and on face value the task appears to require regulation over learned responses, the task itself has rarely been investigated. Until the distractor measure on the letter-crossing task can be validated with other measures of self-regulation errors, there is then the possibility that functional markers of depletion were not identified because self-regulation was not involved in any of the experiments. This could be argued for a number of the self-regulation tasks. For example, it is not overtly evident how standing on one leg and counting backwards requires self-regulation (Tyler & Burns, 2008; Webb & Sheeran, 2003). The letter-crossing task was originally chosen because it was an accepted measure for inducing depletion and performance could be tracked over time in a variety of scoring mechanisms. At the time, it presented as the best measure for the current research.

7.8. Implications for Future Research

The understanding of the depletion effect, and what self-regulation is, is criti-cal research. The current studies attempted to address research questions in the current self-regulation field as other fields, such as animal studies (Parrish, et al., 2018), domestic violence (Quigley et al., 2018), and dietary behaviour (Papies, Stroebe, & Aarts, 2008; Powell, McMinn, & Allan, 2017; Stroebe, et al., 2013), uncritically assert the validity of ego-depletion without considering or assessing the replication and conceptual concerns. The current thesis has addressed one critical flaw in the previous research; performance on the depletion task is rarely directly measured. As such, it is not clear what aspects of the depletion task reflect self-regulation and the critical assumption that resources are depleted with exertion is never directly assessed. The current research has provided insight into how to score the letter-crossing task effectively, what processes are shared between letter-crossing ability and executive functioning, and hints at the conceptual domain behind what has been termed self-regulation under the strength model.

As the letter-crossing task has been accepted as a self-regulation measure, which was correlated to components under executive functioning tasks and believed to employ similar executive processes, then perhaps self-regulation should instead be regarded as and measured by executive functioning tasks. Further work needs to address the complex relationship between behaviour and cognition. Focusing on individual differences in susceptibility to self-regulatory failures may answer whether higher cognitive processes, such as active goal-maintenance and conflict-resolution, largely account for these differences. If this is the case, some individuals who are more structured in their goal-pursuit and greater executive attention should have stronger self-control.

In short, this thesis has addressed some of the experimental issues of self-regulation research in the field. Further research is required to understand the broader questions pertaining to the conceptual origins of self-regulation.

7.9. Conclusion

The merging of two separate fields, executive cognition and self-regulation, plus improved experimental methodologies has permitted a deeper examination of the strength model of self-regulation. This has identified the letter-crossing task, and executive functioning, as measured by the Stroop colour word task (Golden & Freshwater, 1978), OSPAN task (Turner & Engle, 1989), ISR task, and PI-ISR task (Tolan & Tehan, 2002), are underpinned by executive resources, but there is some doubt as to whether these resources are actively involved in self-regulation. Although, much alike executive functioning tasks, it is highly likely that the letter-crossing task has a similar task impurity. That is, multiple abilities are employed for successful performance on the task. These results could be satisfied with a conclusion that the letter-crossing task involves the recruitment of the common EF factor, among other cognitive systems.

What has been described as self-control could be considered a multifaceted mixture of influences, such as motivation towards the goal, executive cognition, physiological factors, mood, and so on. Difficulty in consistently replicating the depletion effect may then be due to depletion effects appearing in different forms due to the generalised nature of self-regulation and task modality. For instance, depletion effects may be evident in self-reported motivation measures, if participants do not feel motivated to participate on a task that requires some form of an investment on their behalf, rather than decreases in glucose levels within the bloodstream that fluctuate during the day naturally.

Ultimately, all tasks inside and outside of the laboratory require some form of goal. Regardless of the domain then, self-regulation is the regulation of an individual’s self in order to meet this goal. Self-regulation should therefore not be considered an energy reserve as the strength model posits, rather it should be considered as a goal-maintenance processing system, of which is powered by goal-orientated cognition systems. Self-regulation can therefore be explained through executive cognition, such as the common EF factor, where the employment of a variety of executive abilities (i.e., inhibition, updating, and binding under interference) under executive attention and WMC, helps to secure the regulation of goal-maintenance and resolve conflicts for successful target performance. Nevertheless, any alternative conclusion made to the depletion effect could be deemed to be powered by an energy reserve, so the strength model of self-regulation cannot be refuted. Nonetheless, the executive control account provides a potentially successful challenge to the energy reserve hypothesised by Baumeister and colleagues (Baumeister, 2002; Baumeister et al., 2018). In conclusion, the current research strongly suggests that self-regulation, as measured by the letter-crossing task, is employing and presumably measuring cognitive processing systems relating to goal-maintenance, rather than processing conducted by an energy reserve as previously proposed.

7.10. Summary of Chapter and Thesis

The current thesis has attempted to address some of the methodological issues associated with the replication and conceptual issues of depletion effects in self-regulation. While recognising that self-regulation is multi-faceted and therefore multi-determined, the current focus was on understanding one particular depletion task, the letter-crossing task, from a cognitive standpoint. The thesis identified three functional markers of the strength model, two of which required that performance on the depletion task be measured. While the letter-crossing task has been frequently used as a depletion task, it had not been established what aspect of the task was appropriate for measuring self-regulation, nor had it been established that it was depleting some type of resource. Furthermore, it had not been established that the self-regulation task shared resources with any of the previous outcome tasks that had been used, let alone identified these resources as those involved in self-regulation. The research examined four measures on the letter-crossing task, one of which, overall accuracy, did show evidence for resource depletion over time. By using well understood and validated outcome measures, that resource was identified as executive function resources that are common to many cognitive tasks. Although, the letter-crossing accuracy measure failed to reflect the depletion transfer effects that are the key prediction of the strength model. In contrast, using distractor responses on the letter-crossing task as the operational measure of self-regulation failures did produce the expected depletion transfer effects. Although, these depletion transfer effects were inconsistent. The distractor measure, however, did not show the expected functional characteristic of changing over time. Thereby, the current research did not show consistent support for the strength model of self-regulation. Instead, the current data point towards some form of higher cognition systems, including goal-maintenance and conflict-resolution, being responsible for self-regulation and the depletion of self-regulation. Although, this conclusion was limited to self-regulation under the letter-crossing task.

References

Ackerman, P. L., Beier, M. E., & Boyle, M. D. (2002). Individual differences in working memory within a nomological network of cognitive and perceptual speed abilities. *Journal of Experimental Psychology: General, 131*(4), 567.

Arber, M. M., Ireland, M. J., Feger, R., Marrington, J., Tehan, J., & Tehan, G. (2017). Ego depletion in real-time: an examination of the sequential-task paradigm. *Frontiers in Psychology, 8*, 1672. doi:doi: 10.3389/fpsyg.2017.01672

Arndt, J., Das, E., Schagen, S. B., Reid-Arndt, S. A., Cameron, L. D., & Ahles, T. A. (2014). Broadening the cancer and cognition landscape: the role of self-regulatory challenges. *Psycho-oncology, 23*(1), 1-8.

