Depletion and Replenishment: Exploring Self-Regulation Resource Depletion, Activities that Replenish the Resource, and the Corresponding Effects on Mood

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Sara L. Klaphake

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Wilma Koutstaal, PhD.

August 2011

Acknowledgements

I owe a debt of gratitude to many individuals for their support for and assistance with this research and manuscript. First, I would like to thank my advisor, Wilma Koutstaal, for mentoring me throughout my graduate career. Her thoughtful guidance and encouragement have been essential as I have progressed toward a completion of my doctoral degree. I am especially grateful that she suggested I take a course on self-regulation, providing the inspiration for this work. I would also like to thank Mary Kennedy for offering that class and encouraging my interest in the field. Kathleen Vohs was also integral to this research, pioneering the theory of self-regulation depletion on which it is based and assisting in design considerations for these studies. I am also grateful to Angus MacDonald III and Chad Marsolek for their valuable time and advice.

This research was also made possible by the hard work of many undergraduate research assistants, including: Heather Barto, Tatyana Berkhina, Nathan Bjornberg, Lisa Browning, Yoobin Cho, Anna Crist, Alana Feijo, Nate Follen, Qianhui Gong, Annica Hall, Scott Hargarten, Andy Hey, Brooke Huibregtse, Soon Min Kwon, Seongmi Lee, Yue Li, Abbas Mulla, Dominic Mussack, Jason Neff, Matt Russell, Alex Schwegman, Eric Swanson, and Ashley Thuemmler.

Funding for the research was provided in part by a Graduate Research Partnership Program (GRPP) fellowship from the University of Minnesota.

Dedication

This dissertation is dedicated to my family and friends, who have patiently supported me throughout my graduate career, especially my loving husband, Jonathan Klaphake, and my parents, Timothy and Lynn Kvidera.

Abstract

The resource depletion model of self-regulation proposes that people's ability to perform deliberate, effortful thought and behavior is a limited resource, with earlier self-regulation depleting this resource, leaving less for subsequent tasks. The current research investigated the pervasiveness of self-regulation depletion, explored various tasks as potential means of counteracting depletion, and assessed how mood was impacted by both. In a series of seven studies, participants completed an initial task that required either high or low levels of self-regulation and subsequent self-regulation measurement tasks, along with pre- and post-task measurements of mood. Some participants also completed various intervening tasks to assess the potential of different activities to counteract depletion. Our research indicated that self-regulation depletion, while common, is not inevitable when one has completed an earlier self-regulation task. We also found little evidence that completing the various intervening tasks such as exercise, magazine reading, and drawing, replenished participants' self-regulation resources. We did, however, find clear evidence that mood, both in terms of mood valence and arousal levels, is impacted by self-regulation, and some indication that these mood effects played a small mediating role in depletion.

Table of Contents

List of Tables

List of Figures

Background

Self-regulation can be broadly defined as "the many processes by which the human psyche exercises control over its functions, states, and inner processes" (Vohs & Baumeister, 2004b, p. 1). Thus, it is typically characterized as a controlled, intentional process[1] whereby we direct our actions and thoughts in accordance to our goals (see Vohs & Baumeister, 2004a for a comprehensive review). Aspects that fall under the purview of self-regulation include behavioral regulation such as resisting temptation or persisting in an unpleasant but necessary task, emotional regulation, and attentional regulation (see, for example, Baumeister, Vohs, & Tice, 2007). The same or similar concepts in other areas of psychology may be referred to as self-control, controlled processing, "executive control,"[2] or even using the more informal and colloquial term of "willpower" (see Moors & De Houwer, 2006, for an extensive list).

Because self-regulation is generally characterized as intentional or goal-driven, many models of the process of self-regulation begin with a representation of a goal or goal state and the selection of behaviors in order to reach the goal state (e.g., Metcalfe & Mischel, 1999; Schneider & Chein, 2003). Next, a memory process is typically included to maintain a representation of the goal (Schneider & Chein, 2003). Another key aspect of most self-regulation models is the inclusion of monitoring processes (Kahneman, 2003; Kennedy & Coelho, 2005; Schneider & Chein, 2003). In addition to selecting a goal, maintaining the goal, and taking steps to achieve it, the person must determine how

[1] However, see Fitzsimons & Bargh (2004) for an argument that self-regulation should encompass automatic behaviors as well.
[2] "Executive control" is typically used to refer to control of cognitive processes, whereas "self-regulation" and "self-control" more often refer to controlling behaviors and emotions.

much progress he or she is making and when the goal has been accomplished. This monitoring function is often represented in the form of a feedback loop that continuously monitors the discrepancy between the current state and the goal state (Kennedy & Coelho, 2005; Schneider & Chein, 2003).

Assessing Self-Regulatory Capacity

A number of paradigms have been designed to assess participants' self-regulation/executive function. Frequently, the tasks require participants to inhibit a prepotent response. In self-regulation research, this often involves resisting temptation, such as resisting eating tasty snacks (Vohs & Heatherton, 2000) and various other delay-of-gratification tasks (see Mischel & Ayduk, 2004, for a review). Similarly, cognitive researchers and neuroscientists often use response-inhibition paradigms called Go/No-go tasks in which participants are instructed to respond (e.g., by pressing a button) for most stimuli but not to respond (e.g., by abstaining from pressing the button) for certain other stimuli (e.g., Passamonti et al., 2006; see Simmonds, Pekar, & Mostofsky, 2008 for a review). This creates an experimental situation wherein one response becomes highly pre-potent, but infrequently and usually unpredictably, a different response is required. Paradigms such as the Go/No-go task are classified as stimulus-response conflict tasks; this category also includes activities such as flanker tasks, in which participants are asked to respond only to a centrally positioned item (e.g., the middle arrow in a set of five arrows) that may contain the same or conflicting items on either side (Eriksen & Eriksen, 1974; e.g., Verbruggen, Notebaert, Liefooghe, & Vandierendonck, 2006) and Stroop tasks wherein participants must respond to one aspect of a stimulus (e.g., font color) rather than another, often incongruous, stimulus aspect (e.g., the word, "red," spelled out in blue ink) (Stroop, 1935, e.g., Inzlicht and Gutsell, 2007). For additional reviews of self-regulation capacity and ways in which to assess it, see Baumeister (2002), Baumeister et al. (2007), and Muraven and Baumeister (2002).

Neural Correlates

Researchers have also begun to identify potential neural structures that are associated with self-regulation. Generally, the prefrontal cortex (PFC) has been associated with the control aspect of self-regulation, whereas the anterior cingulate cortex (ACC) is typically associated with the monitoring aspect (e.g., Lungu, Liu, Waechter, Willingham, & Ashe, 2007; MacDonald, Cohen, Stenger, & Carter, 2000; see Carter & VanVeen, 2007, for a review). In particular, dorsolateral prefrontal cortex (DLPFC) is most often linked to cognitive control (cf. Brass, Derrfuss, Forstmann, & von Cramon, 2005). DLPFC is often activated during working memory tasks and thus is thought to play a role in retaining the goals of the system (Barch, Braver, Nystrom, Forman, Noll, & Cohen, 1997; Braver, Reynolds, &Donaldson, 2003; Mansouri, Buckley, & Tanaka, 2007). However, DLPFC appears to play a more active role in control and manipulation of cognitions as well (Badre & Wagner, 2004; Wagner, Maril, Bjork, & Schacter, 2001). Schneider and Chein (2003) have also suggested neural analogues to the components of their controlled processing model. They suggest that the DLPFC may be a good candidate for the instantiation of their Goal Processor; these researchers further link the posterior parietal cortex to their Attentional Control Processor and propose the ACC as the site of their Activity Monitor.

Self-Regulation Depletion

Recently, a number of researchers have begun to conceptualize self-regulation/ executive control as a resource that can be depleted (e.g., Baumeister, Bratslavsky, Muraven, & Tice, 1998; Mischel & Ayduk, 2004; Schmeichel, 2007; Vohs & Heatherton, 2000). The resource depletion model of self-regulation proposes that people's ability to perform deliberate, effortful thought and behavior is a limited resource (Baumeister, 2002; Baumeister et al., 2007, Schmeichel & Baumeister, 2004). Therefore, when a person completes one effortful behavior, he/she has less of this resource left to utilize for a subsequent effortful behavior. This person is thus said to be "depleted" and is believed to demonstrate an increased likelihood of relying on more automatic or habitual processing while in this "depleted" state. In sum, when one has previously been performing a demanding control task, one fails to exert control in a subsequent task because one does not have sufficient self-regulation resources remaining in order to exert self-control.

Research on this model has established that when someone is depleted in this manner, his or her performance on a second task that likewise places demands on self-regulatory processes suffers. For example, early experimental demonstrations of the resource depletion effect[3] illustrated that participants who were asked to regulate their behavior in order to eat radishes instead of consuming prominently available freshly baked cookies demonstrated less persistence on a subsequent tracing task; in another experiment, those who first had to inhibit their emotional reactions had lower scores on a

---

[3] Self-regulation resource depletion was originally referred to as ego depletion (Baumeister et al., 1998) and these researchers still sometimes use the "ego depletion" term.

later anagram-solving task (Baumeister, Bratslavsky, Muraven, & Tice, 1998). The researchers interpreted the participants' lack of persistence on the tracing and anagram tasks as indication that the participants were less able to self-regulate in terms of sustaining their behavior and attention. The depleting effects of emotional regulation have also been demonstrated in tasks more similar to real-world situations. For instance, Goldberg and Grandey (2007) found that participants assigned to act in accord with a "Service with a smile" set of rules for customer service representatives made more errors on order forms and even felt more exhausted after their interactions with customer-confederates in a simulated call-center environment than did participants not required to smile in such a rule-bound manner.

Muraven, Tice, and Baumeister (1998) further found that like suppressing a desire or emotion, suppression of a particular thought or topic can be depleting as well. Participants who originally suppressed a prohibited thought (e.g., do not think about a white bear) were later less able to regulate their emotional reactions. Many additional studies have demonstrated similar depleting effects of thought suppression on a number of subsequent self-regulation measures including decision-making (e.g., Fischer, Greitemeyer, & Frey, 2008) stop-inhibition tasks (e.g., Muraven, Rosman, & Gagné, 2008) and persistence tasks (e.g., Muraven & Slessareva, 2003).

A number of other situations requiring behavioral self-regulation likewise have been found to both cause and suffer from resource depletion. A series of studies by Vohs, Baumeister, and Ciarocco (2005) illustrated that portraying oneself in a particular manner can be depleting, especially when it requires acting in ways contrary to one's disposition or in opposition to typical social norms. These researchers further established that the reverse was also true; participants who had earlier performed unrelated self-regulation tasks were subsequently less able to regulate their self-presentations effectively. In addition, Finkel, Campbell, Brunell, Dalton, Scarbeck, and Chartrand (2006) found that engaging in difficult social interactions impaired participants' later performance on various problem solving tasks. Lastly, Muraven (2008b) found that earlier regulation attempts also lead people with strong racial biases to respond more prejudicially.

Cognitively demanding tasks seem to rely on this self-regulation resource as well. Schmeichel, Vohs, and Baumeister (2003, p. 33) found that completing earlier self-regulation tasks such as attention and emotion regulation led to deficits on subsequent complex cognitive tasks that "require active guidance by the self" such as logic and reasoning tasks as well as reading comprehension tasks. Schmeichel (2007) likewise found that earlier engagement in self-regulation led to deficits in cognitive tasks such as working memory updating. He also found cognitive, executive control-type tasks (e.g., attentional control) to be depleting. In addition, Vohs, Baumeister, Schmeichel, Twenge, Nelson, and Tice (2008) found that being asked to make numerous choices (as compared to various control conditions, including merely considering the favorability of the options, having choices made for one by others, and implementing the choices of others) hindered performance on a number of subsequent self-regulation tasks.

However, the theory as well as empirical evidence suggests that not all cognitive tasks tap the self-regulation resource. Schmeichel et al. (2003) demonstrated that earlier emotional regulation did not impair later performance on a general knowledge task, and earlier attentional regulation did not impair later memorization and recall performance. Thus, there are apparently limits or boundary conditions on the activities that rely on the self-regulation resource (and that are thus adversely affected by self-regulatory depletion).

In addition, the self-regulation depletion effect is thought to be limited to tasks that require self-regulation and is not simply a generalized mental fatigue or difficulty effect (e.g., see Baumeister, Galliott, DeWall, & Oaten, 2006). In support of this position, participants have been shown to perform more poorly after self-regulation tasks than after non-self-regulation tasks even when the two tasks were perceived to be equally difficult, unpleasant, and frustrating (Muraven et al., 1998; Muraven & Slessareva, 2003). Furthermore, Schmeichel (2007) found no differences in the ability of participants to regulate their emotions after an easy or hard *non-self-regulation* task. Taken together, such findings appear to indicate that the crucial factor in determining whether an initial task will generate self-regulatory depletion is the degree to which it requires self-regulation, not task difficulty per se.

Schmeichel et al. (2003) further explored whether depletion effects are also dependent on the degree to which the *subsequent* task requires self-regulation, again as opposed to its overall degree of difficulty. In one study, participants achieved a lower percentage correct on a subsequent non-self-regulation task than on a subsequent self-regulation task, suggesting that the non-self-regulation task was the more difficult task. In line with the theory that only self-regulation tasks "tap the resource", however, performance of the initial self-regulation task only affected performance on the subsequent self-regulation task. Thus, overall, it appears that the degree to which a task requires self-regulation, not its overall difficulty, is the defining feature required in order for a task to either produce or be susceptible to depletion.

Biological Foundations

Inzlicht and Gutsell (2007) provide interesting event-related potential (ERP) electroencephalography (EEG) evidence suggesting that it is particularly the monitoring component of self-regulation that is depleted. Specifically, they found that the error-related negativity (ERN), an ERP component frequently associated with detecting that one has made an error and generally localized to the anterior cingulate cortex, was weaker while participants were completing a Stroop task if they had previously had to suppress their emotional responses to a graphic film. The work of Inzlicht and Gutsell (2007) further implies that when one is depleted, one may not be as likely to re-engage self-regulation processes when needed because the monitoring component that cues one of the need to regulate is not functioning normally.

