

Evaluating the Strength Model and Willpower Beliefs Accounts of the Ego-
Depletion Effect

By

Parker Benson

Psychology, University of Colorado Boulder

Defense Date: April 9, 2018

Thesis Advisor: Akira Miyake, Psychology

Defense Committee: Michael Saddoris, Psychology and Neuroscience

Andrew Chapman, Philosophy

Abstract

The ego-depletion effect was originally found to be highly robust and has often been explained through the strength model of self-control. This model states that exerting self-control depletes a limited pool of domain-general resources and makes subsequent self-control attempts less effective. However, some researchers proposed an alternative account based on the notion of people's implicit beliefs about willpower; specifically, only those individuals who hold the belief that willpower is limited are susceptible to the ego-depletion effect. This study evaluated these two theoretical accounts of the ego depletion effect while trying to address some of the methodological and conceptual limitations present in most prior research. The current study used the sequential-task paradigm on two 20-minute tasks (N-back and SART) that have been shown to overlap in their self-control task demands. N-back task differed in their self-control demands where the easy 1-back version was administered in the non-depletion condition and the more challenging 2-back version in the depletion condition. It was hypothesized that if the strength model is valid, there should be more inhibitory errors and attentional lapses in the second SART task. Further it was hypothesized that if the mindset theory is valid, the ego-depletion effect should be moderated by one's belief about willpower. The preliminary results based on 98 subjects ($n = 49$ in each condition) showed that there were no significant differences in SART performance or mind wandering between conditions. Further, no significant interactions were found between beliefs about willpower and condition. These findings do not only challenge both the strength model and willpower belief accounts of the ego-depletion effect but more fundamentally add to the growing body of research that questions the robustness and replicability of the ego-depletion effect.

Evaluating the Strength Model and Willpower Beliefs Accounts of the Ego-Depletion Effect

There has been a wide variety of research done on different aspects self-control. Self-control plays into every day decisions for situations like dieting and can be relevant in important areas like work or school. Self-control is an individual's ability to inhibit unproductive tendencies or impulses and choose something less desirable to achieve a long-term goal (Baumeister et al., 1998).

Research examining self-control is often focused on the phenomenon of ego-depletion (Baumeister et al., 1998). This phenomenon is the idea that when an individual exerts self-control on a certain task or choice, their ability to continue to exert self-control is reduced in future instances. As will be reviewed later, findings for this theory have been varied. Early research has shown that ego-depletion is real and significant (Baumeister et al., 1998), but other more recent large-scale replications and various studies on the topic have concluded both non-significant or null results (Hagger et al., 2016). The present research aims to address some methodological and conceptual issues in the field of ego-depletion research in hopes of assessing how changes in research design such as task length and non-domain analogous tasks influence the ego-depletion effect.

Ego-Depletion Effect and the Strength Model of Self-Control

Ego-Depletion

Baumeister, Bratslavsky, Muraven, and Tice's (1998) original finding across 4 different experiments established the ego-depletion effect as a psychological construct. It was originally established as a highly robust effect with a moderate effect size. Through their cookie task, attitude-relevant actions task, emotion regulation task, and self-regulation task they found that

there was some “psychological cost” to doing certain actions. For example, in the study done by Baumeister et al., (1998) it was found that individuals who forced themselves to eat radishes in the presence of tempting cookies quit more readily when doing an unsolvable puzzle than the participants who did not previously exert self-control. It was at this point that it was suggested that the capacity for self-control is limited and can be diminished when doing related or unrelated acts that pull from the same resource (Baumeister et al., 1998).

With the presupposition of the effect of ego depletion, Schmeichel, Vohs and Baumeister (2003) further researched the concept to see if logical reasoning and information processing, actions that require active guidance by the self, deplete this resource more readily than simple mental activities. They found that those who had to regulate attention or emotion performed worse on tasks that required logic and reasoning, cognitive extrapolation, and thoughtful reading comprehension (Schmeichel et al., 2003). It was found in further research by the same authors that blood glucose also plays a role in the effect of ego-depletion (Baumeister Vohs, & Tice, 2007). They concluded that there is a biological aspect to self-control that is beyond just mental capacity. Blood glucose can manipulate the ego-depletion effect in addition to mental tasks (Baumeister et al., 2007).

With this amount of initial support, a variety of researchers found more support for the effect. For example, Webb and Sheeran (2003) tested whether a variety of self-regulatory strategies were effective in reducing the extent of the ego-depletion effect, to which they found supporting evidence for the both ego-depletion and the ability for self-regulatory strategies to mitigate it. Research looking at autonomy also found support for the theory as well as noticed a mediating effect of choice (Moller, Deci, & Ryan, 2006). Mauraven, Shmueli and Burkley (2006) also found similar results to Moller et al; Participants who were expecting to exert self-

control in future situations performed worse on the present self-control task and better at the future tasks (Muraven et al., 2006).

All of the research up to this point provided a basis for a large meta-analysis, which assessed the effect size and robustness of the ego-depletion effect in previous research. This meta-analysis, based on 83 studies, found that the standardized mean difference for the ego-depletion on self-control had a Cohen's d of .62 (Hagger et al., 2010). This study did not rule out moderators or identify specific mechanisms for the ego-depletion effect but concluded that previous research as a whole showed a significant, medium to large effect size for the support of ego-depletion (Hagger et al., 2010).

Strength Model of Self-Control

With the supporting research thus far regarding the ego-depletion effect, researchers have proposed the strength model of self-control to explain those results. According to this theory, ego-depletion is the process of drawing from an internal resource that regulates self-control behaviors and can be temporarily depleted after extensive use but can be strengthened with practice—like using a muscle (Muraven, Shmueli, & Burkely, 2006). As such, self-control can be made stronger through repeated use and conserved by active moderation of current self-control usage. This intuitive model is used to explain how the phenomenon of ego-depletion occurs.