Baddeley, A. (1996). Exploring the central executive. *The Quarterly Journal of Experimental Psychology Section A, 49*(1), 5-28.

Baddeley, A. D. (2000). The phonological loop and the irrelevant speech effect: Some comments on Neath (2000). *Psychonomic Bulletin & Review, 7*(3), 544-549. doi:10.3758/BF03214369

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. *The psychology of learning and motivation, 8*, 47-89.

Baddeley, A. D., Thomson, N., & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior, 14*(6), 575-589.

Barkley, R. A. (2001). The executive functions and self-regulation: An evolutionary neuropsychological perspective. *Neuropsychology review, 11*(1), 1-29.

Baumeister, R. F. (1982). A self-presentational view of social phenomena. *Psychological Bulletin, 91*(1), 3-26. doi:10.1037/0033-2909.91.1.3

Baumeister, R. F. (1984). Choking under pressure: self-consciousness and paradoxical effects of incentives on skillful performance. *Journal of Personality and Social Psychology, 46*(3), 610.

Baumeister, R. F. (2002). Ego depletion and self-control failure: An energy model of the self's executive function. *Self and Identity, 1*(2), 129-136.

Baumeister, R. F., & Alquist, J. L. (2009). Is There a Downside to Good Self-control? *Self and Identity, 8*(2-3), 115-130. doi:10.1080/15298860802501474

Baumeister, R. F., Bratslavsky, E., Muraven, M., & Tice, D. M. (1998). Ego depletion: Is the active self a limited resource? *Journal of Personality and Social Psychology, 74*(5), 1252-1265. doi:10.1037/0022-3514.74.5.1252

Baumeister, R. F., Hamilton, J. C., & Tice, D. M. (1985). Public versus private expectancy of success: Confidence booster or performance pressure? *Journal of Personality and Social Psychology, 48*(6), 1447.

Baumeister, R. F., & Heatherton, T. F. (1996). Self-regulation failure: An overview. *Psychological Inquiry, 7*(1), 1-15.

Baumeister, R. F., & Steinhilber, A. (1984). Paradoxical effects of supportive audiences on performance under pressure: The home field disadvantage in sports championships. *Journal of Personality and Social Psychology, 47*(1), 85.

Baumeister, R. F., Tice, D. M., & Vohs, K. D. (2018). The Strength Model of Self-Regulation: Conclusions From the Second Decade of Willpower Research. *Perspectives on Psychological Science, 13*(2), 141-145. doi:10.1177/1745691617716946

Baumeister, R. F., & Vohs, K. D. (2007). Self-Regulation, Ego Depletion, and Motivation. *Social and Personality Psychology Compass, 1*(1), 115-128. doi:10.1111/j.1751-9004.2007.00001.x

Baumeister, R. F., & Vohs, K. D. (2016). Misguided Effort With Elusive Implications. *Perspectives on Psychological Science*, *11*(4), 574-575. doi:10.1177/1745691616652878

Beedie, C. J., & Lane, A. M. (2012). The role of glucose in self-control: Another look at the evidence and an alternative conceptualization. *Personality and Social Psychology Review, 16*(2), 143-153.

Beilock, S. L., & DeCaro, M. S. (2007). From poor performance to success under stress: working memory, strategy selection, and mathematical problem solving under pressure. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33*(6), 983.

Benton, D., Owens, D. S., & Parker, P. Y. (1994). Blood glucose influences memory and attention in young adults. *Neuropsychologia, 32*(5), 595-607. doi:https://doi.org/10.1016/0028-3932(94)90147-3

Burgess, N., & Hitch, G. J. (1999). Memory for serial order: a network model of the phonological loop and its timing. *Psychological Review, 106*(3), 551.

Burgess, N., & Hitch, G. J. (2006). A revised model of short-term memory and long-term learning of verbal sequences. *Journal of Memory and Language, 55*(4), 627-652. doi:https://doi.org/10.1016/j.jml.2006.08.005

Carter, E. C., Kofler, L. M., Forster, D. E., & McCullough, M. E. (2015). A series of meta-analytic tests of the depletion effect: self-control does not seem to rely on a limited resource. In: American Psychological Association.

Carter, E. C., & McCullough, M. E. (2014). Publication bias and the limited strength model of self-control: has the evidence for ego depletion been overestimated? *Frontiers in Psychology, 5*(823). doi:10.3389/fpsyg.2014.00823

Carver, C. S., & Scheier, M. F. (1981). The self-attention-induced feedback loop and social facilitation. *Journal of Experimental Social Psychology, 17*(6), 545-568. doi:https://doi.org/10.1016/0022-1031(81)90039-1

Carver, C. S., & Scheier, M. F. (1998). *On the self-regulation of behavior*. New York, NY, US: Cambridge University Press.

Conway, A. R., & Engle, R. W. (1994). Working memory and retrieval: A resource-dependent inhibition model. *Journal of Experimental Psychology: General, 123*(4), 354.

Conway, A. R. A., Kane, M. J., & Engle, R. W. (2003). Working memory capacity and its relation to general intelligence. *Trends in Cognitive Sciences, 7*(12), 547-552. doi:http://dx.doi.org/10.1016/j.tics.2003.10.005

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilheim, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review, 12*(5), 769-786. doi:10.3758/BF03196772

Cook, J. (2012). The Walk. Retrieved from http://www.shortbreadstories.co.uk/story/view/the_walk/#axzz5CnMF38sv

Dahm, T., Neshat-Doost, H. T., Golden, A.-M., Horn, E., Hagger, M., & Dalgleish, T. (2011). Age Shall Not Weary Us: Deleterious Effects of Self-Regulation Depletion Are Specific to Younger Adults. *PLoS ONE, 6*(10), e26351. doi:10.1371/journal.pone.0026351

Dang, J. (2016). Commentary: A Multilab Preregistered Replication of the Ego-Depletion Effect. *Frontiers in Psychology, 7*, 1155. doi:10.3389/fpsyg.2016.01155

Dang, J. (2017). An updated meta-analysis of the ego depletion effect. *Psychological Research*. doi:10.1007/s00426-017-0862-x

Dang, J., Björklund, F., & Bäckström, M. (2017). Self-control depletion impairs goal maintenance: A meta-analysis. *Scandinavian Journal of Psychology*.

Danziger, S., Levav, J., & Avnaim-Pesso, L. (2011). Extraneous factors in judicial decisions. *Proceedings of the National Academy of Sciences, 108*(17), 6889-6892. doi:10.1073/pnas.1018033108

Darby, D., Maruff, P., Collie, A., & McStephen, M. (2002). Mild cognitive impairment can be detected by multiple assessments in a single day. *Neurology, 59*(7), 1042-1046.