There is some preliminary evidence that depletion impacts brain activity in the regions associated with the control aspect of self-regulation as well. Hedgcock, Vohs, and Rao (under review) have recently provided functional magnetic imaging (fMRI) evidence that activity in the right middle frontal gyrus – roughly equivalent to the DLPFC, which as stated above, has been linked to the control aspect of controlled processing – appears to be influenced by depletion. In a rare use of a repeated measures design in depletion research, participants were asked to ignore words in one experimental phase (depleting condition) and were allowed to attend to them at their whim in another phase (non-depleting condition). After each of these phases, participants were asked to make choices on topics such as hotels and cars, and the impulsiveness of their choices was measured based reaction times. The results indicated that participants made more impulsive decisions and showed less activity in the right middle frontal gyrus during the subsequent decision task when it followed the depletion phase than when it followed the non-depletion phase.

It is difficult from the data of Hedgcock et al. to parse whether the reduction in brain activity was the cause or the result of participants' reduced decision times. For example, one might assume that the brain would be less active when less time is spent on a decision. However, the differential in activation in the right middle frontal gyrus they found does not appear to reflect a global processing reduction, as other brain regions did not show similar differences in activity between conditions. Furthermore, Richeson et al. (2003) have also linked the right middle frontal gryus to self-regulation depletion in a study on racial bias. The brain activity of white participants was assessed via fMRI while they viewed black faces. The authors found that participants' brain activity in this region when viewing black faces (but not white) was positively correlated with their racial bias questionnaire scores. The authors argued that this likely reflected the fact that, due to the values of the current culture, those with racial biases felt pressure to control their thoughts and behavior in order to seem more egalitarian. Of greater relevance to the depletion theory, however, was the fact that participants' brain activity in the right middle frontal gyrus when viewing the black faces was also positively correlated with the degree of Stroop interference they displayed after an interracial interaction outside of the scanner. Indeed, the activity in the region mediated the relationship between prejudice and depletion (i.e., Stroop interference).

Moderating and Counteracting Depletion

In sum, numerous studies over the past decade have demonstrated the reality of self-regulation depletion effects and there appears to be some biological basis for them. However, it also appears that the strength or existence of depletion can depend on certain characteristics of the participant and the situation. For example, Vohs and Heatherton (2000) found that the extent to which a task proves to be depleting can depend on one's goals. In one study, dieters were less able than were non-dieters to exert self-control on the subsequent task when the first task involved resisting tempting candy.

Furthermore, in contrast to the implications regarding impaired monitoring during depletion suggested by Inzlicht and Gutsell (2007), people – even when depleted – seem to be able to monitor their reserves of this self-regulation resource and to control their actions accordingly (Muraven, Shmueli, & Burkley, 2006). Muraven et al. (2006) found that the participants who expected another challenging task to follow "conserved" some of the self-regulation resource, and this was especially true if the participant was already somewhat depleted from a previous task. Moreover, the participants' conservation efforts were effective; those who conserved their resources (as behaviorally indicated by poorer performance on an earlier Stroop task) did achieve higher scores on the final task.

Researchers are also just beginning to explore what one might be able to do to counteract depletion. Some of the earliest findings have demonstrated that consuming glucose (sugar) replenished the self-regulation resource by providing energy (e.g., Gailliot, 2008; Gailliot et al., 2007; see also Denson, von Hippel, Kemp, & Teo, 2010 for a recent replication, as well as Miller, Pattison, deWall, Rayburn-Reeves, & Zentall, 2010, for a generalization of the effect to domestic dogs). However, consuming sugar to replenish cognitive resources after every challenging task is not likely to be recommended, given the problems with diabetes and obesity in this country and abroad (World Health Organization: http://www.who.int/en/) as well as other potential drawbacks such as tooth decay (American Dental Association: http://www.ada.org). Therefore, research is needed to assess the restorative properties of other non-consumptive activities.

Since the establishment of the theory, originators of the self-regulation resource theory have argued that rest and "time off" from self-regulation tasks should have replenishing effects (e.g., Muraven & Baumeister, 2000), pointing, for anecdotal support, to such observations as increased failures in self-regulation in the evening – presumably following upon many daytime self-regulation efforts. More direct empirical evidence has also been found in support of the idea that "time off" may be replenishing (Knowles, Brennan, & Linn, 2002, in Vohs & Ciarocco, 2004; Oaten, Williams, Jones, & Zadro, 2008; Tyler & Burns, 2008). Oaten et al.'s (2008) main focus was comparing the depleting effects of ostrascism between those who were socially anxious and those who were not. However, in doing so, they compared how long the depletion effects lasted. They found that shortly after the ostracism, all of the ostracized participants struggled to self-regulate (e.g., ate more cookies). After a delay of 45 minutes, though, only those with high anxiety continued to have difficulties with self-regulation. Tyler and Burns (2008) set out to systematically examine the usefulness of breaks between self-regulation tasks and the impact of the duration of that break on subsequent performance. They found that a break of 10 minutes (but not of 3 minutes) wherein participants completed basic questionnaires allowed depleted participants to perform as well as non-depleted participants on a subsequent persistence task.

The type of activity engaged in between self-regulation tasks also seems to influence the restorative properties of the break. Tyler and Burns (2008) found that a shorter, three minute break with an emphasis on relaxation and the addition of relaxing music reduced depletion. Further research has suggested that enhancing mood can reduce or even eliminate depletion effects (Tice, Baumeister, Schmueli, & Muraven, 2007). Depleted participants who were shown an amusing video or were given a surprise gift were able to perform as well as non-depleted participants on various persistence tasks. Additional evidence that one's mindset during the break is influential comes from Schmeichel and Vohs (2009). They found that having depleted participants engage in self-affirmation exercises (e.g., writing about the importance of one of their core values) helped participants to perform subsequent self-regulation tasks. The self-affirmation exercises were found to increase the level of abstraction at which participants thought about common actions, suggesting that a shift toward a more abstract (rather than concrete) mental construal level may benefit self-regulatory control (cf. Schmeichel, Vohs, & Duke, in press). Lastly, intervening tasks can influence depletion effects through subconscious processes. Research has demonstrated that participants can be primed for persistence by seeing words relating to persistence or reading an article about a persistent individual (Alberts, Martijn, Greb, Merckelbach, & de Vries, 2007; Martijn, Alberts, Merckelbach, Haermans, Huijts, & de Vries, 2007). Furthermore, this implicit intervention led depleted participants to perform nearly as well or as well as non-depleted participants.

Lastly, the way in which the self-regulation tasks are presented appears to play a role. Muraven, Gagné, and Rosman (2008) found that presenting the initial self-regulation task in a manner that supported autonomy reduced or eliminated its depleting effects. Muraven and Slessareva (2003) also demonstrated that by increasing participants' motivation levels for the final task, depletion effects can be overcome. Depleted participants who were motivated to help others (i.e., told the task would help to develop treatments for Alzheimer's disease) or for personal gain (e.g., monetary compensation) were able to perform as well as non-depleted participants. Nevertheless, the authors argued that the increased motivation more likely encouraged participants to use even more of their already reduced resource, rather than actually replenishing the resource. Whether this an accurate conceptualization is still in debate. However, the argument fits well with the the conservation of resources findings by Muraven et al. (2006) reported above, wherein depleted participants expended less effort on a second self-regulation task if they expected a third self-regulation task was yet to come. One would expect that a strong enough motivation to accomplish a given task could overcome the desire to conserve resources. Similar arguments and findings were provided by Stewart, Wright, Hui, and Simmons (2009) in their study exploring mental fatigue, expectations of reward, and cardivascular measurements. The authors argued that fatigued participants would put forth more effort when motivated by a higher probability reward. As evidence, their data indicated that fatigued participants' cardiovascular responses rose when their expectations of reward were high, but fell slightly when reward expectations were low.

Mood, Self-Regulation, and Depletion

Another issue beginning to garner more attention is the relationship between self-regulation and mood. For instance, Metcalfe and Mischel (1999) highlight mood in their theory of self-regulation and self-regulation failure. Their theory focuses on contrasting what they call "hot" and "cold" systems. Their "cool" system is "complex, slow, and contemplative" and generates "rational, reflective, and strategic behavior" (Mischel & Ayduk, 2004, p. 109), and therefore is akin to the controlled processes of self-regulation and executive function. In contrast, the "hot" system is more automatic – it is described as reflexive/under stimulus control, and is characterized by rapid emotional responding. Metcalfe and Mischel (1999) argue that high levels of stress heighten the predominance and effectiveness of the hot system. In contrast, while the cool system is amplified by low levels of stress, higher levels of stress are detrimental to the functioning of the cool system, leading the individual to rely on the hot system.

Other researchers have more directly addressed the interaction of the cognitive and emotional domains. In his theory of mood-as-information, Schwarz (1990; 2001) posits that moods – or feelings more generally – act as an indicator of one's situation. Positive moods signify "benign" situations in which no specific action is required.

Because of this, people in good moods often show more flexible thinking and creativity, and are more likely to spontaneously use less systematic and more heuristic approaches to problem-solving. In contrast, negative moods signal to the individual that one's situation is problematic in some manner and requires some action to remedy it. Thus, those in negative moods are more likely to be detail-oriented and to use analytic reasoning. Support for this account has been provided in a number of studies (e.g., Converse, Lin, Keysar, & Epley, 2008; Forgas, 1998; Gasper, 2004; Wyland & Forgas, 2010; see Clore & Huntsinger, 2007 for a review).

Other investigators have developed theories with similar, but more nuanced, implications. As summarized previously, self-regulation is frequently conceptualized as involving a feedback loop, often a feedback loop that works to reduce the distance between some desired goal state and the present situation. Carver and Scheier (1990) developed a theory wherein a second, "meta-monitoring" loop tracks the rate of progress of this first loop and influences mood. Specifically, if the rate of progress is slower than desired, negative affect occurs; if faster, postive affect occurs. If the rate of progress is at the desired pace, no affect occurs. Over the years, Carver and collegues have tested and refined this theory. One important addition has been the distinction between moods induced by negative feedback loops utilized when trying to approach a desired goal and those induced by positive feedback loops utilized when trying to avoid an undesired situation (see Carver 2004; 2006). According to this new distinction, somewhat slower than expected progress in approach is typically associated with frustration and anger, along with increases in effort. Carver (2004) suggests that these emotions may therefore act as cues that progress is slower than desired, but that there is the potential to correct the problem with greater exertion. In contrast, larger discrepancies in approach are more often associated with feelings of sadness and depression, along with reductions in effort, perhaps as a result of the emotions acting as cues that further efforts would be futile and the more prudent choice would be to conserve resources. Finally, slower than expected progress in avoiding an undesirable result normally produces feelings of anxiety and fear.

Thus, a number of theoretical connections have been made between mood and self-regulation. A number of experiments have also demonstrated that self-regulation can indeed be influenced by mood. For instance, Schwarz and Pollack (1977) found that children who were put in a positive mood were more likely to delay gratification than were children who were put in a negative mood. As stated above, Tice et al. (2007) also found that positive mood inductions can enhance previously depleted self-regulation abilities. Tice, Bratslavsky, and Baumeister (2001) further concluded that negative moods prompt participants to focus on self-regulating their emotions rather than their behavior (e.g., eating), causing them to struggle with behavioral self-regulation.

Some evidence also indicates that self-regulation can impact one's psychological state. In a series of experiments, Fischer, Greitemeyer, and Frey (2007) demonstrated that depleted participants were less optimistic about their own abilities than were non-depleted participants. They further found that this seemed to stem from the depleted participants' overall lower self-efficacy and their calling to mind less positive aspects of themselves. That is, participants' experiences of self-efficacy and the availability of positive self-characteristics mediated the effect of initial self-regulation on their views of themselves. The researchers, however, did not test the relationship between these feelings and any subsequent behavioral measures of self-regulation.

In sum, there is, therefore, clear theoretical and empirical evidence connecting self-regulation and mood. However, proponents of the self-regulation resource theory have consistently argued that depletion effects are due solely to the expenditure of the limited self-regulation resource rather than to other factors such as general fatigue or negative mood states. To demonstrate the lack of involvement of mood, many researchers have asked participants to complete mood inventories after completing their initial depleting or non-depleting task. Inventories used have included the Brief Mood Introspection Scale (BMI/BMIS; e.g., Baumeister et al., 1998; Muraven et al. 1998), the Positive and Negative Affect Schedule (PANAS, e.g., Schmeichel et al., 2003; Vohs et al., 2008) and the UWIST Mood Adjective Checklist (e.g., Gailliot et al., 2007), as well as shorter measures developed for particular studies (e.g., Oaten et al., 2008). Consistent with the self-regulation resource theory account, most researchers found no significant differences in post-task mood between their depleted and non-depleted groups. A few, however, did find significant negative impacts on mood (e.g., Alberts et al., 2007; Ciarocco et al., 2001; Vohs & Schmeichel, 2009), and others have reported effects that, while non-significant, showed trends in the expected direction (e.g., Baumeister et al., 2005; Schmeichel, 2007). Indeed, a recent meta-analysis (Haggar, Wood, Stiff, & Chatzisarantis, 2010) has confirmed that self-regulation manipulations do have a small but significant impact on negative affect (average effect size of .14 across 36 experiments). In the studies where significant mood effects were detected, however, additional analyses did not indicate that mood acted as a mediator of the relationship between the depletion manipulations and the dependent measures (e.g., Ciarocco et al., 2001; Schmeichel & Vohs, 2009, cf. Alberts et al., 2007). Thus, while self-regulation impacted mood in these cases, that impact was not thought to explain the effect of the initial self-regulation on the effectiveness of subsequent self-regulatory attempts.

Study 1 Overview

Based on the previous research on self-regulation depletion, we chose two main avenues of further exploration regarding the phenomenon to investigate in our initial experiments (i.e., Study 1). First, as stated above, previous research has suggested some promising ways to counteract depletion. However, numerous additional options have yet to be explored, especially regarding the various tasks that could be engaged in during a "break" between self-regulation-intensive activities. Establishing how the nature of intervening tasks impacts subsequent self-regulation performance may provide further insights into the exact nature of the self-regulatory depletion phenomenon. Determining which activities are restorative would also have strong practical applications, both in the academic world and in industry. For example, teachers could be advised as to how they could optimize the scheduling of their class time to alternate depleting and restorative activities. Employers could likewise determine the best daily routines for employees in order to maximize their performance (Elsbach & Hargadon, 2006).

Second, in our judgment, it appeared that the relationship between self-regulation depletion and mood also needed to be better understood. As summarized above, the potential involvement of mood in self-regulation depletion studies has typically been downplayed, with non-significant differences in post-task mood inventories taken as indication of no effect. Given the potential theoretical importance of mood in relation to self-regulation (for instance, in the theoretical views of Carver and Scheier, 1990, and Schwarz, 2000) as well as Haggar et al.'s (2010) meta-analytic findings of significant impacts of self-regulation on negative affect, we wished to put a greater emphasis on investigating the possibility of small but real mood effects in our studies. In order to do so, we used an experimental design that would be more sensitive to such small changes, specifically by measuring mood both before and after the initial depleting (or non-depleting) task. Our concern was that small mood effects were being overshadowed by individual variability in mood when only post-task measurements were used.