A major assumption behind this theory is that self-control draws on a single pool of the same domain-general resources regardless of the activity being done. To test this, researchers tried various combinations of very different tasks to intentionally test this theory. One such study was one that was done by Muraven, Baumeister, and Tice (1999), who looked at improving self-regulation through practice. They required individuals to practice different self-control exercises such as improving posture and compared them to a control group on an outcome hand-grip

exercise task after doing a thought-suppression activity (Muraven et al., 1999). Other studies that used widely different tasks were those done by Baumeister, Bratslavsky, Muraven, and Tice (1998) that used a cookie temptation resistance measure and comparing that to things like emotion and self-regulation tasks. Because such rather different combinations of tasks led to ego-depletion effects, researchers suggested that the underlying self-control resources are domain-general.

Challenges to the Ego-Depletion Effect and the Strength Model

Replication Crisis

Until recently, ego-depletion seems to be a robust phenomenon, and there has a lot of support for the strength model. However, in the last 7-8 years, there has been an accumulation of evidence that the effect may not be as robust as initially seemed. Although Hagger et al., (2010) 83-study meta-analysis originally supported the ego depletion effect, the validity of the estimated effect size ($d=.62$) has been called into question. Namely, Carter and McCullough (2014) conducted a meta-analysis using a different method and challenged the original suggestion that the effect size for ego-depletion is medium to high. The researchers took into account small-study effects and publication bias and found that the true effect size given the collected studies is close to, if not, zero (Carter & McCullough, 2014). This put a spotlight on many older studies and showed some of the more major issues regarding research in ego-depletion.

One example of a failed replication was research done by Xu et al., (2014) which has shown some opposite findings regarding the strength model of ego-depletion. Across 4 different studies with varying samples, they used the sequential-task paradigm as a measure of ego depletion. This sequential-task paradigm is one where a participant either completes a depleting or non-depleting task followed by an outcome self-control task. They found null results in

performance changes across different outcome self-control tasks after being required to exert self-control in the initial task (Xu et al., 2014). They suggest that with certain tasks and methodological approaches, the ego-depletion effect is more limited than originally thought (Xu et al., 2014).

Another influential replication attempts came from a pre-registered study by Lurquin et al., (2016). Lurquin and his colleagues used the sequential-task paradigm to test for the effect of ego-depletion. With a large sample size and modified task characteristics they were able to confront many of the issues brought up by Carter and McCullough (2014). It was found that despite modified tasks and increased sample size, there was no difference between the depleting and non-depleting condition when compared on the outcome self-control task (Lurquin et al., 2016). Additionally, no moderating effects were found for the ego-depletion effect, which lead researchers conclude that the ego-depletion is not as robust as previously reported (Luquin et al., 2016).

From this, there was a further need to assess the validity of some of the original studies in the field. One of the biggest attempts to do so is the Hagger et al., (2016) study. This 23-lab, pre-registered replication and meta-analysis used the sequential-task paradigm. The methodology was that of the Sripada et al., (2014) study which used a computerized E-cross task for the manipulation and a multi-source interference task (MIST) as the outcome task. Across 23 labs, Hagger et al. found that the Cohen's d effect size for the meta-analysis was less than .1 and the confidence interval contained zero (Hagger et al., 2016). Thus, the large-scale replication which included every lab's results concluded that if there is an ego-depletion effect, it is incredibly small.

Methodological Concerns

One reason for this replication crisis might have to do with several major methodological limitations of the original research that demonstrated significant ego-depletion effects. One such methodological concern has to do with task duration. Typical time on task for the depleting conditions were originally around 5 minutes and the outcome measures were also quite short. With such short tasks it seems unlikely that individuals are completely “depleted” or fatigued after completing the first task. As such, future research should aim to increase task duration to ensure that the first task is actually fatiguing and requiring of some amount of mental energy.

Another methodological problem is smaller sample sizes, which make effect size estimates to be imprecise and error-prone. With lower sample sizes, minor differences in participant responses will create major impacts on outcome measures. Combined with inherent publication biases favoring statistically significant results, some research could be exploiting this by stopping data collection at low participant numbers when it showed a significant effect. To adequately assess the effect size of ego-depletion and the strength model, higher participant numbers will be needed. This will not only allow for a more accurate picture of effect size but allow for moderator variables to be tested for.

Lastly, there is a notable lack of manipulation checks in many, if not most, studies. This is problematic in that there is no way of showing that the two tasks, depleting and non-depleting, are adequately different from each other. Measures of perceived fatigue, effort, difficulty, etc. helps researchers understand the task characteristics and allows for researchers to test exactly why there may be individual differences in performance. Additionally, it gives some further insight in to what might be causing the ego-depletion effect. If participants are highly motivated for one condition over the other, or one condition reports higher levels of difficulty and fatigue,

one may be able to make a conclusion about general fatigue causing changes in outcome measures rather than other mental resources.

Conceptual Crisis

Although the methodological concerns can be prevalent, there are notable conceptual issues that affect research as well, which Lurquin and Miyake (2017) called a “conceptual crisis” in ego-depletion research. The first issue raised by Lurquin and Miyake (2017) is the lack of clear operational definition of self-control. While the strength model is often referred to, definitions of self-control are often broad and ambiguous. Additionally, many tasks haven’t been validated as effective measures of self-control (Lurquin & Miyake, 2017). This is problematic in that many researchers are using previously used, or even novel tasks for measuring self-control solely based off of intuition. These tasks can sometimes be wildly different from each other and not clearly be pulling from the same resources. Lastly, Lurquin and Miyake (2017) mention that although the strength model is prevalent, the theories surrounding ego-depletion are too vague and are often stretched to fit a certain finding. This makes each theory lacking in testability and falsifiability.

Although these issues should be addressed in further research, some researchers question whether the resource notion of self-control is even the best explanation of the results in past experiments. Another explanation of results has to do with mindset about self-control, where beliefs about willpower can be seen as a predictor of performance rather than mental resource depletion. To test this alternate view, Job, Dweck, and Walton (2010) measured participants’ belief about willpower and the extent to which individuals had control over their self-regulation. It was found that individuals who did not believe self-control was a limited resource did not exhibit any signs of ego-depletion after a depleting experience. The researchers concluded that

reduction in self-control, or the ego-depletion effect, could likely be due to people beliefs about their control of their willpower, and not some actual mental resource (Job et al., 2010). This is significant as it calls into question the legitimacy of the strength model and the common self-control resource that was originally studied.