Das, A. (2013). The Magician. Retrieved from http://www.shortbreadstories.co.uk/story/view/the_magician/#axzz5CnMF38sv

De Houwer, J. (2011). Why the cognitive approach in psychology would profit from a functional approach and vice versa. *Perspectives on Psychological Science, 6*(2), 202-209.

Delaney, P. F., Godbole, N. R., Holden, L. R., & Chang, Y. (2018). Working memory capacity and the spacing effect in cued recall. *Memory, 26*(6), 784-797. doi:10.1080/09658211.2017.1408841

Donohoe, R. T., & Benton, D. (1999). Cognitive functioning is susceptible to the level of blood glucose. *Psychopharmacology, 145*(4), 378-385. doi:10.1007/s002130051071

Drummond, A., & Philipp, M. C. (2017). Commentary: "Misguided Effort with Elusive Implications" and "A Multi-Lab Pre-Registered Replication of the Ego Depletion Effect". *Frontiers in Psychology, 8*(273). doi:10.3389/fpsyg.2017.00273

Duff, K., Callister, C., Dennett, K., & Tometich, D. (2012). Practice Effects: A Unique Cognitive Variable. *The Clinical Neuropsychologist, 26*(7), 1117-1127. doi:10.1080/13854046.2012.722685

Engels, R., Finkenauer, C., & Den Exter, B. (2000). Parental influences on self-control and juvenile delinquency. *Manuscript in preparation, Utrecht University, Netherlands*.

Engle, R. W. (2018). Working Memory and Executive Attention: A Revisit. *Perspectives on Psychological Science, 13*(2), 190-193. doi:10.1177/1745691617720478

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G* Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*(4), 1149-1160.

Fayant, M.-P., Sigall, H., Lemonnier, A., Retsin, E., & Alexopoulos, T. (2017). On the limitations of manipulation checks: An obstacle toward cumulative science. *International Review of Social Psychology, 30*(1).

Francis, Z. L., Milyavskaya, M., Lin, H., & Inzlicht, M. (2018). Development of a Within-Subject, Repeated-Measures Ego Depletion Paradigm: Inconsistent Results and Future Recommendations.

Friedman, N. P., & Miyake, A. (2017). Unity and diversity of executive functions: Individual differences as a window on cognitive structure. *Cortex, 86*, 186-204.

Friedman, N. P., Miyake, A., Young, S. E., DeFries, J. C., Corley, R. P., & Hewitt, J. K. (2008). Individual differences in executive functions are almost entirely genetic in origin. *Journal of Experimental Psychology: General, 137*(2), 201.

Friese, M., Loschelder, D. D., Gieseler, K., Frankenbach, J., & Inzlicht, M. (2018). Is ego depletion real? An analysis of arguments. *Personality and Social Psychology Review*, 1088868318762183.

Gailliot, M. T., Baumeister, R. F., DeWall, C. N., Maner, J. K., Plant, E. A., Tice, D. M., . . . Schmeichel, B. J. (2007). Self-control relies on glucose as a limited energy source: willpower is more than a metaphor. *Journal of Personality and Social Psychology, 92*(2), 325.

Golden, C. J., & Freshwater, S. M. (1978). Stroop color and word test.

Hagger, M. S., & Chatzisarantis, N. L. D. (2016). A Multilab Preregistered Replication of the Ego-Depletion Effect. *Perspectives on Psychological Science, 11*(4), 546-573. doi:10.1177/1745691616652873

Hagger, M. S., Wood, C., Stiff, C., & Chatzisarantis, N. L. D. (2010). Ego depletion and the strength model of self-control: A meta-analysis. *Psychological Bulletin, 136*(4), 495-525. doi:10.1037/a0019486

Hall, P. A., & Fong, G. T. (2007). Temporal self-regulation theory: A model for individual health behavior. *Health Psychology Review, 1*(1), 6-52.

Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York, NY, US: Guilford Press.

Healey, M. K., Hasher, L., & Danilova, E. (2011). The stability of working memory: Do previous tasks influence complex span? *Journal of Experimental Psychology: General, 140*(4), 573.

Hedge, C., Powell, G., & Sumner, P. (2018). The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behavior Research Methods, 50*(3), 1166-1186. doi:10.3758/s13428-017-0935-1

Henson, R. N. A. (1998). Short-Term Memory for Serial Order: The Start-End Model. *Cognitive psychology, 36*(2), 73-137. doi:https://doi.org/10.1006/cogp.1998.0685

Hofmann, W., Schmeichel, B. J., & Baddeley, A. D. (2012). Executive functions and self-regulation. *Trends in Cognitive Sciences, 16*(3), 174-180. doi:http://dx.doi.org/10.1016/j.tics.2012.01.006

Inzlicht, M., & Berkman, E. (2015). Six Questions for the Resource Model of Control (and Some Answers). *Social and Personality Psychology Compass, 9*(10), 511-524. doi:10.1111/spc3.12200

Inzlicht, M., Legault, L., & Teper, R. (2014). Exploring the Mechanisms of Self-Control Improvement. *Current Directions in Psychological Science, 23*(4), 302-307. doi:10.1177/0963721414534256

Inzlicht, M., & Schmeichel, B. J. (2012). What Is Ego Depletion? Toward a Mechanistic Revision of the Resource Model of Self-Control. *Perspectives on Psychological Science, 7*(5), 450-463. doi:10.1177/1745691612454134

Kane, M. J., Bleckley, M. K., Conway, A. R., & Engle, R. W. (2001). A controlled-attention view of working-memory capacity. *Journal of Experimental Psychology: General, 130*(2), 169.

Kane, M. J., & Engle, R. W. (2000). Working-memory capacity, proactive interference, and divided attention: Limits on long-term memory retrieval. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 26*(2), 336-358. doi:10.1037/0278-7393.26.2.336

Kane, M. J. E., R. W. (2003). Working-Memory Capacity and the Control of Attention: The Contributions of Goal Neglect, Response Competition, and Task Set to Stroop Interference. *Journal of Experimental Psychology: General, 132*(1), 47-70. doi:10.1037/0096-3445.132.1.47

Kaplan, S., & Berman, M. G. (2010). Directed attention as a common resource for executive functioning and self-regulation. *Perspectives on Psychological Science, 5*(1), 43-57.

Karoly, P. (1993). Mechanisms of self-regulation: A systems view. *Annual Review of Psychology, 44*(1), 23-52.

Kurzban, R. (2010). Does the brain consume additional glucose during self-control tasks? *Evolutionary Psychology, 8*(2), 147470491000800208.

Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity?! *Intelligence, 14*(4), 389-433. doi:http://dx.doi.org/10.1016/S0160-2896(05)80012-1

Lee, N., Chatzisarantis, N., & Hagger, M. S. (2016). Adequacy of the Sequential-Task Paradigm in Evoking Ego-Depletion and How to Improve Detection of Ego-Depleting Phenomena. *Frontiers in Psychology, 7*, 136. doi:10.3389/fpsyg.2016.00136

Lewandowsky, S., & Farrell, S. (2008). Phonological Similarity in serial recall: Constraints on theories of memory. *Journal of Memory and Language, 58*(2), 429-448. doi:10.1016/j.jml.2007.01.005

Lurquin, J., & Miyake, A. (2017). Challenges to Ego-Depletion Research Go Beyond the Replication Crisis: A Need for Tackling the Conceptual Crisis. *Frontiers in Psychology, 8*(568). doi:10.3389/fpsyg.2017.00568

Lurquin, J. H. (2017). *Evaluating the Current Conceptualization and Measurement of Self-Control: a Systematic Review and an Individual Differences Study.* (PhD), University of Colorado Boulder,

Lustig, C. M., C. P., & Hasher, L. (2001). Working Memory Span and the Role of Proactive Interference. *Journal of Experimental Psychology: General, 130*(2), 199-207. doi:10.1037/0096-3445.130.2.199

Mackworth, J. F. (1964). Auditory short-term memory. *Canadian Journal of Psychology/Revue canadienne de psychologie, 18*(4), 292-303. doi:10.1037/h0083309

Madsen, P. L., Hasselbalch, S. G., Hagemann, L. P., Olsen, K. S., Bülow, J., Holm, S., . . . Lassen, N. A. (1995). Persistent resetting of the cerebral oxygen/glucose uptake ratio by brain activation: evidence obtained with the Kety—Schmidt technique. *Journal of Cerebral Blood Flow & Metabolism, 15*(3), 485-491.

Marcora, S. M., Staiano, W., & Manning, V. (2009). Mental fatigue impairs physical performance in humans. *Journal of applied physiology, 106*(3), 857-864.

McEvoy, C. L., & Nelson, D. L. (1982). Category name and instance norms for 106 categories of various sizes. *The American Journal of Psychology*, 581-634.

Miller, H. C. (2012). The Role of Glucose on Executive Control Vigor: Evidence inSupport of Comparative Research. *International Journal of Comparative Psychology, 25*(3).

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive psychology, 41*(1), 49-100.

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

Muraven, M., & Baumeister, R. F. (2000). Self-regulation and depletion of limited resources: Does self-control resemble a muscle? *Psychological Bulletin, 126*(2), 247.

Muraven, M., Baumeister, R. F., & Tice, D. M. (1999). Longitudinal improvement of self-regulation through practice: Building self-control strength through repeated exercise. *The Journal of social psychology, 139*(4), 446-457.

Muraven, M., & Slessareva, E. (2003). Mechanisms of self-control failure: Motivation and limited resources. *Personality and Social Psychology Bulletin, 29*(7), 894-906.

Muraven, M., Tice, D. M., & Baumeister, R. F. (1998). Self-control as a limited resource: Regulatory depletion patterns. *Journal of Personality and Social Psychology, 74*(3), 774-789. doi:10.1037/0022-3514.74.3.774

Myers, L. (2013). Me and Mine. Retrieved from http://www.shortbreadstories.co.uk/story/view/me_and_mine/#axzz5CnMF38sv

Nelson, M. J., & Denny, E. C. (1960). *The Nelson-Denny Reading Test: Forms A & B*: Houghton Mifflin.

Nes, L. S., Roach, A. R., & Segerstrom, S. C. (2009). Executive functions, self-regulation, and chronic pain: a review. *Annals of Behavioral Medicine, 37*(2), 173-183.

Nordgren, L. F., & Chou, E. Y. (2012). A Devil On Each Shoulder: When (and Why) Greater Cognitive Capacity Impairs Self-Control? *Social Psychological and Personality Science*. doi:10.1177/1948550612451622

Norman, D. A., & Shallice, T. (1986). Attention to action. In *Consciousness and self-regulation* (pp. 1-18): Springer.

Oberauer, K., Süß, H.-M., Wilhelm, O., & Sander, N. (2007). Individual differences in working memory capacity and reasoning ability. *Variation in working memory*, 49-75.

Owens, D. S., & Benton, D. (1994). The Impact of Raising Blood Glucose on Reaction Times. *Neuropsychobiology, 30*(2-3), 106-113.

Palmer, C. (2014). Im Falschen Fenster. Retrieved from http://www.shortbreadstories.co.uk/story/view/im_falschen_fenster/#axzz5CnMF38sv

Parrish, A., James, B., Rossettie, M., Smith, T., Otalora-Garcia, A., & Beran, M. (2018). Investigating the depletion effect: Self-control does not waiver in capuchin monkeys. *Animal Behavior and Cognition, 5*, 118-138.

Persson, J., Welsh, K. M., Jonides, J., & Reuter-Lorenz, P. A. (2007). Cognitive fatigue of executive processes: Interaction between interference resolution tasks. *Neuropsychologia, 45*(7), 1571-1579. doi:http://dx.doi.org/10.1016/j.neuropsychologia.2006.12.007

Posner, M. I., Rothbart, M. K., Sheese, B. E., & Tang, Y. (2007). The anterior cingulate gyrus and the mechanism of self-regulation. *Cognitive, Affective, & Behavioral Neuroscience, 7*(4), 391-395. doi:10.3758/cabn.7.4.391

Quigley, B. M. Q., Levitt, A., Derrick, J. L., Testa, M., Houston, R. J., & Leonard, K. E. (2018). Alcohol, Self-regulation, and Partner Physical Aggression: Actor-Partner Effects Over a Three Year Time Frame. *Frontiers in Behavioral Neuroscience, 12*, 130.

Raichle, M. E., & Mintun, M. A. (2006). Brain work and brain imaging. *Annu. Rev. Neurosci., 29*, 449-476.

Rogers, R. D., & Monsell, S. (1995). Costs of a predictible switch between simple cognitive tasks. *Journal of Experimental Psychology: General, 124*(2), 207.

Schmeichel, B. J. (2007). Attention control, memory updating, and emotion regulation temporarily reduce the capacity for executive control. *Journal of Experimental Psychology: General, 136*(2), 241.

Schmeichel, B. J., & Vohs, K. (2009). Self-affirmation and self-control: Affirming core values counteracts ego depletion. *Journal of Personality and Social Psychology, 96*(4), 770-782. doi:10.1037/a0014635

Schmeichel, B. J., Vohs, K. D., & Baumeister, R. F. (2003). Intellectual performance and ego depletion: role of the self in logical reasoning and other information processing. *Journal of Personality and Social Psychology, 85*(1), 33.