Therefore, with the original goal of examining how various intervening tasks counteracted the depletion effect, as well as further exploring the relationship between self-regulation and mood, we designed and performed a series of studies. As we wished to first ensure that depletion was indeed occurring before we endeavored to counteract it, Studies 1.1-1.4 began by testing participants only in the "control" conditions, i.e., the conditions without any intervening task conditions. These conditions thereby corresponded to what typically would be the depletion and non-depletion conditions in most of the earlier research on self-regulatory depletion.

Furthermore, as our end goal was to determine if particular intervening tasks could counteract depletion, we wished to equate the different conditions on all other aspects. One concern was that by the time they reached the dependent variable measurement task, participants who completed an intervening task would have been in the lab, engaging in activities, for a longer total amount of time than would the controls who did not engage in an intervening task. As a result of this concern, we designed Studies 1.1-1.3 in such a way that there would be a preliminary task, thought to be self-regulation-neutral, that the experimental participants could engage in for a short time and the control participants could engage in for a longer time. Participants then completed an initial task known from previous research to be depleting or a non-depleting control condition. Finally, all participants performed a subsequent self-regulation/executive-function activity.

In each of these studies, participants were told that they were completing a number of short activities in order to pilot the tasks for future research in the lab. This deception was used in order to: *(1)* reduce the likelihood of participants determining the hypotheses of the study, *(2)* provide a good explanation for the need to complete the number and variety of tasks necessary, and *(3)* make the task and mood ratings required less suspect.

Study 1.1

Participants

A total of 42 participants (13 male, 29 female) were conveniently sampled from the University of Minnesota – Twin Cities area using the Research Experience Program (REP) pool. In this program, students receive extra credit for certain introductory courses in exchange for participation in research projects. Additional participants were recruited via postings in the psychology building for those who wished to receive monetary compensation. Participants were required to be between the ages of 18-30 years old, native speakers of English, and to have normal or corrected-to-normal vision; these requirements were enforced for all the experiments reported in this manuscript. The participants in Study 1.1 ranged in age from 18-23 years old (*M* = 18.90, *SD* = 1.30) and represented these racial/ethnic groups: 75.60% Caucasian, 4.90% African or African American, 17.10% Asian or Asian American, and 2.40% Hispanic.

Materials

In Study 1.1, the preliminary task, implemented in order to be able to equate the total time in the study session for experimental and control participants, required participants to watch a video of waves breaking on the shore obtained from YouTube® (droolpurr, 2007). The video stimulus was simple including only visuals and sounds of the waves; no other features (e.g., people, animals, vehicles) were present.

In order to manipulate the usage of self-regulation resources, two different tasks[4] were used – (1) an inhibition of typical writing habits task, or (2) a working memory updating task – with a depleting and non-depleting[5] version of each. The first task, used successfully by previous self-regulation depletion researchers including Schmeichel (2007), required participants to write a story about a trip they had taken recently. The non-depleting version included only those simple instructions. The depleting version included the additional restriction that participants were not allowed to use the letters “a” or “n” anywhere in their stories and needed to instead use only words that did not contain

[4] The use of two depletion manipulations was implemented because, originally, we had hoped to look at an interaction between the type of task that produced the depletion and the type of task that might be replenishing. The writing task was chosen as our generative task, requiring participants to create something. The working memory updating task was chosen as a memory task, requiring participants to learn new information.

[5] Throughout this manuscript, the task versions thought to put fewer demands on the self-regulation resource will be referred to as the non-depleting versions. However, I acknowledge that few, if any, tasks are completely free of demands on self-regulation.

those letters. Thus, the depleting version required participants to inhibit their habitual word choices and to determine an alternative way of expressing their ideas.

The second task, also used successfully by Schmeichel (2007), involved evaluating mathematical equations and recalling sets of words. In the depleting version, participants engaged in an operation span (OpSpan) task (Turner & Engle, 1989). Participants were simultaneously visually presented with an equation and an accompanying word (e.g., (9 x 2) – 3 = 16 ??? hat) using EPrime® software. They were required to first read and evaluate the accuracy of the equation and then to read and remember the word. Participants continued this process of evaluation and word reading until they were presented with a screen giving a memory recall cue, at which time they were to write down all of the words they could recall from that set. Set sizes varied randomly and ranged from 2-5 equation-word pairs before a recall screen was introduced. A total of 27 equation-word pairings were used. The stated mathematical equations were correct approximately 50% of the time: 13 of the 27 trials were correct and 14 of the 27 trials were incorrect. In the non-depleting version, participants completed the memory and math aspects of the operation span task separately. Specifically, they were first shown sets of sequentially presented words (e.g., back, fish, bread, bush, cheek, jar) and asked to recall them. This version included the same 27 words as the depleting version, but here the words were divided into four sets of six and one set of three. Participants then separately evaluated the accuracy of the 27 math equations (e.g., (8 x 4) + 2 = 32).

The primary dependent measure of self-regulation, administered after the participant completed their assigned depleting or non-depleting initial task, was a reasoning task from the Graduate Records Exam (GRE), previously found by Schmeichel et al. (2003) to be useful in measuring self-regulation depletion. In this task, participants were given 13 items from the analytical subtest of the GRE, requiring complex relational problem solving, and were asked to complete (correctly) as many as possible within the time limit. An example question is as follows:

> The primary schools in a city range from one to six stories in height. If a classroom in a primary school is above the second floor, it must have a fireproof door.
>
> If the statements above are true, which of the following statements must also be true about primary-school rooms in the city?
>
> (A) Some third-floor rooms in primary schools do not have fireproof doors.
> (B) No second-floor classrooms in primary schools have fireproof doors.
> (C) In primary schools, rooms above the second floor that are not classrooms do not have fireproof doors.
> (D) Any fourth-floor classrooms in primary schools have fireproof doors.
> (E) Primary schools with classrooms on the first floor only do not have any fireproof doors.
> Ans: D

In addition to this primary dependent measure of self-regulation/self-regulation depletion, we also measured participants' moods as well as their perceptions of the tasks throughout the study. To measure mood, the Self-Assessment Manikin (SAM) was used (Lang, 1980; Hodes, Cook, & Lang, 1985). This measure requires participants to rate their mood valence (positive to negative emotions) and mood arousal level (calm to excited) using simple pictorial rating scales. As is common with this measure, participants were asked to rate their feelings on each dimension by placing an X either on one of the five pictographs or between two, thus creating 9-point scales. The SAM measure was selected because it has been well validated in previous research (e.g., Bradley & Lang, 1994) and because its brevity and straightforward nature facilitated our ability to make repeated measurements of mood. To gauge participants' subjective experiences of the tasks, they were asked to rate the pleasantness, difficulty, and "frustratingness" for each task on 10-point scales. These measures were included to provide a manipulation check of the experimental interventions, as well as to reduce demand characteristics by fostering the impression (consistent with the cover story) that we were simply piloting the various tasks for use in future research.

Procedure

All participants first watched the waves video for 8 minutes[6], and then rated this video watching task on the various dimensions (e.g., pleasantness). Next, they all completed an initial SAM measurement, providing both valence and arousal ratings.

Participants were then randomly assigned to one of four conditions, dictating which of the initial depleting/non-depleting tasks they would complete: (1) Depleting Writing Task, (2) Non-depleting Writing Task, (3) Depleting OpSpan Task, or (4) Non-Depleting Memory and Math Task. Once the participant understood the instructions, experimenters timed those completing the writing tasks and stopped them after they had been writing for seven minutes. The OpSpan/Memory and Math tasks were more self-paced, with no time limits given for participants to complete the recall portions; however, participants took approximately seven minutes to complete these tasks as well. After participants completed their assigned task, they were asked to rate it on the various dimensions, and to complete another SAM.

[6] The plan was for the experimental participants (i.e., replenished participants) – none of which were run due to the results obtained from this initial phase of only control participants – to watch the video for two minutes and then, after the depleting task had been administered, to engage in their intervening task for 6 minutes.

Participants were then given the self-regulation/self-regulation depletion-measuring GRE Analytical Problems Task. In order to ensure adequate motivation for this task, participants were reminded of the importance placed by society on problem-solving abilities, being told something like, “Problem-solving abilities are highly valued by those who look over graduate school and job applications.” They were also informed that they had 10 minutes to correctly answer as many problems as they could. After the participants had completed the subsequent, final, self-regulation task, they were debriefed with an explanation regarding the true purpose of the study and provided with their compensation.

Results[7]

Depletion. As this study was designed to determine if completing the depleting tasks would lead to significantly poorer performance on the subsequent GRE task than completing the non-depleting tasks, 2-way ANOVAs were first conducted using depletion condition (depleting vs. non-depleting) and task type (writing vs. memory and math) as the between-subjects factors. Three dependent variables were assessed: the number of GRE items participants answered, the number of items participants answered correctly, and the percentage of the items that the participants answered that they were able to answer correctly. Means and effect sizes (Cohen’s *d*) can be found in Table 1.

[7] For each study, all data were first screened for outliers using box-plots, i.e., outliers were defined based on quartiles. Specifically, the “box” was defined as the middle two quartiles (i.e., 50%) of the distribution, and an outlier was defined as at least 1.5 box lengths above or below the outer edges of that box. With the data first categorized by experimental group, the screening for and removal of outliers was done separately for each dependent variable. Thus, the final Ns/dfs reported for each analysis differ somewhat, depending on the number of outliers removed for that variable. Across all of the studies, the minimum percentage of outliers removed for a variable was 0% (numerous instances) and the maximum percentage of outliers removed for a variable was 11% (SAM pre-task valence in study 1.2), with a mean removal rate of 2.52% and a median of 1.82%.

Table 1. Depletion means and effect sizes for Study 1.1

Depletion Variable	Condition	*N*	*M*	*SD*	Effect sizes: Depletion condition	Effect sizes: Task type x Depletion condition
GRE: # attempted	Depleted	21	8.52	2.73	-.22	.02
	Writing	10	7.70	2.98		
	Memory/Math	11	9.27	2.37		
	Non-depleted	21	7.81	3.20		
	Writing	11	7.00	2.41		
	Memory/Math	10	8.70	3.83		
GRE: # correct	Depleted	21	4.57	1.99	-.54	.03
	Writing	10	4.50	2.27		
	Memory/Math	11	4.64	1.80		
	Non-depleted	20	3.55	1.73		
	Writing	10	3.56	1.74		
	Memory/Math	10	4.30	2.87		
GRE: % correct	Depleted	20	53.84	21.54	-.21	.08
	Writing	10	59.15	22.83		
	Memory/Math	10	48.52	19.88		
	Non-depleted	20	49.64	22.56		
	Writing	11	52.89	23.47		
	Memory/Math	9	45.66	22.08		

Number answered. There was no main effect of depletion condition, $F(1,38) = .49$, $p = .49$, with means actually trending in the opposite direction than expected ($M_{\text{depleted}} = 8.52$, $M_{\text{non-depleted}} = 7.81$). There was also no interaction between type of task (writing vs. memory and math) and depletion condition, $F(1,38) = .005$, $p = .94$.

Number correct. Again, there was no main effect of depletion condition, $F(1,37) = 2.87$, $p = .099$, with trends toward an effect in the opposite direction ($M_{\text{depleted}} = 4.57$,

$M_{\text{non-depleted}} = 3.55$). There was also no interaction between type of task and depletion condition, $F(1,37) = .01$, $p = .92$.

Percent of those answered that were correct. Once more, there was no main effect of depletion condition, $F(1,36) = .42$, $p = .52$, with means in the opposite direction ($M_{\text{depleted}} = 53.84$, $M_{\text{non-depleted}} = 49.64$), and no interaction between type of task and depletion condition, $F(1,36) = .06$, $p = .81$.

Mood. Next, we assessed the effects of the depleting/non-depleting tasks on participants' moods, as indicated by their SAM valence and arousal ratings. For each dimension, a 2-way ANOVA was conducted using depletion condition (depleting vs. non-depleting) and task type (writing vs. memory and math) as the between-subjects factors to gauge their impact on the *post-task* SAM ratings. We then conducted a 3-way ANOVA for each dimension, adding time of measurement (ratings from before vs. after the depleting/non-depleting task) as a within-subjects factor. Means and effect sizes can be found in Table 2.

SAM valence. A significant main effect of depletion condition (depleting vs. non-depleting) was found, $F(1,38) = 6.71$, $p = .014$; participants' moods were significantly more negative after the depleting tasks ($M = 5.05$, $SD = 1.88$) than after the non-depleting tasks ($M = 3.81$, $SD = 1.33$).[8] There was also a marginal interaction effect between depletion condition and task type (writing vs. memory and math) $F(1,38) = 3.41$, $p = .073$, with a trend for participants' mood valence to be more affected by the writing task versions than by the memory and math task versions.

[8] Note that higher scores on valence actually indicate a more negative mood.

Table 2. Mood means and effect sizes for Study 1.1

Mood Variable	Condition	*N*	*M*	*SD*	Effect sizes: Depletion condition	Effect sizes: Depletion condition x Time of measurement	Effect sizes: Depletion condition x Time of measurement x Task type
SAM: Valence	**Before**						
	Depleted	18	3.72	1.23	-.06		
	Writing	8	4.38	.92			
	Memory/Math	10	3.20	1.23			
	Non-depleted	21	3.86	1.06			
	Writing	11	3.91	1.30			
	Memory/Math	10	3.80	.79			
	After						
	Depleted	18	5.11	1.84	.82	.65	.30
	Writing	8	6.13	1.25			
	Memory/Math	10	4.30	1.89			
	Non-depleted	21	3.81	1.33			
	Writing	11	3.64	1.69			
	Memory/Math	10	4.00	.82			
SAM: Arousal	**Before**						
	Depleted	19	8.00	.94	-.23		
	Writing	9	8.11	.93			
	Memory/Math	10	7.90	.99			
	Non-depleted	21	7.76	1.48			
	Writing	11	7.73	1.79			
	Memory/Math	10	7.80	1.14			
	After						
	Depleted	19	5.42	1.64	.68	.61	.33
	Writing	9	6.11	1.83			
	Memory/Math	10	4.80	1.23			
	Non-depleted	21	6.19	.93			
	Writing	11	6.18	.98			
	Memory/Math	10	6.20	.92			

Note: Higher scores on valence actually indicate a more negative mood, and higher scores on arousal actually indicate a calmer mood.