Further research testing the mindset theory was that done by Savani and Job (2017), where they tested Indian cultures against western cultures in their beliefs about willpower and how that affected the phenomenon of ego-depletion. They found that whereas the western populations exhibited an ego-depletion effect, individuals from the Indian culture exhibited reverse ego-depletion effects (Savani & Job, 2017). This is because people from Indian cultures believe they would be more energized after exerting self-control and hence likely to be able to exert even more self-control on future tasks due to mindset and culturally held views of willpower (Savani & Job, 2017).

Current Study

The current study seeks to address many of the methodological and conceptual issues regarding ego depletion. First and foremost, the present research will be a large-scale study with validated measures. The intended number of participants when finished with data collection is 200 usable participants with 100 in each condition. For the sake of this Honors Thesis, a cutoff date of participants was made to ensure adequate time for analyses and write-up. The goal for this thesis was 50 per condition. Despite being half of the intended sample size, this is higher than most self-control research. Additionally, all intended analyses, conditions, and procedures are in line with previous research.

Figure 1 demonstrates the overall procedure of this study. For this thesis I chose the N-Back working memory task as the manipulated depleting (2-back) or non-depleting (1-back) task

and the Selective Attention to Response Task (SART) (McVay & Kane, 2009) as the dependent measure. The first task, N-Back, is regarded as a working memory capacity or updating task. This task requires participants to keep a short list of letters in their head and determine if the current letter they are seeing matches a previously seen letter. Different levels of this task are more difficult than others. This was chosen for the manipulated depleting and non-depleting task due to the variability in difficulty between the 1 and 2-back versions of the task. The 2-back task, which required one to keep a longer list in their head is depleting because of the notable difficulty in needing to consistently keep a 2-letter list of previous letters. Because the 1-back task requires matching the current letter to the previous, the task is fairly simple and non-depleting.

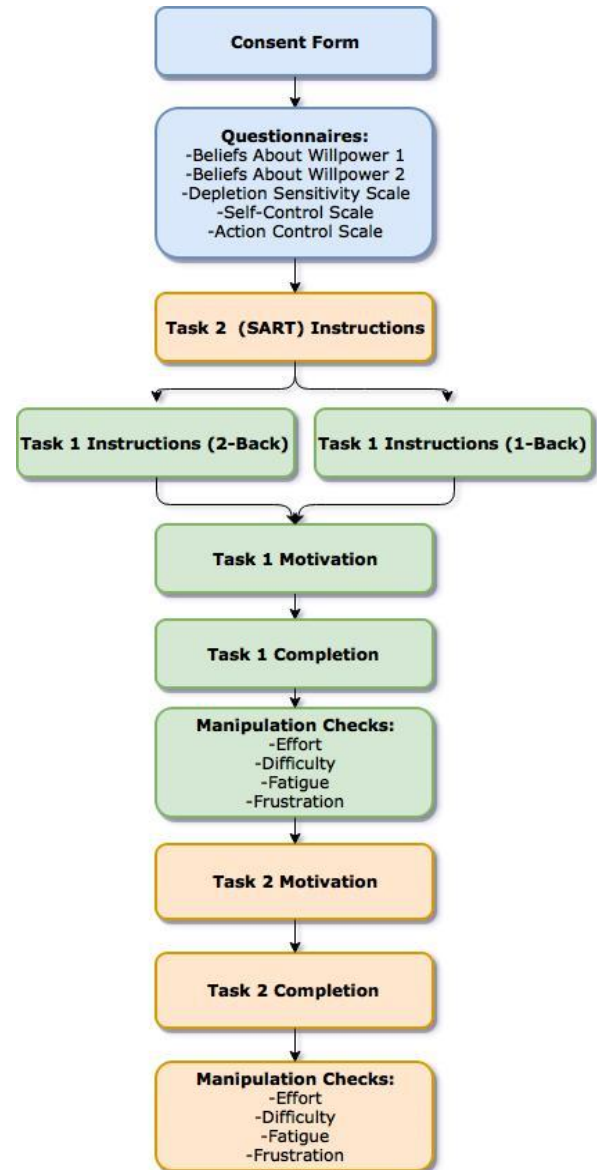


Figure 1. Task layout and procedure.

The second task for this study was chosen due to its inhibitory nature. The task required participants to press the space bar if they see a word that appears on the screen from a specific category (which occurred more often), and to actively *not* press the space bar if the word is from the incorrect category. While every trial of this task is not a measure of inhibition, it is measured nonetheless by assessing participant ability to overwrite a habituated response of pressing the

space bar. Throughout the task there were randomly inserted mind-wandering probes which also assessed participant tendency to stay mentally on the task at hand.

These tasks were chosen because they are in different realms of executive function but have been significantly correlated in previous research. The previous research that showed this correlation was that done by McVay and Kane (2009) which used multiple SART versions to test the construct of working memory capacity and mind wandering (McVay & Kane, 2009). They found that SART demands did not affect mind wandering rates, but rather found a complex interaction where mind-wandering acted as a partial mediating effect of the working memory capacity and SART performance relation (McVay & Kane, 2009). They concluded that working memory related differences in performance were due in part to the “control of conscious thought.” (McVay & Kane, 2009). These findings show that there is likely a significant overlap in the N-Back task (working memory capacity) and the SART (inhibitory control capability).

Moreover, the N-back and SART tasks were designed such that they were both 20 minutes in length. This was ensured by tweaking individual trial time such that the average time to complete all trials was approximately the same. This extended time was done to ensure that the depletion condition was done for a sufficient amount of time and there were enough inhibitory control trials to accurately test the effect. Task instruction order was given in the demonstrated order to ensure that participants had minimal break between the first and second task, to not “regenerate” and self-control resources when not on task.