Scholey, A. B., Harper, S., & Kennedy, D. O. (2001). Cognitive demand and blood glucose. *Physiology & Behavior, 73*(4), 585-592.

Sham, N. (2013). Picture Speaking... Retrieved from http://www.shortbreadstories.co.uk/story/view/picture_speaking/#axzz5CnMF38sv

Shamosh, N. A., & Gray, J. R. (2007). The relation between fluid intelligence and self-regulatory depletion. *Cognition and Emotion, 21*(8), 1833-1843.

Shapiro, S. I., & Palermo, D. S. (1970). Conceptual organization and class membership: Normative data for representations of 100 categories. In *Psychonomic Monograph Supplements* (Vol. 3, pp. 107-127).

Shipstead, Z., Redick, T. S., & Engle, R. W. (2012). Is working memory training effective? *Psychological Bulletin, 138*(4), 628.

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-Positive Psychology: Undisclosed Flexibility in Data Collection and Analysis Allows Presenting Anything as Significant. *Psychological Science, 22*(11), 1359-1366. doi:10.1177/0956797611417632

Sripada, C., Kessler, D., & Jonides, J. (2014). Methylphenidate Blocks Effort-Induced Depletion of Regulatory Control in Healthy Volunteers. *Psychological Science*. doi:10.1177/0956797614526415

Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18*(6), 643-662. doi:10.1037/h0054651

Tangney, J. P., Baumeister, R. F., & Boone, A. L. (2004). High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. *Journal of personality, 72*(2), 271-324.

Tehan, G., Arber, M., & Tolan, G. A. (2018). Working Memory Capacity as a Determinant of Proactive Interference and Auditory Distraction. *Journal of Cognition, 1*(1).

Tehan, G., Hendry, L., & Kocinski, D. (2001). Word length and phonological similarity effects in simple, complex, and delayed serial recall tasks: Implications for working memory. *Memory, 9*(4-6), 333-348.

Tehan, G., & Humphreys, M. S. (1995). Transient phonemic codes and immunity to proactive interference. *Memory & Cognition, 23*(2), 181-191.

Tehan, G., & Humphreys, M. S. (1998). Creating proactive interference in immediate recall: Building a DOG from a DART, a MOP, and a FIG. *Memory & Cognition, 26*(3), 477-489.

Tice, D. M., Baumeister, R. F., Shmueli, D., & Muraven, M. (2007). Restoring the self: Positive affect helps improve self-regulation following ego depletion. *Journal of Experimental Social Psychology, 43*(3), 379-384.

Tolan, A. G., & Tehan, G. (2002). Testing Feature Interaction: Between-stream Irrelevant Speech Effects in Immediate Recall. *Journal of Memory and Language, 46*(3), 562-585. doi:10.1006/jmla.2001.2820

Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language, 28*(2), 127-154. doi:https://doi.org/10.1016/0749-596X(89)90040-5

Tyler, J. M., & Burns, K. C. (2008). After Depletion: The Replenishment of the Self's Regulatory Resources. *Self and Identity, 7*(3), 305-321. doi:10.1080/15298860701799997

Unsworth, N., & Engle, R. W. (2006). Simple and complex memory spans and their relation to fluid abilities: Evidence from list-length effects. *Journal of Memory and Language, 54*(1), 68-80.

Unsworth, N., & Engle, R. W. (2007). On the division of short-term and working memory: an examination of simple and complex span and their relation to higher order abilities. *Psychological Bulletin, 133*(6), 1038.

Unsworth, N., Redick, T. S., Spillers, G. J., & Brewer, G. A. (2012). Variation in working memory capacity and cognitive control: Goal maintenance and microadjustments of control. *Quarterly Journal of Experimental Psychology, 65*(2), 326-355.

Wallace, H. M., & Baumeister, R. F. (2002). The Effects of Success versus Failure Feedback on Further Self-Control. *Self and Identity, 1*(1), 35-41. doi:10.1080/152988602317232786

Wan, E. W., & Sternthal, B. (2008). Regulating the effects of depletion through monitoring. *Personality and Social Psychology Bulletin, 34*(1), 32-46.

Webb, T. L., & Sheeran, P. (2003). Can implementation intentions help to overcome ego-depletion? *Journal of Experimental Social Psychology, 39*(3), 279-286. doi:https://doi.org/10.1016/S0022-1031(02)00527-9

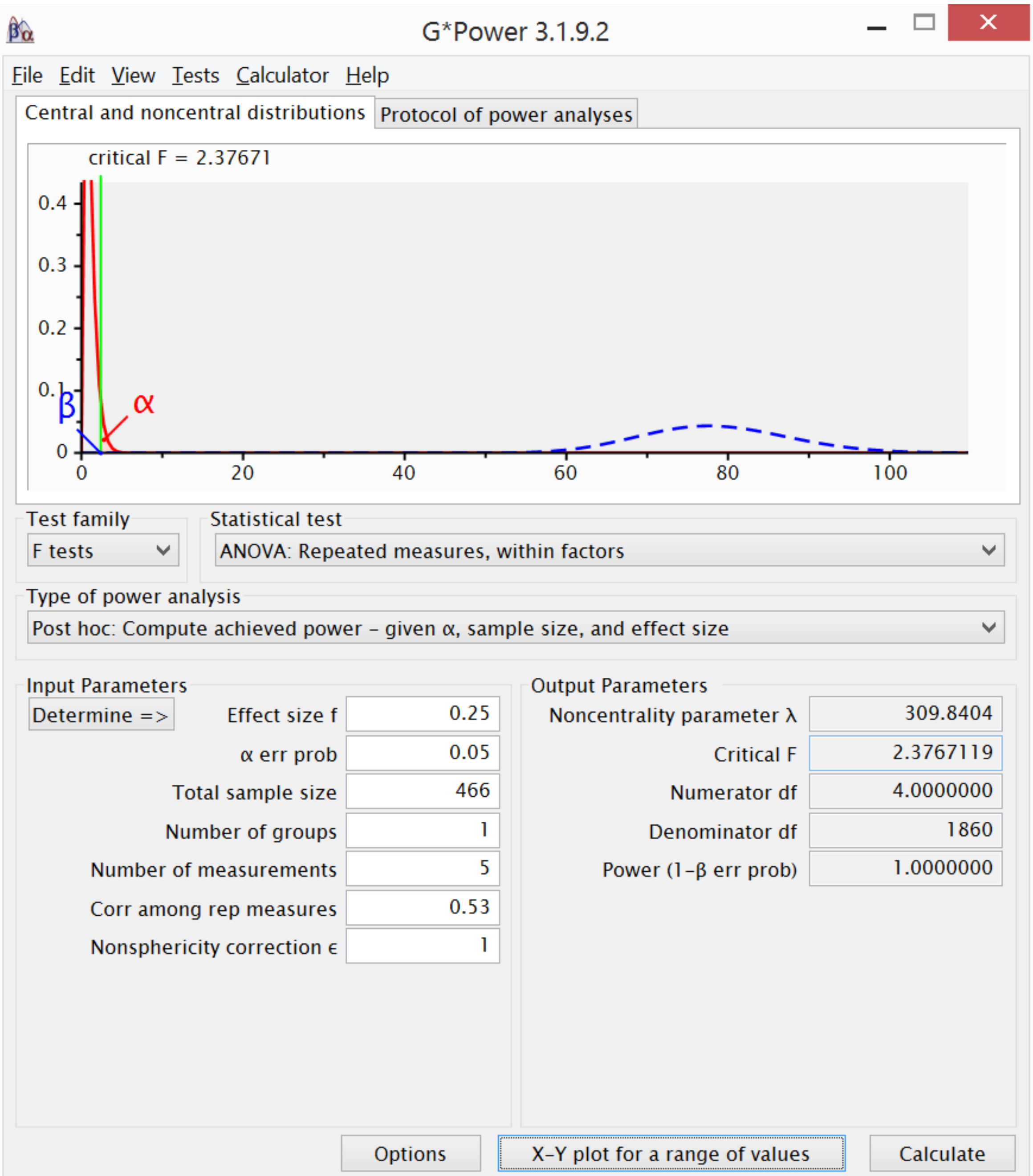
Wicherts, J. M., Veldkamp, C. L. S., Augusteijn, H. E. M., Bakker, M., van Aert, R. C. M., & van Assen, M. A. L. M. (2016). Degrees of Freedom in Planning, Running, Analyzing, and Reporting Psychological Studies: A Checklist to Avoid p-Hacking. *Frontiers in Psychology, 7*, 1832. doi:10.3389/fpsyg.2016.01832