Additionally, there was a significant interaction of depletion condition and time of measurement (before vs. after depleting/non-depleting task), $F(1,35) = 6.19$, $p = .018$, with the non-depleting versions having little effect on mood, but the depleting versions negatively impacting mood valence (see Figure 1). No significant three-way interaction (depletion condition x task type x time) was found, $F(1,35) = .91$, $p = .35$, reflecting the consistency of the pattern of change in participants' mood valence from pre- to post-task due to the degree of self-regulation required on the initial task (depleting vs. non-depleting), regardless of the nature of the task (writing vs. memory and math).

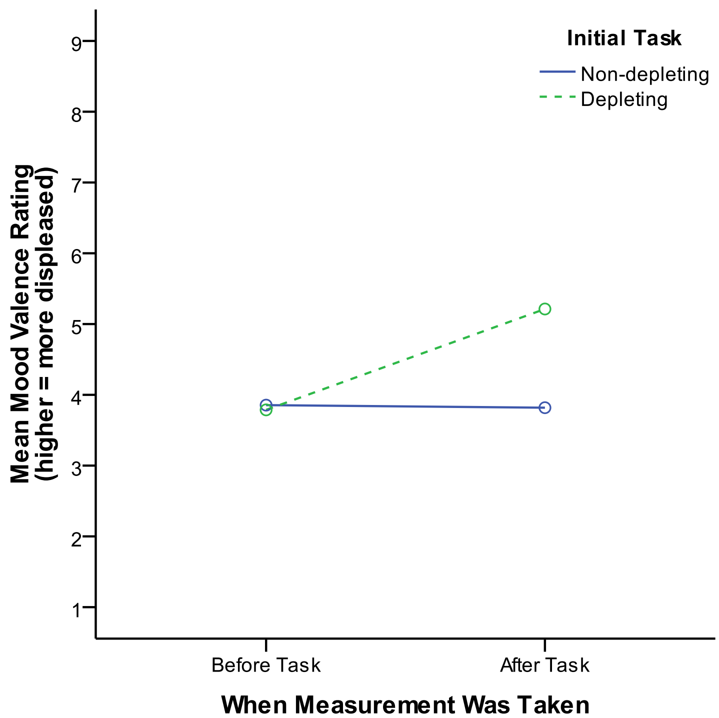

Figure 1. Mood valence by depletion condition (whether participants completed a depleting or non-depleting initial task) and time of measurement in Study 1.1.

SAM Arousal. For this dimension as well, a significant effect of depletion condition was found, reflecting less arousal for the non-depleting ($M = 6.19$, $SD = .93$)

than the depleting (*M* = 5.30, *SD* = 1.69) versions[9], *F*(1,37) = 4.52, *p* = .04, with no significant interaction effect between depletion condition and task type, *F*(1,37) = 1.48, *p* = .23. There was also a marginally significant interaction of depletion condition and time of measurement, *F*(1,36) = 3.35, *p* = .075, with the non-depleting versions tending to increase arousal less than the depleting versions (see Figure 2). No significant three-way interaction (depletion condition x task type x time) was found, *F*(1,36) = .96, *p* = .33.

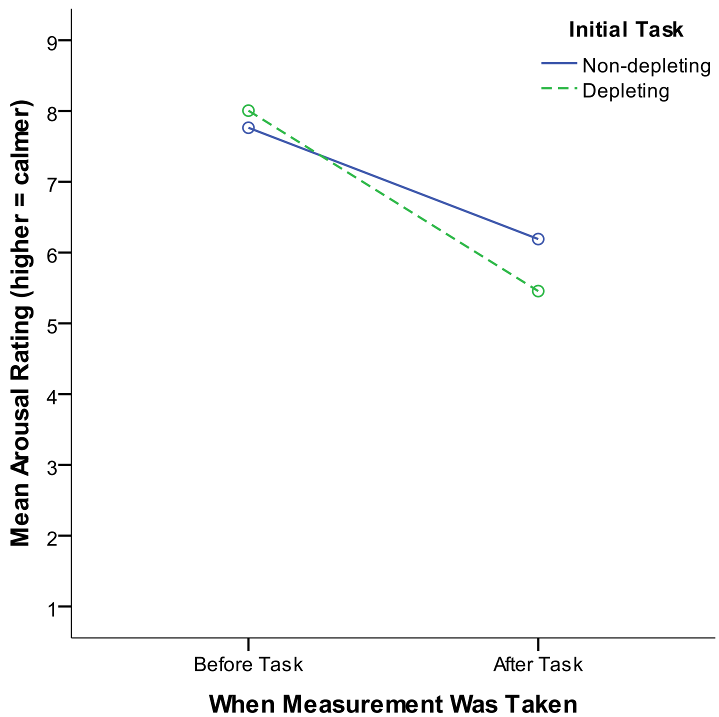

Figure 2. Mood arousal by depletion condition and time of measurement in Study 1.1.

Task Ratings. Because of the surprising lack of depletion effects on the main GRE measure, we analyzed participants' responses to the task ratings for both the initial depleting/non-depleting tasks and the preliminary waves video task. Participants who

[9] Note that higher scores on arousal actually indicate a more calm mood.

completed the depleting versions rated their task as significantly less pleasant $F(1,38) = 6.21$, $p = .017$, $d = .81$, more difficult, $F(1,36) = 48.18$, $p < .001$, $d = 2.31$, and more frustrating, $F(1,35) = 22.18$, $p < .001$, $d = 1.59$, than those who completed the non-depleting versions. This led support to the belief that these tasks truly do create subjectively different experiences for the participants.

Regarding the new waves video-watching task, the ratings data indicated that participants found it to be relatively pleasant ($M = 7.07$, $SD = 1.94$ on a scale of 1 = very unpleasant and 10 = very pleasant) and easy ($M = 1.59$, $SD = 0.82$ on a scale of 1 = very easy and 10 = very difficult), and their arousal ratings afterward were generally relaxed ($M = 7.90$, $SD = 1.24$ on a scale of 1 = very agitated and 9 = very relaxed). However, participants also reported that watching the video was mildly frustrating ($M = 2.82$, $SD = 2.13$ on a scale of 1 = not at all frustrating and 10 = very frustrating). The somewhat larger standard deviation here also suggested that there was more variability in participants' ratings on this dimension, with some considering it non-frustrating and others rather frustrating, as they needed to sustain their attention to the monotonous film.

Study 1.2

Study 1.1 provided some interesting data on the effects of self-regulation tasks on mood but failed to demonstrate any depletion effects. Indeed, there were trends for the GRE performance measures in the opposite direction. Therefore, we made changes to the experimental design and developed Study 1.2 with the goal of producing clear depletion effects. We also wished to determine if our novel mood findings would be replicated under different experimental conditions.

Because participants' task ratings suggested that watching the waves video produced some frustration, we were concerned that the preliminary task may have taxed the self-regulation resource of *all* the participants, diminishing the potential for our experimental manipulation of depletion condition to differentially exhaust the resource. The fact that the video induced such a highly relaxed state of mind also appeared to have the potential to be problematic (e.g., by influencing participants' effort levels or processing styles on the initial depletion manipulation task). As a result, we replaced the waves video with a word finds task for Study 1.2. In addition, there was some concern that emphasizing the importance placed by society on problem-solving abilities directly before the GRE task actually produced *too much* motivation and was counteracting any depletion effects, as had been reported by Muraven and Slessareva (2003). Therefore, the importance of problem-solving abilities was no longer emphasized in Study 1.2.

Participants

A new convenience sample of 80 participants (24 male, 56 female) from the University of Minnesota – Twin Cities area was recruited using the REP pool and paper postings for those who wished to receive monetary compensation. Participants in this phase ranged in age from 18-30 years old (M = 19.67, SD = 2.81) and represented these racial/ethnic categories: 72.50% Caucasian, 7.50% African or African American, 15.00% Asian or Asian American, 2.50% Hispanic, and 2.50% who chose "Other" or not to respond.

Materials

Four word finds for the preliminary task were created for this study, one for each of the following words: AND, IN, THE, and TO. Each word find consisted of a grid of capital letters in Courier New font size 13.5 that was 43 letters long by 10 letters high, containing a random mix of the letters of the word the participant was supposed to find. For example, the AND word find contained only the letters "A," "N," and "D" in various configurations. Instructions at the top said: "Please circle all instances of the word "AND" [or "IN" or "THE" or "TO"] that you can find. The words may be in a horizontal, vertical, or diagonal direction. Words written backwards also are permitted. The same letter may contribute to more than one instance. You will have 2 minutes."

All other materials for this study were the same as those used in Study 1.1.

Procedure

All participants first completed the four word finds, spending two minutes on each, for a total of 8 minutes[10] for the task, and subsequently rated the word finds task on the different dimensions. They then all completed an initial SAM. As in Study 1.1, participants were then randomly assigned to one of the four conditions: (1) Depleting Writing Task, (2) Non-depleting Writing Task, (3) Depleting OpSpan Task, or (4) Non-Depleting Memory and Math Task. The procedures for completing these tasks, rating the tasks, and completing a second SAM were retained from Study 1.1. Participants were then given the depletion-measuring GRE Analytical Problems Task. However, unlike in Study 1.1, we no longer emphasized the importance of problem-solving abilities when we gave participants this task. Participants were still informed that they had 10 minutes to

[10] The initial plan was for the later experimental (i.e., "replenished") participants to complete one word find for two minutes and then, after the depleting task had been administered, to engage in their "replenishing" intervening task for 6 minutes. However, this planned "replenished" condition was not run given that the control participants did not demonstrate consistently strong self-regulatory depletion (and so there was not consistent depletion for the replenishment to "counteract").

correctly answer as many problems as they could. After the participants had completed the GRE task, they were thanked, debriefed, and compensated.

Results

Depletion DVs. To assess whether or not previously completing the depleting tasks produced significantly poorer performance on the subsequent GRE task than completing the non-depleting tasks, we again ran ANOVAs using depletion condition (depleting vs. non-depleting) and task type (writing vs. memory and math) as the between-subjects factors to assess the three dependent variables of number of GRE problems answered, number answered correctly, and percent answered correctly. Means and effect sizes can be found in Table 3.

Number answered. This time, there was a significant main effect of depletion condition, $F(1,75) = 5.06$, $p = .027$, with means in the expected direction ($M_{\text{depleted}} = 6.18$, $M_{\text{non-depleted}} = 7.45$). There was also no interaction between type of task (writing vs. memory and math) and depletion condition, $F(1,75) = .76$, $p = .39$.

Number correct. Here too, there was a significant main effect of depletion condition, $F(1,73) = 6.50$, $p = .013$, with means in the expected direction ($M_{\text{depleted}} = 3.58$, $M_{\text{non-depleted}} = 4.59$), and no interaction between type of task and depletion condition, $F(1,73) = .27$, $p = .61$.

Percent correct. However, there was no significant depletion for percent correct, $F(1,76) = .19$, $p = .67$, although the trend of the means was in the expected direction ($M_{\text{depleted}} = 63.06$, $M_{\text{non-depleted}} = 65.22$). Thus, of the items they were able to complete, depleted participants were able work through and answer them as well as non-depleted participants. There was also no interaction between type of task and depletion condition, $F(1,76) = .24$, $p = .63$.

Table 3. Depletion means and effect sizes for Study 1.2

					Effect Sizes	
Depletion Variable	Condition	*N*	*M*	*SD*	Depletion condition	Task type x Depletion condition
GRE: # attempted	Depleted	39	6.18	2.85	.51	.20
	Writing	19	4.84	1.61		
	Memory/Math	20	7.45	3.20		
	Non-depleted	40	7.45	2.69		
	Writing	20	6.65	2.60		
	Memory/Math	20	8.25	2.59		
GRE: # correct	Depleted	38	3.58	1.61	.59	.12
	Writing	19	3.00	1.37		
	Memory/Math	19	4.16	1.64		
	Non-depleted	39	4.59	2.01		
	Writing	19	3.79	1.65		
	Memory/Math	38	3.58	1.61		
GRE: % correct	Depleted	40	63.06	22.75	.10	.11
	Writing	20	63.96	26.21		
	Memory/Math	20	62.16	19.33		
	Non-depleted	40	65.22	21.42		
	Writing	20	63.67	21.32		
	Memory/Math	20	66.77	21.95		

Mood. Again, the effects of the depleting/non-depleting tasks on participants' SAM ratings were assessed via 2-way ANOVAs using depletion condition (depleting vs. non-depleting) and task type (writing vs. memory and math) as the between-subjects factors, as well as via 3-way ANOVAs adding time of measurement (before vs. after the depleting/non-depleting task) as a within-subjects factor. For means and effect sizes, see Table 4.

Table 4. Mood means and effect sizes for Study 1.2

Mood Variable	Condition	*N*	*M*	*SD*	Effect sizes: Depletion condition	Depletion condition x Time of measurement	Depletion condition x Time of measurement x Task type
SAM: Valence	**Before**						
	Depleted	35	3.09	1.10	-.40		
	Writing	18	2.44	.98			
	Memory/Math	17	3.76	.75			
	Non-depleted	32	3.38	.79			
	Writing	13	3.15	.56			
	Memory/Math	19	3.53	.91			
	After						
	Depleted	35	4.49	1.65	.88	1.17	.87
	Writing	18	4.22	1.77			
	Memory/Math	17	4.76	1.52			
	Non-depleted	32	3.47	1.39			
	Writing	13	2.46	.78			
	Memory/Math	19	4.16	1.30			
SAM: Arousal	**Before**						
	Depleted	39	7.08	1.56	-.31		
	Writing	19	7.26	1.49			
	Memory/Math	20	6.90	1.65			
	Non-depleted	40	6.55	1.95			
	Writing	20	7.25	1.83			
	Memory/Math	20	5.85	1.84			
	After						
	Depleted	39	5.41	1.55	.50	.87	.10
	Writing	19	5.21	1.36			
	Memory/Math	20	5.60	1.73			
	Non-depleted	40	6.27	2.12			
	Writing	20	6.75	2.47			
	Memory/Math	20	5.80	1.64			

Note: As before, higher scores on valence actually indicate a more negative mood, and higher scores on arousal actually indicate a calmer mood.