As shown in Figure 1, before conducting the tasks, participants completed a variety of questionnaires. These questionnaires were chosen to evaluate the mindset theory of ego-depletion. The primary questionnaires were two general Beliefs About Willpower Scales. One measured participant’s beliefs about their ability to exercise self-control (Job et al. 2010) and the

other measured whether people believed that exercising willpower on one task was energizing (Savani & Job, 2017). Also included were three other measures to examine if they are correlated with outcome tasks or the beliefs about willpower scales. These other measures were the Depletion Sensitivity Scale (Salmon et al., 2014), Action Control Scale (Kuhl & Kazén, 1994), and a general Self-Control Scale (Baumeister et al., 1998). These all give a different perspective self-control and allow a better picture of how willpower is exerted in an individual.

Further measures included in the study were manipulation checks for each task. Before each task was presented, motivation for completion was assessed, and after each task there were questions regarding difficulty, effort, fatigue, and frustration. These were implemented to verify that there was a significant difference in task characteristics on important facets like depletion capacity and difficulty to inhibit responses for the depleting (2-back) and non-depleting (1-back) tasks.

Hypotheses and Predictions

Through experimental design and the SART characteristics, there are 4 major variables being measured: (i) Motivation to do well on the SART task, (ii) number of correct button presses for the correct category, (iii) number of incorrect button presses for the incorrect category (inhibitory control errors), and (iv) average mind wandering rates throughout the task. My hypothesis is based on these measures as well as the questionnaires presented before the first task.

Hypothesis 1: Testing the Ego-Depletion Effect

It is hypothesized that, in light of the current design and previous research, if the ego-depletion effect is real and the strength model of self-control is valid, then there should be an expected increase in (iii) inhibitory control errors and (iv) attentional lapses during the outcome

task. This is expected due to the habituated responses making it more difficult to not respond to the incorrect category. Additionally, if a participant is depleted, it is often asserted that it is due to fatigue which would increase mind wandering rates. Due to (ii) the correct button presses being a habituated response and (i) motivation being recorded before the task happens, those are expected to be equal and unaffected by condition differences. Those in the depleted condition (2-back) though should exhibit more inhibitory control errors and levels of mind-wandering due to fatigue than the non-depleting (1-back) condition.

Hypothesis 2: Testing the Effect of Willpower Mindset

To address the mindset view of ego-depletion as well, questionnaire measures will be assessed. It is hypothesized that if the ego-depletion effect is due to mindset, as proposed by previous research, then there should be an interaction between the questionnaire measures regarding beliefs about willpower and the condition that the participant is in. The other questionnaire measures (Action Control Scale, Depletion Sensitivity Scale, and the Self-Control Scale) which may supplement the beliefs about willpower scales, will not necessarily show an interaction with condition if the mindset theory is true.

Methods

Participants

In total, data was collected from 116 participants, 98 of which had usable data. These students were all enrolled in the General Psychology 1001 course offered through the University of Colorado Boulder's psychology program. These students were predominantly freshman with ages ranging from under 18 to over 25 ($M=19.1$, $SD=2.0$). Of the usable participants, 34 identified as male, and 64 identified as female. Participants signed up voluntarily through the

Department of Psychology-CU Psychology 1001 Subject Pool website. The participants were given course credit for participation in the study.

Design

This study utilized an experimental design, where ability to exert self-control on an outcome task was measured on a few variables in relation to previously completed depleting or non-depleting task. There were additional questionnaires administered prior to the first task, which were analyzed as a correlation to other questionnaires and for interactions with condition on performance. The independent variable that was manipulated in this experiment was the depleting or non-depleting version of the N-back task. The dependent variable was consistent across all participant which was the SART task. The correlational analyses done were on the participant ratings on the questionnaires of beliefs about willpower, depletion sensitivity, perceived self-control and action control on outcome task performance. The task performance data and manipulation checks were analyzed by using an independent samples t-test. The correlates were analyzed through multiple regression to test the influence of each variable on the SART task.

Measures and Tasks

Questionnaires. The questionnaires were implemented through Qualtrics. There were two different scales that measured beliefs about willpower. The first (Job et al., 2010) was measured on a 10-point Likert scale, with 1 indicating “strongly disagree” and 10 indicating “strongly agree” responses to willpower statements. The first beliefs about willpower measure assess individual’s idea regarding their ability to exert self-control in a variety of situations. An example belief about willpower statement read: “After a strenuous mental activity your energy is depleted, and you must rest to get it refueled again.”

The second belief about willpower questionnaire (Savani & Job, 2017) assessed whether individuals believe that exerting self-control is a tiring or an energizing process. This questionnaire required participants to indicate on the same 10-point Likert scale the extent to which they agree with the sentiment in a given question. An example of one of these questions would read: “Imagine you are working on a very difficult task that requires a lot of concentration for 1 hr. Do you believe that immediately after this, you would need a break before you can work on another difficult task that also requires a lot of concentration, or would you be able to concentrate on another difficult task right away without any break?”

The following scales were implemented to supplement the belief about willpower scales and give more specific insight regarding individual perception of ego-depletion. The Depletion Sensitivity Scale (Salmon et al., 2014) was also on a 10-point Likert scale with 1 indicating “strongly disagree” and 10 indicating “strongly agree.” This scale measures the degree to which an individual believes they are sensitive to not being able to exert self-control following a difficult task. An example of a depletion sensitivity scale question would be “it is hard for me to persist with a difficult task.”

The Self-Control Scale (Baumeister et al., 1998) was measured with a 10-point Likert scale as well, with 1 indicating “strongly disagree” and 10 indicating “strongly agree.” An example of this would read “I am good at resisting temptation.” Although this is similar to the beliefs about willpower scale, this scale was also validated to measure self-control and provides a slightly different interpretation of participants self-control behavior.