Wilhelm, O., Hildebrandt, A., & Oberauer, K. (2013). What is working memory capacity, and how can we measure it? *Frontiers in Psychology, 4*.

Xu, X., Demos, K. E., Leahey, T. M., Hart, C. N., Trautvetter, J., Coward, P., . . . Wing, R. R. (2014). Failure to Replicate Depletion of Self-Control. *PLoS ONE, 9*(10), e109950. doi:10.1371/journal.pone.0109950

Appendices

Appendix A. The analysis in G*Power software (Faul et al., 2009) conducted to assess effect size for the total sample size in Study 1.



Appendix B. Participant information sheet and consent forms.

University of Southern Queensland

Participant Information for USQ Research Project

Project Details

Title of Project:	Testing the Conservation Hypothesis in Ego Depletion
Human Research Ethics Approval Number:	**H16REA031**

Research Team Contact Details

Principal Investigator Details

Madeleine Arber
Email: Madeleine.Arber@usq.edu.au
Mobile: 0411726657

Other Investigator/Supervisor Details

Prof Gerry Tehan
Email: Gerry.Tehan@usq.edu.au
Mobile: 0421025256

Description

The purpose of this project is to investigate how fatigue might affect performance on certain aspects of cognitive functioning.
This project has put together a series of tasks that will test your executive functioning under different conditions.

Participation

Your participation will involve you completing several cognitive performance tasks. In all, it will take approximately 45 minutes of your time.
The experiment will include a cognitive search and identification task (crossing out letters in a story), and a short-term memory task where you will have to remember four items on each trial. Neither of the tasks is particularly difficult although we will expect you to make some errors on both tasks. We will then get some demographic information from you.
Your participation in this project is entirely voluntary. If you do not wish to take part you are not obliged to. If you decide to take part and later change your mind, you are free to withdraw from the project at any stage. You may also request that any data collected about you be destroyed, however because we will not be collecting any personal identifying information this may be difficult in some cases. If you do wish to withdraw from this project or withdraw data collected about you, please contact the Research Team (contact details at the top of this form). You are encouraged to contact the principle investigator if you have any concerns or questions or wish to discuss your decision to withdraw.
Your decision whether you take part, do not take part, or to take part and then withdraw, will in no way impact your current or future relationship with the researchers involved or with the University of Southern Queensland.

Expected Benefits

It is expected that this project will not directly benefit you. However, by improving our understanding of the variables under study, it may benefit our understanding of the problems people face in regulating their behaviour.

Risks

There are minimal risks associated with your participation in this project. These include

1) Frustration and boredom associated with the nature of the tasks may impact your mood or wellbeing adversely.
2) Beliefs that performance tasks are an indication of cognitive ability – which they are not.

Sometimes thinking about the sorts of issues raised in the project create some uncomfortable or distressing feelings. If you need to talk to someone about this immediately please contact centacare on 1300 236 822 or beyondblue on 1300 22 4636. You may also wish to consider consulting your General Practitioner (GP) for additional support.

Privacy and Confidentiality

All performance data will be treated confidentially unless required by law.
Any data collected as a part of this project will be stored securely as per University of Southern Queensland's Research Data Management policy. The data collected in this experiment will also contribute to verifying the same experiment that was conducted in 2015.
Any data provided as part of this study can be used in future studies. It should be noted that in the future it may be impossible for us to locate your data should you wish to have it removed from the dataset. Non-identifiable data will be managed as per the requirements of the National Statement (Chapter 3.2) and external requests to access the non-identifiable data set will be assessed by the Principal investigator and distributed where appropriate.

Consent to Participate

We would like to ask you to sign a written consent form (enclosed) to confirm your agreement to participate in this project. Please return your signed consent form to a member of the Research Team prior to participating in your interview.

Questions or Further Information about the Project

Please refer to the Research Team Contact Details at the top of the form to have any questions answered or to request further information about this project.
To receive a summary of the results please put a contact email on the consent form before returning it.

Concerns or Complaints Regarding the Conduct of the Project

If you have any concerns or complaints about the ethical conduct of the project you may contact the University of Southern Queensland Ethics Coordinator on (07) 4631 2690 or email ethics@usq.edu.au. The Ethics Coordinator is not connected with the research project and can facilitate a resolution to your concern in an unbiased manner.

Thank you for taking the time to help with this research project. Please keep this sheet for your information.