SAM valence. There was a significant main effect of depletion condition (depleting vs. non-depleting), $F(1,72) = 14.18$, $p < .001$, replicating the finding from Study 1.1 that the depleting task adversely affected mood valence. There was also a marginally significant interaction effect between depletion condition and task type (writing vs. memory and math), $F(1,72) = 3.65$, $p = .06$. With these significant effects of depletion condition on mood valence, coupled with significant depletion on two of the GRE measures reported above, we explored the potential for mood valence to mediate the significant depletion effects. However, post-task mood valence was not significantly correlated with either number of GRE problems attempted, r (73) = -.05, $p = .70$, or number correct, r (71) = -.14, $p = .23$. Thus, when post-task mood valence was added as a covariate, depletion effects changed very little and remained significant for both the problems attempted and the number correct measures ($p = .024$, $d = .54$ and $p = .019$, $d = .57$, respectively).

There was also a significant interaction of depletion condition and time of measurement (before vs. after depleting/non-depleting task) $F(1,63) = 21.61$, $p < .001$, with the non-depleting versions having less of a negative impact on changes in mood. We again explored the possibility that these mood effects mediated depletion. Here, *change in mood valence* was marginally significantly correlated with both number of GRE problems attempted, r (64) = -.23, $p = .061$, and number correct, r (62) = -.23, $p = .063$. With change in mood valence added as a covariate, depletion effects were reduced somewhat and dropped to marginal significance for the number of problems attempted ($p = .071$, $d = .46$), but remaining significant for number correct ($p = .026$, $d = .58$).

Finally, there was a significant three-way interaction (depletion condition x task type x time): $F(1,63) = 11.85$, $p = .001$. As can be seen in Figure 3, for participants completing the writing tasks, the depleting version negatively impacted mood valence, whereas the non-depleting version actually boosted it. In contrast, both versions of the memory and math tasks negatively influenced mood valence, but the depleting version did so to a greater degree.

SAM arousal. Here too, there was a significant effect of depletion condition on post-task arousal ratings, $F(1,76) = 4.97$, $p = .029$, with greater post-task emotional arousal in the depleted compared with non-depleted conditions, but with a trend toward an interaction effect between depletion condition and task type, $F(1,76) = 3.06$, $p = .085$. We again assessed whether these mood effects mediated the significant depletion effects. Post-task arousal was again not significantly correlated with number of GRE problems attempted, $r\ (77) = .13$, $p = .26$, although it was marginally correlated with number correct, $r\ (75) = .20$, $p = .08$. However, when post-task mood arousal was added as a covariate, depletion effects were reduced and dropped to only marginally significant for number of problems attempted ($p = .073$, $d = .42$), whereas they remained significant for number correct ($p = .042$, $d = .48$).

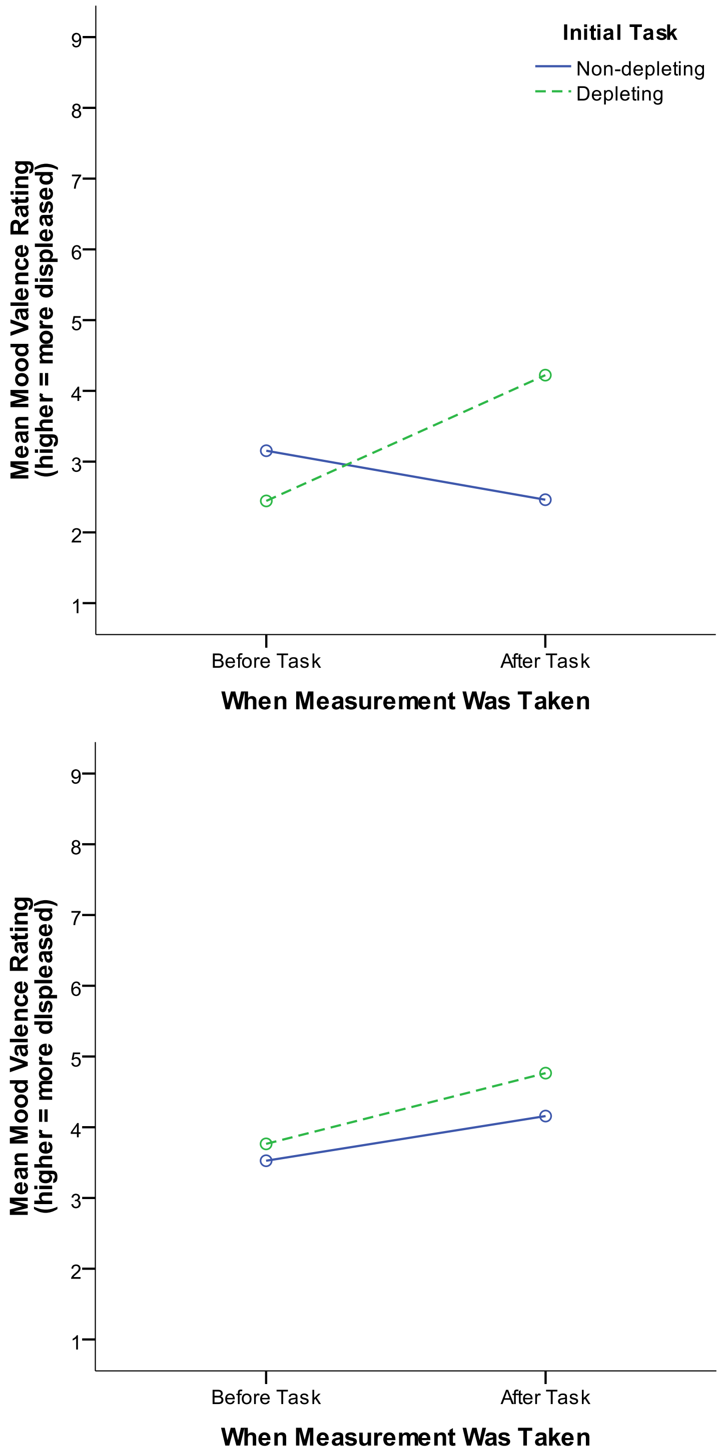

Figure 3: Valence for writing (top) and memory and math (bottom) tasks by depletion condition and time of measurement in Study 1.2.

There was also a significant interaction of depletion condition and time of measurement, $F(1,75) = 14.18$, $p < .001$, with the non-depleting versions increasing emotional arousal less than the depleting versions (see Figure 4). As before, we assessed the mediating potential of these mood effects. Change in arousal was marginally significantly correlated with number of GRE problems attempted, $r(76) = .19$, $p = .09$, and significantly correlated with number correct, $r(74) = .36$, $p = .001$. Furthermore, when change in arousal was added as a covariate, depletion effects dropped considerably and fell below significance for both the number attempted and the number correct measures ($p = .174$, $d = .31$ and $p = .217$, $d = .29$, respectively).

There was no three-way interaction (depletion condition x task type x time): $F(1,75) = .17$, $p = .69$.

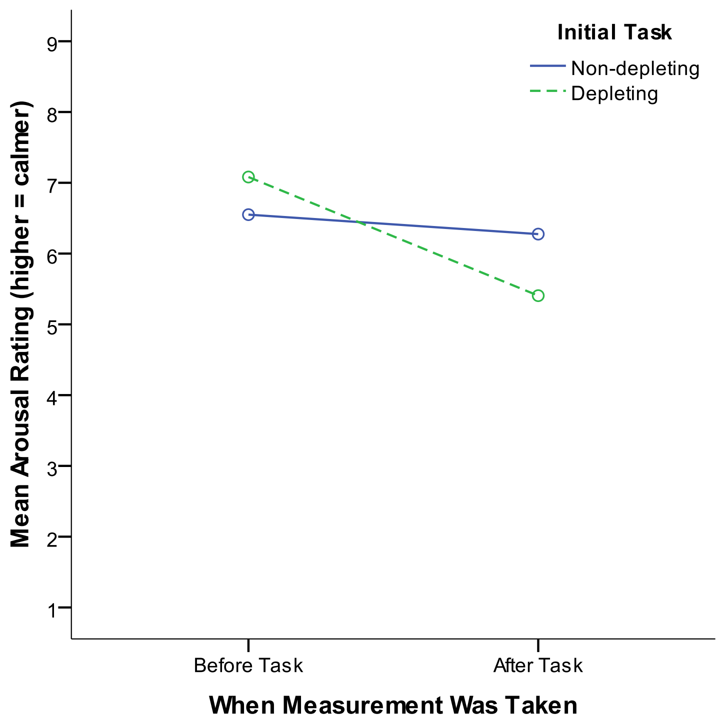

Figure 4. Mood arousal by depletion condition and time of measurement in Study 1.2.

Study 1.3

The results of Study 1.2 were more promising, with significant depletion on two out of the three measures of performance on the GRE task. However, we wanted to be certain we were consistently producing depletion before we began exploring any potential replenishers – so as to ensure that any apparently "replenishing activities" truly were counteracting robust depletion effects. Therefore, we made some slight adjustments to the materials and procedures in Study 1.3 in hopes of boosting the depletion effects. We also wished to further test the effects of self-regulatory depletion on mood valence and mood arousal.

Participants

In all, 80 participants (27 male, 53 female) were conveniently sampled from the University of Minnesota – Twin Cities area through the REP pool. These participants ranged in age from 17-28 years old (M = 19.24, SD = 1.72) and racially consisted of 81.30% Caucasians, 3.80% Africans or African Americans, 11.30% Asians or Asian Americans, and 2.50% Hispanics, and 1.30% who chose "Other" or chose not to report their race.

Materials

We retained the word finds preliminary task; however, we changed the search words to BUT, HOW, THE, and YES. Our motivations were twofold. First, we felt that it may be important to avoid the letters "A" and "N," so as to prevent any priming of those letters for participants who were assigned to complete the depleting writing initial task. Second, we felt that it would be best to keep all of the search words consistently 3 letters in length, as the word finds for the two-letter words (IN and TO) included a rather overwhelming number of instances of the words.

Additionally, for the task ratings measurements, we added two questions to gauge participant's perceived autonomy during task performance. The questions were based on the Intrinsic Motivation Inventory (e.g., Ryan, 1982) used by Muraven et al. (2008) in their study of autonomy as a way to counteract self-regulation depletion. In expectation of running participants in potentially replenishing intervening tasks, we were interested in examining their perceived autonomy. Participants were asked to rate the statements "I felt I had to do this task" and "I felt pressure during the task" on 10-point scales from "strongly disagree" to "strongly agree". The materials used for the initial tasks, SAM, and the GRE task, were all the same as those in Studies 1.1 and 1.2.

Procedure

As in Study 1.2, all participants first completed the four word finds, spending two minutes on each, then rated the word finds task, and then completed the initial SAM. Thereafter, participants were randomly assigned to one of the four conditions: (1) Depleting Writing Task, (2) Non-depleting Writing Task, (3) Depleting OpSpan Task, or (4) Non-Depleting Memory and Math Task. The procedures for completing and then rating these tasks and completing a second SAM were retained from Studies 1.1 and 1.2.

Participants were then given the depletion-measuring GRE Analytical Problems Task. In this study, however, participants were given 15 minutes, rather than 10 minutes, to work on the GRE task. The increased duration was implemented in hopes of preventing any floor effects in terms of number attempted or correct (e.g., our participants in Studies 1.1 and 1.2 were managing to correctly answer only about 4 of the 13 items; Schmeichel et al., 2003, used the same GRE task and their non-depleted participants obtained higher scores than ours did on all three measures). Furthermore, participants in this study were *not* informed of the time limit when they were given the materials. We chose not to notify them of the time limit in order to reduce guessing at the last minute, which could have the potential to mask depletion effects when assessing the number of problems completed (e.g., if depleted participants guessed on more items). Also prompted by this idea that depleted participants might use different strategies than non-depleted participants on the GRE task, we developed one final measure. After participants were finished working on the problems, we asked them to provide feedback on their problem-solving strategies by writing a letter next to each problem on the GRE materials: *W = "I worked through this problem completely until I determined the correct response," E = "I eliminated some options for this problem, and then guessed between the remaining options,"* or *G = "I made a complete guess when answering this problem."* We hoped that this would allow us to determine how far they worked through each problem before answering it. After the participants had completed these ratings, they were debriefed and compensated.

Results

Depletion. As before, the potential effects of depletion were assessed via ANOVAs using depletion condition (depleting vs. non-depleting) and task type (writing vs. memory and math) as the between-subjects factors for the dependent variables of number of GRE problems answered, number answered correctly, and percent answered correct. Means and effect sizes can be found in Table 5.

Table 5. Depletion means and effect sizes for Study 1.3

					Effect Sizes	
Depletion Variable	Condition	*N*	*M*	*SD*	Depletion condition	Task type x Depletion condition
GRE: # attempted	Depleted	39	10.54	2.36	.05	.47
	Writing	20	9.70	2.49		
	Memory/Math	19	11.42	1.90		
	Non-depleted	40	10.45	2.77		
	Writing	20	10.75	2.73		
	Memory/Math	20	10.15	2.85		
GRE: # correct	Depleted	40	6.10	2.48	.39	.14
	Writing	20	5.85	2.52		
	Memory/Math	20	6.35	2.48		
	Non-depleted	40	7.10	2.72		
	Writing	20	7.20	2.51		
	Memory/Math	20	7.00	2.97		
GRE: % correct	Depleted	40	60.27	22.83	.36	.02
	Writing	20	61.20	23.19		
	Memory/Math	20	59.34	23.03		
	Non-depleted	40	68.14	21.01		
	Writing	20	69.51	21.49		
	Memory/Math	20	66.77	20.98		

Number answered. No depletion effect was found for the number of GRE items participants completed, $F(1,75) = .04$, $p = .85$, possibly due to a ceiling effect. However, means were also not in the expected direction ($M_{\text{depleted}} = 10.54$, $M_{\text{non-depleted}} = 10.45$). There was, however, a significant interaction between type of task (writing vs. memory and math) and depletion condition, $F(1,75) = 4.16$, $p = .045$, with participants attempting more problems after the non-depleting writing task, but fewer problems after the non-depleting memory and math task.

Number correct. Here there was a trend toward a depletion effect, $F(1,76) = 2.90$, $p = .093$, with means in the expected direction ($M_{\text{depleted}} = 6.10$, $M_{\text{non-depleted}} = 7.10$). There was no interaction between type of task and depletion condition, $F(1,76) = .36$, $p = .53$.

Percent correct. There was also a trend toward depletion for percent correct, $F(1,76) = 2.52$, $p = .117$, with means in the expected direction ($M_{\text{depleted}} = 60.27$, $M_{\text{non-depleted}} = 68.14$), suggesting that depleting participants may have been having a harder time correctly working through the problems they did complete. There was no interaction between type of task and depletion condition, $F(1,76) = .01$, $p = .93$.