The last scale was the Action Control Scale (Kuhl & Kazén, 1994) which required a binary response from participants, who answered A or B based off of agreement with the associated statement. An example of the action control scale question would be: “When I know I

must finish something soon: A) I have to push myself to get started or B) I find it easy to get it done and over with.” This scale further assessed participants ability to initiate action in difficult situations and whether they have a general tendency to struggle exerting self-control or if they have a higher propensity to control their actions. On all Likert scales, responses were averaged to create an overall individual value for the participant. On the action control scale, number of A and B responses were totaled and compared.

Computer tasks. The two tasks for the study were done through the PsychoPy2 program. The first task, N-back, had two different versions, depleting and non-depleting. Assignment to task version was randomized. The N-back task requires participants to keep a mental list of letters they see on the monitor and press a button indicating yes and no, regarding whether the letter they are seeing matches a letter they have seen previously. Participants pressed “F” on the keyboard for yes and “J” for no. The non-depleting task requires participants to indicate whether the current letter matches the letter that was presented 1 letter ago. In the depleting condition participants have to indicate if it matches the letter they saw 2 letters ago. Participants underwent a practice of 14 trials. There were 4 total blocks each containing 126 trials for a total of 504 trials. Each trial lasted for 2 seconds, with the letter appearing for only the first second. There was a 15 second break between each block which had a simple instruction reminder: “If the letter on the screen matches the one you saw 1 (or 2) letter ago, press “B” for yes. If not, press “N” for no.” There was a 3 second countdown on the screen, it then became blank for 1 second then moved to the next trial. Accuracy on identifying the matching letters was measured.

The SART measures effectiveness in overwriting a habituated response. Participants were shown words from two different categories and were required to press the space bar only when they saw words from one category. The present study used the categories animals and

vegetables, where a correct space bar press was when an animal was presented. Participants were required to press the space bar every time they saw an animal on the screen but not a vegetable. A majority of the words that were presented were animals, where the proportion was 8:1 favoring the animal category. This large number of animal trials (120 per block) ensured that the habituated response was to press the space bar. Percentage of correct space bar presses for animal trials and number of incorrect presses for vegetables trials were measured. There were 20 practice trials with a different category, boy names and girl names. Participants then underwent a total of 135 trials in 4 blocks for a total of 540 trials. Each trial lasted 1.3 seconds, with the letter appearing for only .3 seconds which was then replaced by a string of Xs (e.g. XXXXXXXXXXXX) for the remaining 1 second.

In each block, there were 9 mind-wandering thought probes presented at fixed-random intervals (for a total of 36 mind-wandering probes). The thought probe was 1 question with 15 seconds to answer which read: “What were you thinking about IMMEDIATELY before this screen appeared?” and participants answered with the keyboard either “1. The task” or “2. Something other than the task.” There was a 15 second break between each of the 4 blocks which had a simple instruction reminder: “Only press the spacebar when you see an animal on the screen.”

Manipulation check questions. For the manipulation check, there were additional questions presented before and after each task. Participants were prompted to respond on a scale of 1 (least) to 9 (most) for each of the questions. The question presented before the N-back and SART was “How motivated are you to do well on this task?” Questions presented after asked: “How difficult did you find this task?” “How much effort did this task require?” “How tired are

you after completing the task?” and “How frustrated were you during the task?” Participants responded using the number keys on the keyboard.

Procedure

This experiment was conducted during allotted, 1-hour time slots administered one-on-one with a research assistant. The questionnaires took approximately 8-10 minutes, and each computer task took 20 minutes each with a 1-2 minute debrief. All participant numbers were randomly assigned to a condition prior to the time slot. Research assistants gave a set of basic instructions to begin which included asking them to turn off their phone, telling them they can ask questions at any time, and remind them to try their best on all of the tasks.

Research assistants started by opening the web browser and beginning the questionnaire. They then moved to another part of the room behind a partition. When finished with the questionnaire, participants told the research assistant who would sit next to the participant and open the first task. To ensure immediate transfer from the first to the second task, instructions for the second (SART) task was presented before the first task, then the first task instructions were given. Research assistants read aloud all instructions for both tasks, then again moved to another part of the room when the participant was completing the first task. When finished with the first task, the researchers opened the second task and read aloud a brief reminder of the instructions for the second task. After the second task the researcher provided the participant with a credit sheet and a debriefing form and thanked them for their participation. During each portion of the experiment, the research assistant was instructed to look around the partition every 3-5 minutes to ensure the participant is paying attention and still trying during the tasks.

Results

Exclusion Criteria

The analyses for the current study was run on $n=98$ participants, out of the total $n=116$, 49 in the non-depleting condition and 49 in the depleting condition. Those who were removed from analysis were participants who missed too many catch questions on the questionnaires ($n=1$) or the questionnaire could not be completed ($n=1$). Further, participant data were removed for performance under 75% for the animal trials for the SART ($n=3$). This performance cutoff was implemented to ensure participants exerting little effort on the second task were removed. Additionally, due to some known glitches in PsychoPy version 1.85.4 causing data loss upon completion of the experiment, some participant data were excluded if the first or second task data were lost ($n=13$).

Condition Equivalence

Questionnaire data for this study were analyzed using independent sample t-test to test for between group differences. The results summarized in Table 1 showed that there was no significant difference between samples for beliefs about willpower on the first scale, $t(96)=0.45$, $p=0.65$, or the second scale $t(96)=0.67$, $p=.51$, or on the subset of the first willpower scale $t(96)=0.04$, $p=0.97$. Additionally, there were no differences in the Depletion Sensitivity Scale $t(96)=0.42$, $p=0.68$, Self-Control Scale $t(96)=0.97$, $p=0.33$, and Action Control Scale $t(96)=0.67$, $p=0.50$. These data show that the two groups (depletion vs. non-depletion) were approximately equal on the respective constructs of these questionnaires. As such, neither condition should exhibit any significant difference in performance on either task based on any questionnaire differences. Significant changes in performance could not be attributed to a difference in the aforementioned constructs.

Table 1. Questionnaire Measure Differences.