University of Southern Queensland

Consent Form for USQ Research Experiment

Project Details

Title of Project:	Testing the Conservation Hypothesis in Ego Depletion
Human Research Ethics Approval Number:	**H16REA031**

Research Team Contact Details

Principal Investigator Details

Madeleine Arber
Email: Madeleine.Arber@usq.edu.au
Mobile: 0411726657

Other Investigator/Supervisor Details

Prof Gerry Tehan
Email: Gerry.Tehan@usq.edu.au
Mobile: 0421025256

Statement of Consent

By signing below, you are indicating that you:

1. Have read and understood the information document regarding this project.
2. Have had any questions answered to your satisfaction.
3. Understand that if you have any additional questions you can contact the research team.
4. Understand that you are free to withdraw at any time, without comment or penalty.
5. Understand that you can contact the University of Southern Queensland Ethics Coordinator on (07) 4631 2690 or email ethics@usq.edu.au if you do have any concern or complaint about the ethical conduct of this project.
6. Understand that the data for this project will be used to verify a previous experiment conducted in 2015 and may be used in future experiments looking into these variables. Non-identifiable data will be managed as per the requirements of the National Statement (Chapter 3.2) and external requests to access the non-identifiable data set will be assessed by the Principal investigator and distributed where appropriate
7. Are over 18 years of age?
8. Agree to participate in the project.

Participant Name	
Participant Signature	
Date	

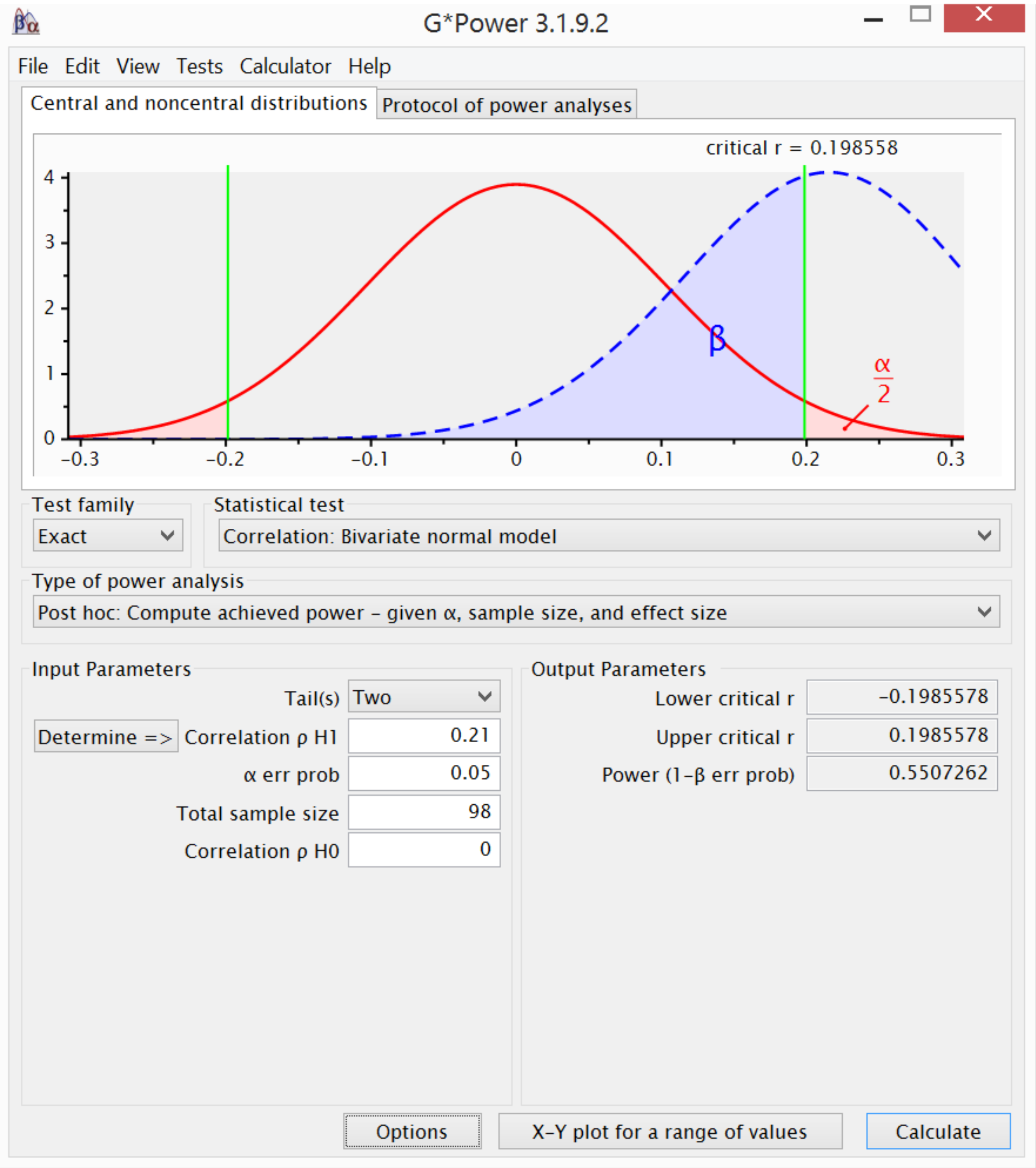
If you would like to receive a summary of results once the study is complete, please include your email below (this will be kept strictly confidential and will not be linked or stored with your study data in any way).