GRE approach ("WEG") ratings. We also explored the possibility that earlier self-regulation may have led participants to give up more easily on problems by analyzing their reported problem-solving strategies (W = worked through completely, E = eliminated some options and then guessed between the remaining options, or G = made a complete guess). The means for these problem-solving approach ratings are presented in Table 6. However, we found no main effects of depletion when assessing the percent of completed items that participants worked through completely, the percent of completed items for which participants used an elimination strategy, or the percent of completed items on which participants guessed completely, *F*s < .31, *p*s > .57. Interactions of depletion condition x task type also did not reach significance, *F*s < 1.81, *p*s > .18.

Table 6. GRE problem solving strategy mean percentages and effect sizes for Study 1.3

Depletion Variable	Condition	*N*	*M*	*SD*	Effect Sizes: Depletion condition	Effect Sizes: Task type x Depletion condition
Worked through completely	Depleted	40	51.12	28.25	.13	.07
	Writing	20	54.18	24.62		
	Memory/Math	20	48.06	31.81		
	Non-depleted	40	54.50	25.68		
	Writing	20	59.40	27.73		
	Memory/Math	20	49.59	23.12		
Elimination strategy	Depleted	40	34.28	20.13	.10	.24
	Writing	20	35.46	17.52		
	Memory/Math	20	33.10	22.86		
	Non-depleted	39	32.15	20.10		
	Writing	20	28.69	23.56		
	Memory/Math	19	35.79	15.50		
Guessed Completely	Depleted	37	10.38	13.20	.07	.32
	Writing	18	5.96	7.84		
	Memory/Math	19	14.57	15.89		
	Non-depleted	38	9.35	11.93		
	Writing	19	8.88	12.80		
	Memory/Math	19	9.82	11.31		

Mood. The effects on mood were again assessed via 2-way ANOVAs using depletion condition (depleting vs. non-depleting) and task type (writing vs. memory and math) as the between-subjects factors, and via 3-way ANOVAs adding time of measurement (before vs. after the depleting/non-depleting task) as a within-subjects factor. The means and effect sizes are presented in Table 7.

Table 7. Mood means and effect sizes for Study 1.3

Mood Variable	Condition	*N*	*M*	*SD*	Effect sizes: Depletion condition	Effect sizes: Depletion condition x Time of measurement	Effect sizes: Depletion condition x Time of measurement x Task type
SAM: Valence	**Before**						
	Depleted	40	3.25	1.39	.06		
	Writing	20	3.00	.86			
	Memory/Math	20	3.50	1.76			
	Non-depleted	35	3.34	1.21			
	Writing	16	3.31	1.20			
	Memory/Math	19	3.37	1.26			
	After						
	Depleted	40	4.18	1.58	.41	.65	.11
	Writing	20	3.75	1.12			
	Memory/Math	20	4.60	1.88			
	Non-depleted	35	3.51	1.12			
	Writing	16	3.44	.73			
	Memory/Math	19	3.58	1.39			
SAM: Arousal	**Before**						
	Depleted	40	6.28	1.92	.01		
	Writing	20	6.40	1.57			
	Memory/Math	20	6.15	2.25			
	Non-depleted	40	6.25	1.89			
	Writing	20	6.65	1.79			
	Memory/Math	20	5.85	1.95			
	After						
	Depleted	40	5.45	2.04	.32	.49	.11
	Writing	20	5.80	1.80			
	Memory/Math	20	5.10	2.25			
	Non-depleted	39	6.08	1.91			
	Writing	20	6.80	1.61			
	Memory/Math	19	5.32	1.95			

Note: As before, higher scores on valence actually indicate a more negative mood, and higher scores on arousal actually indicate a calmer mood.

SAM Valence. There was a significant effect of depletion condition (depleting or non-depleting), $F(1,71) = 4.41$, $p = .039$, indicating higher mood valence among non-depleted participants, with no significant interaction effect between depletion condition and task type (writing vs. memory and math), $F(1,71) = 1.25$, $p = .27$. There was also a significant interaction of depletion condition and time of measurement (before vs. after depleting/non-depleting task), $F(1,71) = 7.39$, $p = .008$, with the non-depleting versions having less of a negative impact on mood valence (see Figure 5). No three-way interaction (depletion condition x task type x time) was found, $F(1,75) = .23$, $p = .64$.

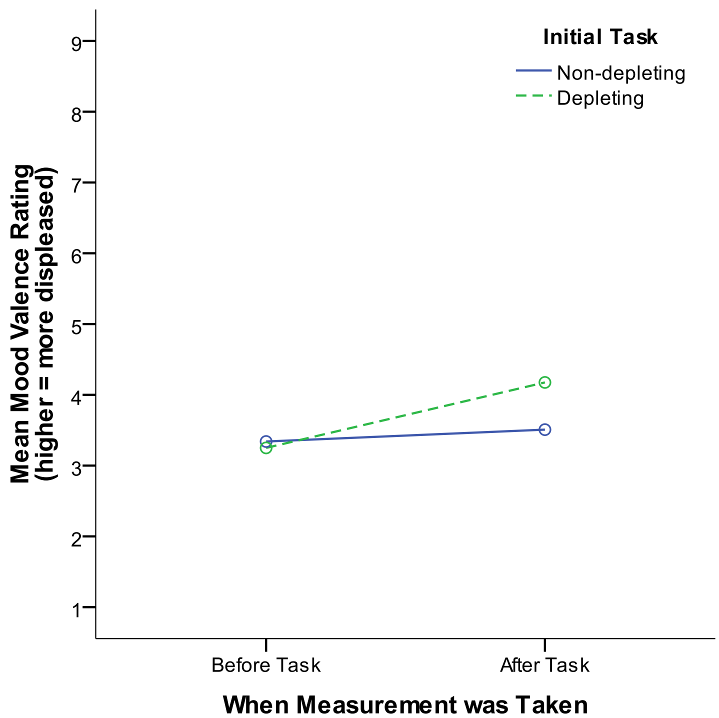

Figure 5. Mood valence by depletion condition and time of measurement in Study 1.3.

SAM Arousal. Here, there was a trend toward an effect of depletion condition, $F(1,75) = 1.99$, $p = .162$, reflecting a tendency toward greater emotional arousal in the depleted compared with non-depleted conditions, with no interaction effect between depletion condition and task type, *F*(1,75) = .83, *p* = .37. There was also a significant interaction of depletion condition and time of measurement, *F*(1,75) = 4.56, *p* = .036, with the non-depleting versions increasing emotional arousal less than the depleting versions (see Figure 6). There was no three-way interaction (depletion condition x task type x time): *F*(1,75) = .24, *p* = .63.

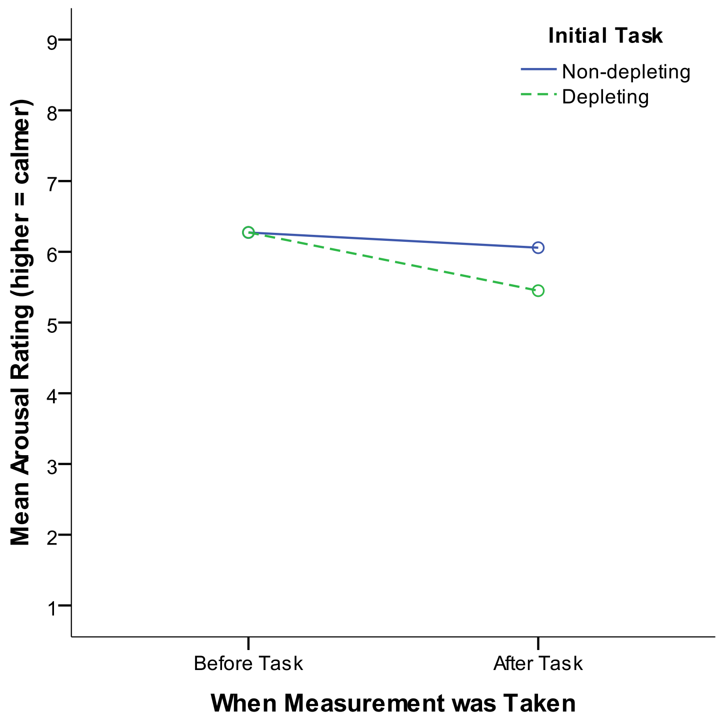

Figure 6. Mood arousal by depletion condition and time of measurement in Study 1.3.

Study 1.4

Our attempts to maximize depletion in Study 1.3 were mainly ineffective. We did see more of an effect of depletion condition on the percent of items participants completed correctly than we had earlier observed in Studies 1.1 and 1.2; however, the effect of depletion condition on this measure of GRE performance still did not reach significance, with effect size *d*s across studies between -.21 and .36. Our efforts to ensure a lack of a floor effect on the GRE task by extending the time limit in Study 1.3 also unfortunately led to somewhat of a ceiling effect instead. Therefore, we returned to the 10-minute time limit for the GRE task in Study 1.4. In addition, we were concerned that, while simple, all of the participants may have been using up some cognitive resources during the preliminary word finds task, dampening the depletion effects that were due to the different initial writing and memory and math tasks. This prompted us to do away with the preliminary task aspect of the design entirely, such that after providing informed consent, and an initial report of mood, participants were immediately assigned to either a depleting or non-depleting task. This procedure, involving no "preliminary" filler task, also more closely mirrors that used in the vast majority of studies on self-regulatory depletion (e.g., Baumeister et al, 1998; Muraven et al, 1998; Schmeichel et al., 2003; Vohs et al., 2008).

Participants

A convenience sample of 55 participants (23 male, 32 female) were recruited from the University of Minnesota – Twin Cities area using the REP pool and postings in the psychology building. The participants in this phase ranged in age from 18-30 years old (*M* = 20.91, *SD* = 2.47) and represented these races: 70.40% Caucasian, 5.60% African or African American, 18.50% Asian or Asian American, 3.70% Hispanic and 1.90% who chose "Other" or chose not to report the information.

Materials

No new materials were used in this study.

Procedure

Rather than a preliminary task, all participants began by completing the initial SAM. Participants were then randomly assigned to one of the four conditions: (1) Depleting Writing Task, (2) Non-depleting Writing Task, (3) Depleting OpSpan Task, or (4) Non-Depleting Memory and Math Task. The procedures for completing and rating these tasks, and completing a second SAM were retained from the previous studies.

Participants were then given the depletion-measuring GRE Analytical Problems Task. In this study, we returned to the 10-minute time limit for this task. As in Study 1.3, participants were not informed of the *length* of their time limit; however, they did know that there would be *a* time limit, as researchers told them that they would let the participants know when it was time to stop working on the problems. We again asked participants to provide us the W/E/G problem-solving strategies ratings after they had completed the task, after which they were debriefed and compensated.

Results

Depletion. As before, possible self-regulatory depletion effects were analyzed via ANOVAs using depletion condition (depleting vs. non-depleting) and task type (writing vs. memory and math) as the between-subjects factors to assess the number of GRE problems answered, number answered correctly, and percent answered correct. Means and effect sizes can be found in Table 8.

Number answered. There was significant depletion on this measure, $F(1,50) = 8.77$, $p = .005$, with means in the expected direction ($M_{\text{depleted}} = 5.19$, $M_{\text{non-depleted}} = 6.70$).

There was no significant interaction between type of task (writing vs. memory and math) and depletion condition, $F(1,50) = 1.41$, $p = .24$.

Table 8. Depletion means and effect sizes for Study 1.4

					Effect Sizes	
Depletion Variable	Condition	*N*	*M*	*SD*	Depletion condition	Task type x Depletion condition
GRE: # attempted	Depleted	27	5.19	1.76	.82	.34
	Writing	13	4.92	1.38		
	Memory/Math	14	5.43	2.07		
	Non-depleted	27	6.70	2.23		
	Writing	14	5.86	2.07		
	Memory/Math	13	7.62	2.10		
GRE: # correct	Depleted	27	3.26	1.53	.64	.04
	Writing	13	3.08	1.38		
	Memory/Math	14	3.43	1.70		
	Non-depleted	27	4.37	1.98		
	Writing	14	4.14	2.18		
	Memory/Math	13	4.62	1.81		
GRE: % correct	Depleted	28	64.69	24.97	.02	.15
	Writing	14	64.99	25.66		
	Memory/Math	14	64.39	25.23		
	Non-depleted	27	65.35	19.24		
	Writing	14	68.73	18.84		
	Memory/Math	13	61.70	19.75		

Number correct. Here, there was also a significant depletion effect, $F(1,50) = 5.31$, $p = .025$, with means in the expected direction ($M_{\text{depleted}} = 3.26$, $M_{\text{non-depleted}} = 4.37$). There was no interaction between type of task and depletion condition, $F(1,50) = .02$, $p = .90$.

Percent correct. There was no depletion on this measure, $F(1,51) = .007$, $p = .93$, with almost equivalent means ($M_{\text{depleted}} = 64.69$, $M_{\text{non-depleted}} = 65.35$). There was also no interaction between type of task and depletion condition, $F(1,51) = .28$, $p = .60$.

Table 9. GRE problem solving strategy mean percentages and effect sizes for Study 1.4

					Effect Sizes	
Depletion Variable	Condition	*N*	*M*	*SD*	Depletion condition	Task type x Depletion condition
Worked through completely	Depleted	28	60.86	32.89	-.32	.55
	Writing	14	59.99	33.67		
	Memory/Math	14	61.73	33.33		
	Non-depleted	26	51.76	26.35		
	Writing	13	66.12	23.85		
	Memory/Math	13	37.41	20.78		
Elimination strategy	Depleted	28	27.47	28.53	.39	.50
	Writing	14	28.15	27.92		
	Memory/Math	14	26.79	30.17		
	Non-depleted	26	37.13	23.26		
	Writing	13	25.71	19.90		
	Memory/Math	13	48.56	21.18		
Guessed Completely	Depleted	27	10.87	14.41	.01	.34
	Writing	14	11.85	15.16		
	Memory/Math	13	9.81	14.08		
	Non-depleted	26	10.68	11.88		
	Writing	13	7.33	11.71		
	Memory/Math	13	14.04	11.51		

GRE approach ("WEG") ratings. As in Study 1.3, we analyzed participants' reported problem-solving strategies (W = worked through completely, E = eliminated some options and then guessed between the remaining options, or G = made a complete guess) in order to explore the possibility that earlier self-regulation may have led participants to use different tactics. The means for these problem-solving approach ratings are presented in Table 9. Main effects of depletion condition were not significant when assessing the proportion of completed items that participants worked through completely, $F(1, 50) = 1.35$, $p = .25$, the proportion of completed items for which participants used an elimination strategy, $F(1, 50) = 1.96$, $p = .17$, or the proportion of completed items on which participants guessed completely, $F(1, 49) = .002$, $p = .97$. Means for the trends were toward depleted participants actually using the working through completely strategy more than those who were not depleted and using the elimination strategy less. There were, however, marginally significant interactions of depletion condition x task type when assessing the working through strategy, $F(1, 50) = 3.79$, $p = .057$, as well as the elimination strategy, $F(1, 49) = 3.08$, $p = .085$. As seen in Figure 7, there was little impact of the depleting vs. non-depleting writing task on participants' problem-solving strategies. However, among the participants who completed the memory and math tasks, the depleted participants tended to be somewhat *more* likely to work through problems completely and *less* likely to use an elimination strategy.