	Total (n = 98)	2-back (n= 49)	1-back (n= 49)	Differences (df=96)
Scales	Means (SD)	Means (SD)	Means (SD)	t (p)
Belief About Willpower 1	5.16 (0.74)	5.20 (0.87)	5.11 (1.06)	0.45 (0.65)
Willpower 1 – Subset 1: Strenuous Mental Activity	4.04 (0.87)	4.04 (1.19)	4.03 (1.44)	0.04 (0.97)
Belief About Willpower 2	4.36 (1.01)	4.48 (1.49)	4.28 (1.52)	0.67 (0.51)
Depletion Sensitivity Scale	5.57 (0.85)	5.63 (1.00)	5.53 (1.30)	0.42 (0.68)
Self-Control Scale	5.55 (0.71)	5.66 (1.08)	5.46 (0.90)	0.97 (0.33)
Action Control Scale	4.36 (1.33)	4.52 (1.74)	4.24 (2.28)	0.67 (0.50)

Note. SART= Sustained Attention to Control Task; Mean values and standard deviation on the 10-point scale for each questionnaire is presented. The right column reports t-values and p-values for the group differences. Significant values are bolded. * $p < .05$.

Manipulation Checks and Task 1 Performance

As shown in table 2, the data showed that there was a significant difference in mean accuracy for the depleting and non-depleting conditions $t(96)=-3.01$, $p=0.003$. The non-depleting 1-back task was found to have a better average performance than the 2-back task (see table 2). This shows that there was a significant difference in task demands which were adequately altered to create a difficulty difference. The subjective participant difficulty ratings reflected this as well. The 2-back task was perceived to be more difficulty on average than the 1-back task $t(96)=7.88$, $p=<.0001$. Other manipulation checks, such as effort and frustration were

found to be significantly higher in the depletion condition than the non-depletion condition. The only non-significant difference was in post-task fatigue, which was unexpected given the significant difference in task difficulty and accuracy. Participants in the 1-back condition ($M=6.86$, $SD=1.89$) showed a marginally higher average fatigue rating compared to the 2-back ($M=6.65$, $SD=1.77$). Finally, although motivation before the onset of the task was significantly higher $t(96)=-2.492$, $p=.014$ for the 1-back task ($M=7.796$, $SD=1.106$) than the 2-back task ($M=6.839$, $SD=2.119$), both conditions still reported high levels of motivation.

Table 2. Task 1 Manipulation Checks.

	2-back (n=49)	1-back (n=49)	Differences (df=96)
Manipulation Checks:	Means (SD)	Means (SD)	t (p)
Motivation	6.949 (2.125)	7.796 (1.106)	-2.492 (.014)*
Accuracy	.715 (.451)	.9322 (.251)	-3.012 (.003)*
Frustration	7.02 (1.531)	5.276 (2.266)	3.061 (.003)*
Effort	8.04 (1.124)	6.387 (1.957)	4.498 (<.0001)**
Difficulty	8.06 (.818)	5.469 (2.021)	7.881 (<.0001)**
Fatigue	6.653 (1.767)	6.857 (1.895)	-0.255 (.799)

Note. Significant values are bolded. * $p < .05$. ** $p < .001$.

Ego-Depletion and Task 2 Performance

Motivation for task completion was analyzed for SART in addition to the first task motivation. Importantly, there was not a significant difference in motivation for completion of the SART before the task was completed, $t(96)=-.06$, $p=0.95$. Due to the categories present in the

SART task, data were analyzed for accuracy for performance for each category, where animal accuracy was based on percentage correct and vegetables based on percentage incorrect. As expected, there was no difference between conditions for the correct proportion of animal trials $t(96)=-0.15, p=0.87$.

As reported in table 3, while there was an expected difference in percentage of incorrect trials when a vegetable was presented, there was not a significant difference between groups $t(96)=-0.41, p=.68$. Further, mind wandering rates were not significantly different between the depletion and non-depletion conditions $t(96)=.47, p=0.64$. The mean levels of mind wandering between the 2-back condition ($M=0.28, SD=0.22$) and the 1-back condition ($M=0.26, SD=0.21$) were only marginally different. These results suggest that there is no effect of the task 1 condition on SART performance and inhibitory control errors which goes against the notion that the resource model of ego-depletion is correct.

Table 3. Task 2 Motivation and Performance Measures.

	2-back (n=49)	1-back (n=49)	Differences (df=96)
Measures:	Mean (SD)	Mean (SD)	t (p)
Motivation	7.10 (1.73)	7.12 (1.78)	-0.06 (0.95)
Correct Animal Responses	0.97 (0.03)	0.97 (0.03)	-0.15 (0.87)
Incorrect Vegetable Responses	0.53 (0.23)	0.55 (0.22)	-0.41 (0.68)
Mind-Wandering Rate	0.28 (0.22)	0.26 (0.21)	0.47 (0.64)

Note. Significant values are bolded. * $p < .05$.

Correlations Among Questionnaires

Correlations were run to test whether certain constructs measured in the questionnaires were commonly concurrent in participants. As represented in table 4, the correlations show that that the “Strenuous Mental Activity” subset of the first beliefs about willpower scale was correlated with the overall Beliefs About Willpower Scale, the second Beliefs About Willpower Scale, the Depletion Sensitivity Scale and the Action Control Scale. The second Belief About Willpower Scale has a medium correlation to the first scale as well as the first subset of the first scale. Depletion sensitivity is correlated highly to both the first and second belief about willpower scale. The general self-control scale had a medium correlation with the depletion sensitivity scale and general mind-wandering rates. Action control was also had a medium sized correlation with the depletion sensitivity and self-control scale. The only questionnaire that had a meaningful sized correlation with inhibitory control in the SART task was the in-task mind-wandering probes which had a slightly significant, small to medium sized correlation.

Table 4. Correlations Among Questionnaires.