Email	

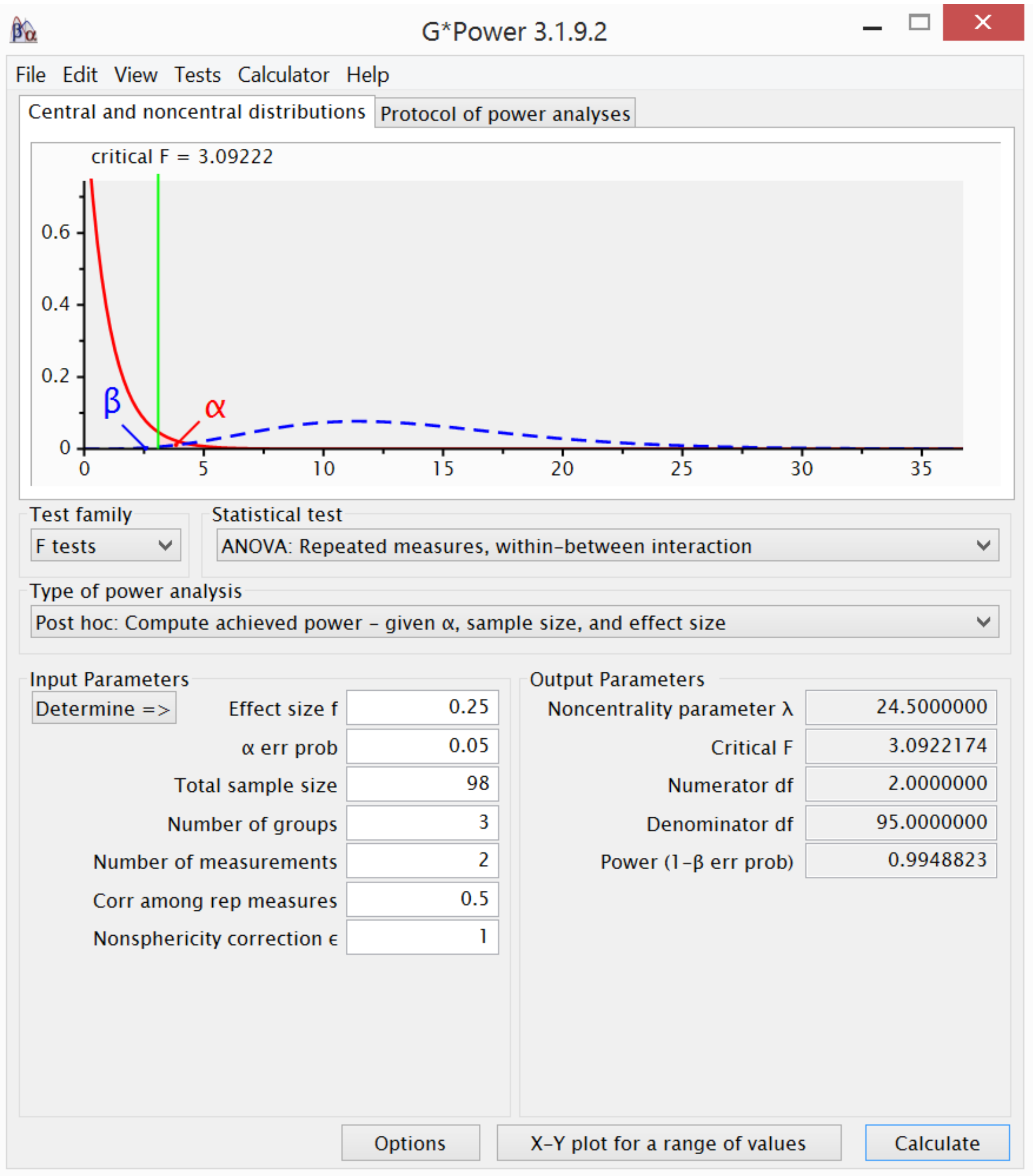
Code:

The first letter or your middle name: _________ (If you have no middle name put in an N)
The month and year of your birth: ___________ (e.g. April, 1988)
The initials of your mothers maiden name: ________ (if you do not know the answer to this put in XY)

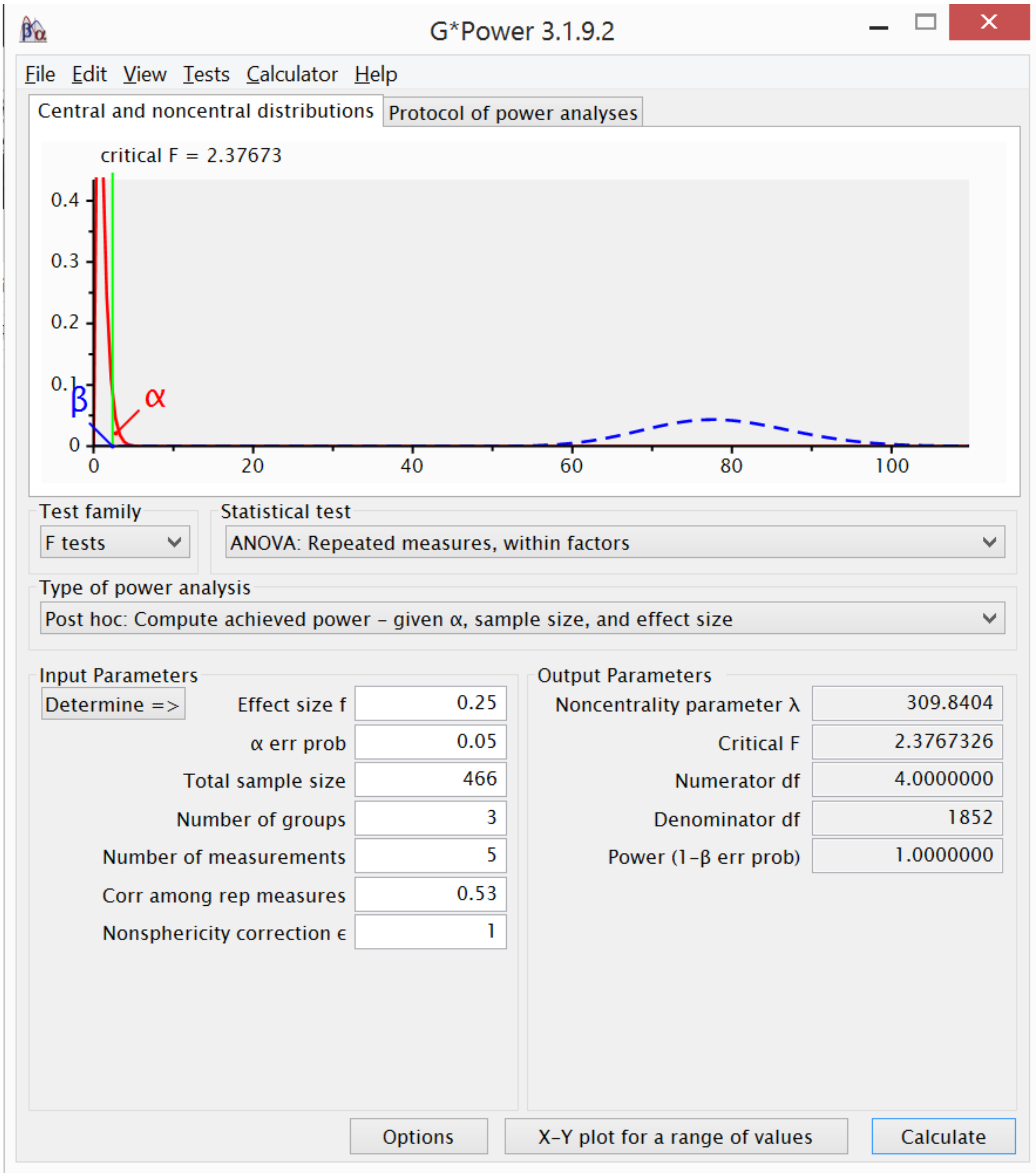
Appendix C. The analysis in G*Power software (Faul et al., 2009) conducted to assess effect size of the smallest sample size achieved in Experiment 1.



Appendix D. The analysis in G*Power software (Faul et al., 2009) conducted to assess effect size of the smallest sample size achieved in Experiment 5.



Appendix E. The analysis in G*Power software (Faul et al., 2009) conducted to assess effect size for the total sample size in the post-hoc analysis in Study 3.



Appendix F. The analysis in G*Power software (Faul et al., 2009) conducted to assess effect size for the smallest sample size in the post-hoc analysis in Study 3.