Mood. Effects on mood were again evaluated via 2-way ANOVAs using depletion condition (depleting vs. non-depleting) and task type (writing vs. memory and math) as the between-subjects factors, and via 3-way ANOVAs adding time of measurement (before vs. after the depleting/non-depleting task) as a within-subjects factor. The means and effect sizes are provided in Table 10.

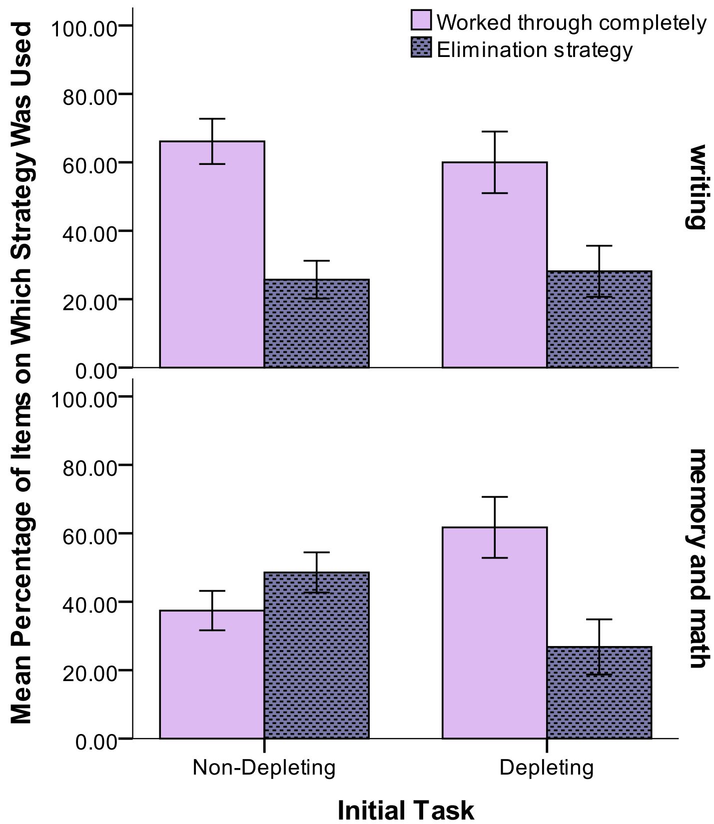

Figure 7. Problem-solving strategies used by participants on the GRE task by task type and depletion condition in Study 1.4. Error bars extend ± one standard error.

SAM Valence. There was no effect of depletion condition (depleting or non-depleting) on post-task emotional valence, $F(1,51) = .014$, $p = .91$. There was a marginally significant interaction effect between depletion condition and task type (writing vs. memory and math), $F(1,55) = 3.86$, $p = .055$, with the depleting writing task producing more negative moods than the non-depleting version, whereas the depleting memory and math task produced better moods than the non-depleting version.

Table 10. Mood means and effect sizes for Study 1.4

Mood Variable	Condition	*N*	*M*	*SD*	Effect sizes: Depletion condition	Effect sizes: Depletion condition x Time of measurement	Effect sizes: Depletion condition x Time of measurement x Task type
SAM: Valence	**Before**						
	Depleted	28	2.79	.83	.63		
	Writing	14	2.79	1.05			
	Memory/Math	14	2.79	.56			
	Non-depleted	26	3.46	1.48			
	Writing	14	3.00	1.18			
	Memory/Math	12	4.00	1.65			
	After						
	Depleted	28	3.25	1.32	.03	.67	.25
	Writing	14	3.43	1.60			
	Memory/Math	14	3.07	1.00			
	Non-depleted	27	3.19	1.55			
	Writing	14	2.64	1.55			
	Memory/Math	13	3.77	1.36			
SAM: Arousal	**Before**						
	Depleted	28	6.36	1.93	.03		
	Writing	14	6.14	2.11			
	Memory/Math	14	6.57	1.79			
	Non-depleted	26	6.27	1.64			
	Writing	14	5.93	1.59			
	Memory/Math	12	6.67	1.67			
	After						
	Depleted	28	6.07	1.88	.12	.28	.06
	Writing	14	5.93	1.98			
	Memory/Math	14	6.21	1.85			
	Non-depleted	27	6.30	2.18			
	Writing	14	6.00	2.29			
	Memory/Math	13	6.62	2.10			

Note: As before, higher scores on valence actually indicate a more negative mood, and higher scores on arousal actually indicate a calmer mood.

In addition, there was a significant interaction of depletion condition and time of measurement (before vs. after depleting/non-depleting task), $F(1, 50) = 5.63$, $p = .022$, reflecting a tendency for participants in the depleting, but not the non-depleting, conditions to show a change towards a more negative mood (see Figure 8).

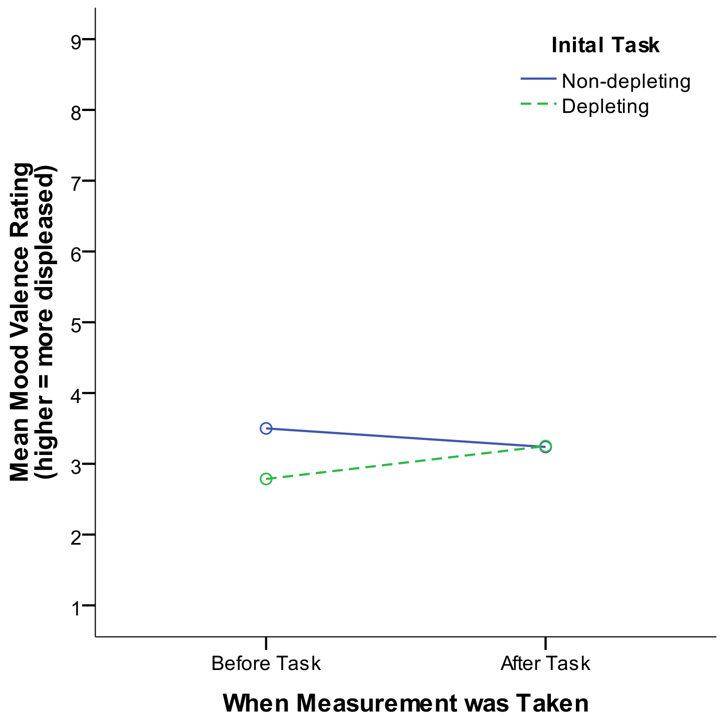

Figure 8. Mood valence by depletion condition and time of measurement in Study 1.4.

As in Study 1.2, when we found both significant depletion and mood effects, we again explored the possibility that the mood effects mediated depletion. Here, change in mood valence was not significantly correlated with either the number of GRE problems attempted, r (51) = -.14, $p = .331$, or number correct r (62) = -.20, $p = .142$. However, when change in mood valence was added as a covariate, depletion effects were reduced.

They remained significant for number of problems attempted (p = .024, d = .66), but dropped below significance for number correct (p = .100, d = .47).

No three-way interaction (depletion condition x task type x time) was found, $F(1,50) = .80$, $p = .38$.

SAM Arousal. There was no effect of depletion condition on emotional arousal, $F(1,51) = .18$, $p = .67$, and no interaction effect between depletion condition and task type, $F(1,51) = .09$, $p = .77$. The interaction of depletion condition and time of measurement was also non-significant, $F(1,50) = .97$, $p = .33$ (see Figure 9 for trends) and there was no three-way interaction (depletion condition x task type x time): $F(1,50) = .04$, $p = .84$.

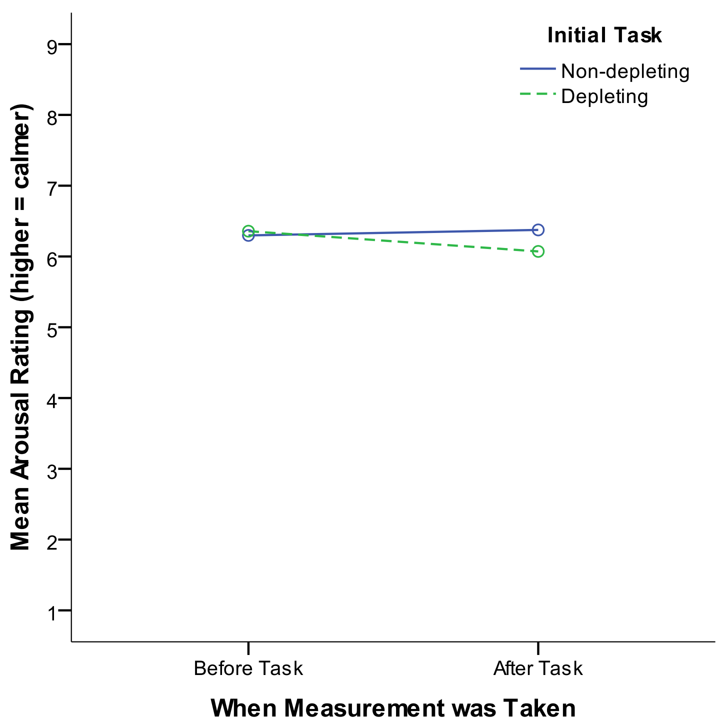

Figure 9. Mood arousal by depletion condition and time of measurement in Study 1.4.

Study 1 Discussion

Overall, the most consistent effects seen in Studies 1.1–1.4 were mood effects. Using only our post-task mood (SAM) measurements, in Studies 1.1–1.3 we consistently found significant differences or trends towards significant differences in both emotional valence and emotional arousal when participants' moods were equated at the beginning of the study, perhaps due to a "common baseline" established via the preliminary tasks used in these studies. Effect sizes (Cohen's *d*) for these analyses were also quite substantial (range: .32-.88; $M_{valence}$ = .70; $M_{arousal}$ = .50), reflecting, overall, a large effect on mood. To our knowledge, previous researchers have not used any preliminary tasks, thus their lack of ability to detect mood effects might be attributed, at least in part, to noisy data due to incidental differences in participants' moods when they arrived at the study. In addition, our paradigm involved a more sensitive design in that we took mood assessments both after the depleting or non-depleting task *and* before the task. Thus, even when participants' moods were not equivalent before starting their tasks, as in Study 1.4, we were still able to assess the degree and direction of the *change* in participants' moods. Overall, the data here indicated that while the non-depleting tasks either produced no change or led to positive effects on mood, the depleting tasks made participants more unhappy and more emotionally aroused (excited/stressed) than they were before the task. Thus, our data provide clear evidence that depletion manipulations can impact mood. This is in line with Carver and Sheier's (1990) theory, which posits mood as being the result of the meta-monitoring of self-regulation processes.

Our results also challenge the claims, made by numerous proponents of the self-regulation resource model (see, for example, Muraven & Baumeister, 2000), that mood does not play a role in depletion effects. In Study 1.2 and Study 1.4, we found both significant depletion effects and significant mood effects. Admittedly, our data did not illustrate much in support of post-task mood valence or arousal as mediators of depletion. However, we did see evidence of small, but clear reductions in depletion effects when *changes* in mood valence and *changes* mood arousal were added to the models (changes in Cohen's *d* for depletion effect when adding changes in valence factor to the model: .01-.17, mean change = .10; changes in Cohen's *d* for depletion effect when adding changes in arousal to the model: .20-.30, mean change = .25). Nevertheless, it is important to note that in some instances, the depletion effects remained statistically significant when changes in mood were added to the models. Furthermore, the size of the impact of mood in our studies indicates that mood's mediating role was often small, with many of the effects of self-regulatory depletion remaining moderately sized with the addition of changes in mood to the model. Therefore, those examining self-regulation depletion are clearly correct to argue that depletion effects cannot be explained away as being simply mood effects.

The second overall conclusion to be drawn from Studies 1.1-1.4 is in regards to the prevalence of the self-regulation depletion effect. Previously published studies often make it seem as though it is a foregone conclusion that completing one self-regulating task will make you less able to complete another (cf. Muraven et al., 2008). However, given the wide range of effect sizes for depletion effects in the current four studies (effect sizes range: -.54 to +.82, *M* = .21), and corresponding inconsistency in whether the depletion effects were statistically significant, it is clear that self-regulation depletion, at least as assessed by the GRE task that was used here, is not an invariable or inevitable outcome. In Study 1.1, no depletion was seen; indeed, trends were in the opposite direction. In Studies 1.2 and 1.4, we saw significant effects of the depletion manipulations using some measures, but none using others. Never did we see depletion on all three measures (number of GRE problems attempted, number of problems answered correctly, and percentage of problems answered correctly given that they were attempted). Additional efforts (in Studies 1.3 and 1.4) to clarify why some measures showed depletion whereas others did not by asking participants to retrospectively report their problem solving strategies (e.g., their tendency to simply guess vs. to systematically work through all of the alternatives) showed little in the way of evidence that depleted and non-depleted participants felt they used different problem solving strategies. Trends were as one might expect in Study 1.3, with depleted participants being slightly less likely to work through problems completely and more likely to guess, but the size of these effects was quite small (d = .13 and .07, respectively). Furthermore, in Study 1.4, there were actually trends toward depleted participants being *more* likely to work through problems (d = .32) and less likely to use elimination strategies (d = .39). In addition, all of the studies used depletion manipulations (the writing task and OpSpan memory and math task) that had been reported as leading to self-regulation depletion in at least some previously published research (e.g., Schmeichel, 2007; Schmeichel & Vohs, 2009) and the sample sizes were typically comparable to those in reported studies using those manipulations. Taken together, these factors suggest that self-regulation depletion is not as prevalent as it may seem. However, the effects of our attempts at implementing a self-regulation “neutral” preliminary task also speak to the issue of the prevalence of depletion, and suggest a different conclusion. Tasks which appear to be extremely undemanding, such as watching a simple waves video in Study 1.1, still appear to require some regulation in order to simply to sustain one’s attention to the mundane task, at least when it is required of participants in an experimental setting.