Scale #		1	2	3	4	5	6	7	8
1	Belief About Willpower 1	--							
2	Willpower 1 – Subset 1: Strenuous Mental Activity	0.72***	--						
3	Belief About Willpower 2	0.31*	0.46***	--					
4	Depletion Sensitivity Scale	-0.53***	-0.70***	-0.49***	--				
5	Self-Control Scale	0.23	0.08	-0.02	-0.48***	--			
6	Action Control Scale	0.20	0.29*	0.08	-0.47***	0.42**	--		
7	Mind-Wandering	-0.14	-0.08	0.10	0.14	-0.38**	-0.19	--	
8	Inhibitory Control Errors (% correct responses on vegetable trials)	0.09	0.13	0.18	-0.06	-0.13	0.05	0.34*	--

Note. Bolded values show a significant correlation. Number of * corresponds to level of significance.

Moderating Effects

To assess moderating effects of willpower beliefs, regressions were run to test each questionnaire and condition interaction on SART performance and mind-wandering rates. These analyses would show whether ego-depletion is mediated by various perceived levels of self-control, willpower, action control, etc. Additionally, interactions of questionnaire responses by condition were run. The results of each questionnaire with each condition interaction is reported in Table 5. The results show that the action control scale to be a significant predictor of mind wandering rates for the non-depletion condition, $\beta = -0.28$, $p = .009$. Importantly, none of the other

questionnaires or interactions (including willpower belief questionnaires) were shown to be a significant predictor of mind wandering rates or incorrect vegetable trials. These results are inconsistent with previous research that suggests that mindset about self-control predicts performance on self-control tasks.

Table 5. Regression Analyses for Condition and Questionnaires on Inhibitory Control Errors and Mind Wandering.

	Inhibitory Control Errors		Mind Wandering	
	β	p	β	p
Condition	-0.04	0.69	0.05	0.64
BAW1	0.02	0.83	-0.1	0.33
BAW1xCond	-0.01	0.99	-0.17	0.09
Condition	-0.04	0.68	0.05	0.64
BAW1-1	0.05	0.65	-0.04	0.71
BAW1-1xCond	0.04	0.69	-0.17	0.1
Condition	-0.04	0.68	0.05	0.65
BAW2	0.01	0.95	0.03	0.77
BAW2xCond	0.06	0.54	0.17	0.09
Condition	-0.03	0.74	0.06	0.52
SCS	-0.11	0.29	-0.23	0.03
SCSxCond	0.07	0.48	-0.01	0.91
Condition	-0.04	0.69	0.05	0.64
DSS	-0.04	0.69	0.14	0.18
DSSxCond	-0.01	0.97	0.11	0.29
Condition	-0.04	0.72	0.07	0.47
ACS	-0.06	0.57	-0.28	0.009*
ACSxCond	0.06	0.56	-0.03	0.79

Note. Condition = Depletion (2-back) vs. Non-Depletion (1-back); SCS= Self-Control Scale; BAW1= Beliefs About Willpower Scale 1; BAW1-1= Beliefs About Willpower Scale-Subset 1; BAW2= Beliefs About Willpower Scale 2; DSS= Depletion Sensitivity Scale; ACS= Action Control Scale. Significant values are bolded. * p < .05.

Discussion

In this study, the ego-depletion effect was assessed using the sequential-task paradigm. This was done in conjunction with questionnaires and mind wandering measures to test the reliability of the strength model of self-control and assess for potential moderating effects. The main goal of this study was to address many glaring issues with previous research and test the robustness of the ego-depletion effect. I tested two main hypotheses, and these are the general, but tentative, conclusions regarding the results of the study.

Evaluation of the Two Main Hypotheses

Hypothesis 1. The first hypothesis that was measured looked at inhibitory control failures as well as attentional lapses (i.e., mind wandering) after completing the first task. It was hypothesized that if the ego-depletion effect is real then participants in the depleting condition should perform worse on the inhibitory control and mind-wandering measures compared to their non-depleted counterparts. The results of the study show that there was no significant difference between conditions for either the inhibitory control failures or attentional lapse measures.

Although the sequential-task paradigm has been used in the past to test for the effect of ego-depletion, the current study was different in a few important ways. Namely, the study assessed mind-wandering rates, increased task duration, reduced time between tasks, and increased sample size. These were implemented to increase the likelihood of fatigue during task completion and to give adequate power for analyzing multiple factors of the ego-depletion effect. Further, using a working memory task in place of a self-control task for the first task aimed to test whether self-control is pulled from a common mental resource or is independent from inhibitory resources.

Due to the important changes in task characteristics and implemented measurements for the effect of ego-depletion, this study has a potential for impacting the narrative surrounding ego-depletion. The general assumption in the field based on previous research is that the strength model is a robust effect that explains how people operate in inhibitory situations throughout daily life. The current study adds to the growing body of research that challenges this prevalent support for the strength model.

Based on the results, this study contributes to the field of ego-depletion insofar that it contributes to the replication efforts for sequential-task studies of self-control. If evidence was found in previous research for ego-depletion, it has shown a much lower effect size. This study remedies some of the complaints about the previous studies while simultaneously remedying the major methodological and conceptual issues. Namely the issues laid out by Lurquin and Miyake (2017) have all been addressed, both theories of ego-depletion are tested, and all of the methodological concerns are covered. The only proviso to these changes and general limitations for this study was regarding first task fatigue and motivation ratings. It was found that there was not a significant difference in fatigue after completing the depleting task vs the non-depleting task. This is problematic in that although they differed in difficulty, we cannot know at this time if there was a so-called “depleting” experience. Further, there was a significant difference in motivation ratings for the first task. Although it was expected that there would be no difference in motivation ratings, the non-depleting condition was rated higher on average than the depleting condition. Despite the existence of this mean difference, both conditions had moderately high levels of motivation.

Hypothesis 2. The second hypothesis used the lens of one’s beliefs about willpower and how those beliefs affect inhibitory ability. It was hypothesized that if the mindset theory of self-

control is true then there would be a notable interaction between belief about willpower and self-control scales and the condition differences. There was no such support of that hypothesis. There was no evidence that any of the primary self-control scales, or any of the supplementary scales, found any significant interaction with the condition on performance measures.

More specifically correlation between questionnaires were measured to see if there was a true supplementary aspect of the additional questionnaires on the primary beliefs about willpower questionnaires. Although there were a few notable correlations between the measures, only the self-control scale was correlated with mind-wandering, and none of the measures showed an interaction with the condition. As such it can be asserted that mindset has no significant impact on inhibitory control ability or “reduce” the ego-depletion effect or tendency to mind wander. Although these are the most convincing explanation for the results based on the data, these results are still tentative. More participants are being run for the study and due to some of the measures being close to significant, there is a potential for some correlational measures to change.

Future Directions

Although the current study did address many methodological issues with the research regarding ego-depletion there were still some shortfalls. At the point of analyses for the current study, there were not as many collected participants as would be preferred. With an extended piloting stage to ensure task difficulty and structure was fine-tuned, there was less time to collect real participant data. The current number of participants per condition is adequate given the cutoff of this Honors Thesis, but there is always derived benefit from increasing participant numbers. That being said, 49 participants per cell is markedly higher than some other original

studies looking at the effect of ego-depletion. Going forward there will be more participants being run to increase participant total to 200 usable subjects.

There will also be more analyses run on the increased sample size. In addition to measuring the general accuracy of participant responses and comparing group means, there will be accuracy measures by block of each task. Accuracy by block would give more insight into the ego-depletion effect as it would allow a closer look into task performance. For example, if a participant had a generally high accuracy it could be concluded that ego-depletion is not at work, but if there is a decline in performance as the task progresses, it could indicate some level of ego-depletion. D-prime analyses would be beneficial moving forwards as it would provide a signal detection type analyses of task performance. If a participant had a general tendency to not press the spacebar, they would have low accuracy but perform better on inhibitory control trials. These analyses would help distinguish correct button presses, incorrect button presses, false-alarms, and correct rejections.

More generally, further directions of ego-depletion research should be focused around task choice and lab-based tasks. Although it was decided to use the sequential-task paradigm which has been widely used in past ego-depletion research, the model is not applicable to everyday inhibitory situations. The idea of the study design is intuitive and appealing and within the lens of ego-depletion it makes sense. Although there is a derived benefit of using a lab-oriented task to improve consistency and ease of implementation, the lack of supporting findings could be due to the lab-based tasks themselves. Often, lab-based tasks are short and contrived and may not be a good representation of the actual effect. Since ego-depletion is a convincing idea, future studies that use real life inhibitory situations that are intuitive in nature to potentially show the effect better than lab-oriented tasks.

References

- 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., Vohs, K. D., & Tice, D. M. (2007). The Strength Model of Self-Control. *Current Directions in Psychological Science*, 16(6), 351-355. doi:10.1111/j.1467-8721.2007.00534.x
- 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.
- Hagger, M. S., Chatzisarantis, N. L., Alberts, H., Anggono, C. O., Batailler, C., Birt, A. R., ... & Calvillo, D. P. (2016). A multilab preregistered replication of the ego-depletion effect. *Perspectives on Psychological Science*, 11(4), 546-573.
- 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. <http://dx.doi.org/10.1037/a0019486>
- Job, V., Dweck, C. S., Walton, G. M. (2010). Ego Depletion--Is It All in Your Head?. *Psychological Science*, vol. 21, no. 11, 2010, pp. 1686–1693., doi:10.1177/0956797610384745.
- Kuhl, J., & Kazén, M. (1994). Self-discrimination and memory: State orientation and false self-ascription of assigned activities. *Journal of Personality and Social Psychology*, 66(6), 1103-1115. doi:10.1037//0022-3514.66.6.1103

- Lurquin, J. H., Michaelson, L. E., Barker, J. E., Gustavson, D. E., Bastian, C. C., Carruth, N. P., & Miyake, A. (2016). No Evidence of the Ego-Depletion Effect across Task Characteristics and Individual Differences: A Pre-Registered Study. *Plos One*, 11(2). doi:10.1371/journal.pone.0147770
- Lurquin J.H., & Miyake A. (2017) Challenges to Ego-Depletion Research Go beyond the Replication Crisis: A Need for Tackling the Conceptual Crisis. *Front. Psychol.* 8:568. doi: 10.3389/fpsyg.2017.00568
- McVay, J. C., & Kane, M. J. (2009). Conducting the train of thought: working memory capacity, goal neglect, and mind wandering in an executive-control task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35(1), 196.
- Moller, Arlen C., et al. (2006). Choice and Ego-Depletion: The Moderating Role of Autonomy. *Personality and Social Psychology Bulletin*, vol. 32, no. 8, 2006, pp. 1024–1036., doi:10.1177/0146167206288008.
- 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., Shmueli, D., & Burkley, E. (2006). Conserving self-control strength. *Journal of personality and social psychology*, 91(3), 524.
- Salmon, S. J., Adriaanse, M. A., De Vet, E., Fennis, B. M., & De Ridder, D. T. (2014). “When the going gets tough, who keeps going?” Depletion sensitivity moderates the ego-depletion effect. *Frontiers in psychology*, 5, 647.

- Savani, K., & Job, V. (2017). Reverse ego-depletion: Acts of self-control can improve subsequent performance in Indian cultural contexts. *Journal of Personality and Social Psychology*, 113(4), 589-607. <http://dx.doi.org/10.1037/pspi0000099>
- 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), 33-46.
- Sripada, C., Kessler, D., and Jonides, J. (2014). Methylphenidate blocks effort- induced depletion of regulatory control in healthy volunteers. *Psychol. Sci.* 25, 1227–1234. doi: 10.1177/0956797614526415
- Webb, T. L., & Sheeran, P. (2003). Can implementation intentions help to overcome ego-depletion?. *Journal of Experimental Social Psychology*, 39(3), 279-286.
- Xu, X., Demos, K. E., Leahey, T. M., Hart, C. N., Trautvetter, J., Coward, P., et al. (2014). Failure to replicate depletion of self-control. *PLoS ONE* 9:e109950. doi: 10.1371/journal.pone.0109950