Study 2 Overview

Given our problems in obtaining consistent depletion effects in Studies 1.1–1.4, we conferred with an expert in the field of self-regulation depletion research (Kathleen Vohs, personal communication, June, 2009). We were advised to change our dependent measure. Although the GRE task we had been using had been used with success by Schmeichel et al. (2003), it had not been validated in much additional self-regulation research. It was recommended that we replace the GRE task with a measure involving completing unsolvable tracing puzzles that has been successfully used numerous times to demonstrate self-regulation depletion (e.g., Baumeister, DeWall, Ciarocco, & Twenge, 2005; Moller, Deci, & Ryan, 2006; Muraven & Slessareva, 2003; Tice, Baumeister, Shmueli, & Muraven, 2007). We implemented this measure in Study 2.1. Encouraged by the belief that changing our dependent measure would produce the clear and consistent depletion effects we were looking for, we also began exploring potential ways to counteract the expected self-regulation depletion in Studies 2.1 and 2.2.

Study 2.1

In Study 2.1, we wished to explore exercise in particular as a potential self-regulation “replenisher.” The benefits of exercise to our physical health are frequently publicized today (see Boreham & Riddoch, 2001; Ignarro, Balestrieri, & Napoli, 2007; and Taylor, Cable, Faulkner, Hillsdon, Narici, & van der Bij, 2004, for reviews), but exercise is often thought of as a way to “clear one’s head” as well. Evidence in support of this position can be found in recent correlational data indicating that children with at least one recess period each day are better able to focus on class work (Barros, Silver, & Stein, 2009), as well as additional experimental research demonstrating that, at least in the long-term, exercise can boost cognitive function, and especially for tasks that place demands on central executive function (e.g., Colcombe & Kramer, 2003; Colcombe et al., 2004). However, exercise had not yet been experimentally examined as a potential “short-term” or immediate way to counteract self-regulation depletion. We were also interested in exploring the restorative properties of exercise in particular because evidence indicating that exercise acts as a self-regulation resource replenisher would also address the biological basis of the resource. Some evidence, as reviewed earlier, indicates that glucose (sugar) seems to be involved (e.g., Gailliot, 2008; Galliott et al, 2007; Denson, et al., 2010; Miller et al., 2010). However, if exercise were found to be replenishing, this would provide strong evidence that additional mechanisms need to be considered, as exercise depletes rather than restores glucose in the body (for a review, see Goodyear & Kahn, 1998).

Participants

A convenience sample of 79 undergraduate students (19 male, 58 female, 2 opted not to report gender) were recruited from the University of Minnesota – Twin Cities psychology REP pool. These participants ranged in age from 18-28 years old (*M* =

19.53, *SD* = 1.78) and the sample was racially composed of 83.80% Caucasians, 5.00% African or African Americans, 6.30% Asian or Asian Americans, and 5.10% who chose "Other" or chose not to report the information.

Materials

The materials for the story-writing task from the previous studies were retained for the initial self-regulation task.

For the measurement of depletion, the suggested puzzle tracing persistence task was used. This task included four geometric shapes that were to be traced in their entirety without lifting one's pencil (see Figure 10). The first puzzle was solvable and served as a demonstration puzzle wherein the researcher showed the participant how to solve it. The second puzzle was also solvable and was given to participants as an initial test of their abilities and to support the illusion that the final two puzzles were also solvable although, in fact, they were not. A desk bell was also used for this task so that the participants could let the researchers know that they had finished the puzzles or wanted to stop trying.

Two intervening tasks were also used to investigate the potentially replenishing properties of exercise. For the exercise task, participants were given a page of instructions regarding how to do four simple exercises as they sat in their chair. The directions were broadly based on those recommended by commercial airlines to counteract physical fatigue and enhance blood circulation during long-distance flights (see http://www.delta.com/traveling_checkin/travel_tips/health/during/index.jsp; http://airplanefootrest.com/exercises.php). The instructions indicated how many of each

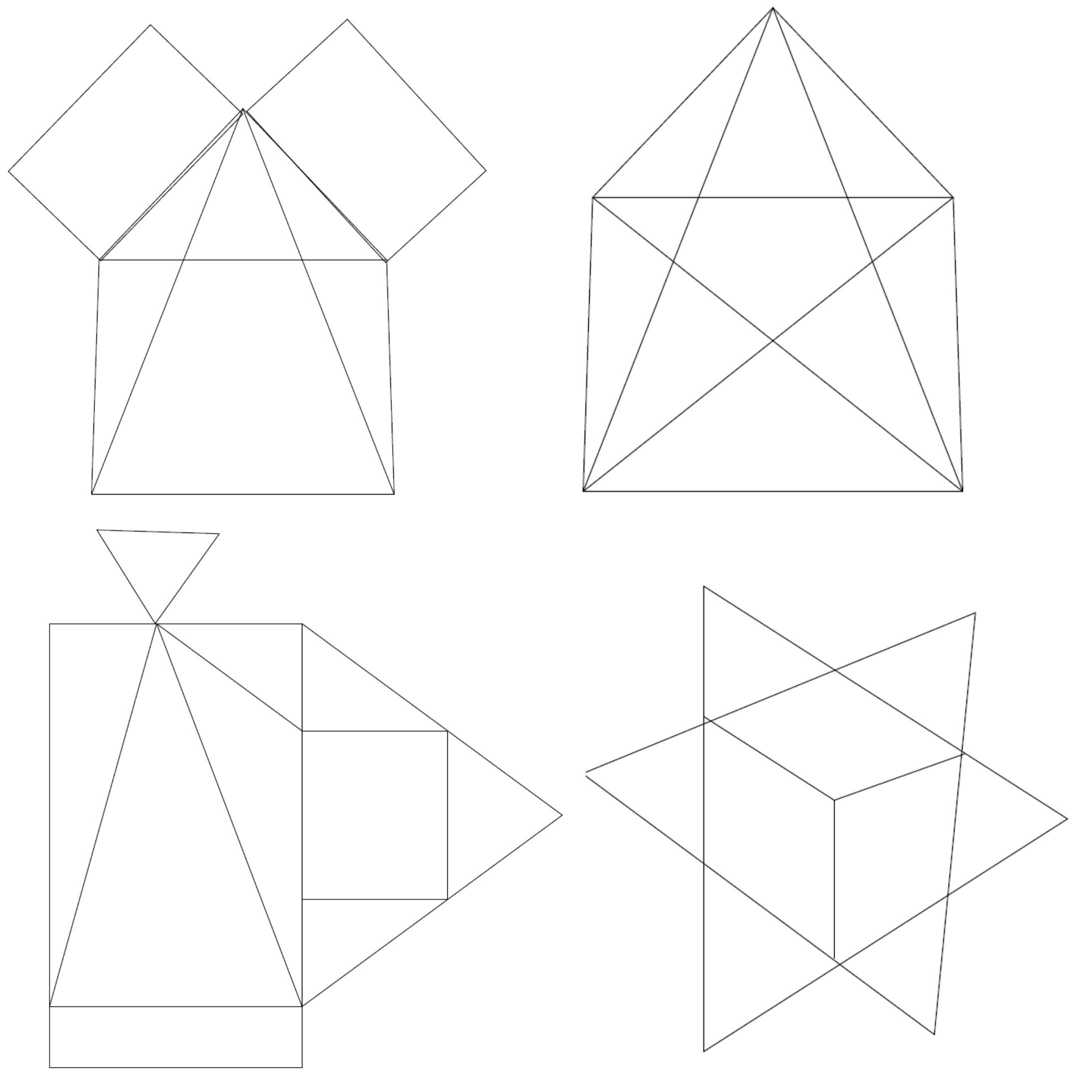

Figure 10. Four shapes used in the puzzle tracing task. On the top left is Shape 1, the demonstration puzzle. On the top right is Shape 2 the easy test puzzle. On the bottom are Shape 3 and Shape 4, the two "difficult" (i.e., unsolvable) test puzzles.

motion to do per set: knee lifts (10 reps/leg), elbow lifts (20 reps), foot lifts (5 reps), and arm lifts (10 reps/arm). The instructions also indicated that the participant should continue completing sets until the experimenter asked them to stop. To act as a control intervening condition, participants were asked to engage in a task that, on the basis of previous findings (Mary Kennedy, personal communication, June, 2009) was thought to be self-regulation "neutral": browsing magazines. To this end, participants were provided with recent copies of People®, Sports Illustrated®, and Self® magazines.

In addition, a brief post-experimental questionnaire was created for this study in hopes of exploring how much of the self-regulation resource participants had left (or perceived themselves as having left) after the final task. On the questionnaire, participants were asked: (1) how long of a break they would want before they would engage in another problem-solving task, (2) how many sugary treats (e.g., M&Ms) they would eat if available, (3) how long they would be willing to continue sitting in the testing room for $5/hr, and (4) how long they would want to engage in an activity (e.g., listening to music, taking a walk) that they found to be helpful in preparing to excel on difficult tasks that require close attention and careful responding, before they would feel "at their best".

Procedure

As in study 1.4, participants began by providing initial SAM ratings. They were then randomly assigned to one of the four conditions: (1) Non-Depleted with Magazine Intervening, (2) Depleted with Magazine Intervening, (3) Depleted with Exercise Intervening, and (4) Depleted with no Intervening Task. Participants then completed their assigned initial depleting or non-depleting writing task as they had in the earlier studies, rated it, and completed their second SAM ratings.

For the participants assigned to the exercise intervening condition, the researcher demonstrated each motion. He/she then asked the participant to follow the instructions as to how many of each motion to do, started a stopwatch, and stopped the participant after they had been exercising for 7 minutes. The participants assigned to a magazine intervening condition were told that they would take a short break at this point (i.e., after the second SAM rating). They were given the three magazines and told that they could look through them during the break. The researcher started a stopwatch and stopped the participants after they had been browsing for 7 minutes. Both the participants assigned to the exercise intervening task and the magazine intervening task completed task ratings and another SAM at this point. Participants who were assigned to the "no intervening" condition moved directly from the SAM rating following the initial writing task to the dependent measure, the puzzle-solving task.

At this point, all participants completed the puzzle-solving task. The experimenter read these instructions to the participant:

> "The next thing I would like you to do is work on this spatial abilities task. What you need to do is to attempt to trace the puzzles in front of you. You need to trace each puzzle without picking up your pencil and without tracing over the same line segment twice. Make sure that you include all parts of the figure and that you do NOT pick up your pencil as you trace. It is not a matter of how many tries it takes or how long it takes you to solve the puzzle."

The experimenter then showed the participant Puzzle 1 and demonstrated by first tracing it incorrectly and then correctly. Next, the experimenter gave the participant a stack of Puzzle 2s (the other solvable puzzle) and asked the participant to complete it, using as many copies as they needed to. Participants were covertly timed as they completed this puzzle. Finally, participants were given stacks of Puzzles 3 and 4 (the unsolvable puzzles) and were told:

> "These are a little more difficult. You can pick either of the puzzles to work on or go back and forth; it's up to you. When you finish both puzzles just ring the bell [*they were given the desk bell at this point*] to let me know. Sometimes people get really frustrated with these puzzles and if that happens and you want to stop trying then that's fine too – just ring the bell. Again, feel free to use as many papers as you need to start over and take as much time as you need; there is still lots of time left in the study. Try your best to solve them both and let me know when you are finished by ringing the bell. Do you have any questions? You can start whenever you wish."

At this point, the experimenter left the testing room, taking the timer with them. Out of earshot, they began a stopwatch and waited to hear the participant ring the bell. When the participant rang the bell, or when 20 minutes had elapsed, whichever came first, the experimenter returned to the room. Finally, participants were asked to complete the post-experimental questionnaire. Afterward, they were debriefed with an explanation of the true nature of the study, with special emphasis on explaining the deception involved in the unsolvable puzzles task. They were then thanked, compensated, and debriefed.

Results

Depletion. Participants' depletion levels were scored based on their performance on the two unsolvable puzzles, both in terms of how long they spent persisting in trying to solve the puzzles, and the number of attempts they made at solving them (as measured by the number of different sheets used). Means and effect sizes can be found in Table 11. These were first analyzed via one-way ANOVAs using condition (Non-Depleted w/Magazine Intervening, Depleted w/Magazine Intervening, Depleted w/Exercise Intervening, and Depleted w/No Intervening) as the between-subjects factor to explore overall differences. More focused analyses were then run via independent samples *t*-tests comparing pairs of conditions. First, in order to test for depletion, we compared the scores of those in the Non-Depleted w/Magazine Intervening condition to those in the Depleted w/Magazine Intervening condition. As these people completed the same, purportedly "self-regulation neutral", magazine intervening task, any differences in puzzle score should be due to completing the different versions of the initial writing tasks. Then, we compared the two depleted groups who completed intervening tasks (Depleted w/Magazine Intervening and Depleted w/Exercise Intervening) to the one depleted no-intervening task group (Depleted w/No Intervening) to explore the effects of the intervening break on depletion.

Table 11. Depletion means and effect sizes for Study 2.1

					Effect Sizes	
Depletion Variable	Condition	*N*	*M*	*SD*	Depleting vs. Non-depleting w/magazines intervening	Intervening vs. None for Depleteds
Puzzles: min. persisting	Depleted: No Intervening	17	17.71	3.82	.01	-.70
	Depleted: Magazine Intervening	20	15.17	4.90		
	Depleted: Exercise Intervening	20	13.82	5.18		
	Non-Depleted: Magazines Intervening	20	15.22	5.25		
Puzzles: # attempts made	Depleted: No Intervening	17	17.53	6.61	-.16	.08
	Depleted: Magazine Intervening	20	19.40	7.59		
	Depleted: Exercise Intervening	19	16.74	8.63		
	Non-Depleted: Magazines Intervening	20	18.15	8.22		

Minutes persisting. The overall ANOVA did not indicate significant differences across conditions, $F(3,73) = 2.01$, $p = .12$. The pattern of the means, shown in Table 11, was unexpected, with those in the Depleted w/Exercise Intervening condition persisting the least time in attempting to solve the puzzles and those in the Depleted w/No Intervening condition persisting the longest time. The focused comparison of the depleted and non-depleted magazine interveners also provided no evidence of depletion, $t(38) = .03$, $p = .49$[11]. The focused comparison of those depleted participants who had, versus had not, completed an intervening task demonstrated a lack of replenishing effects of such a break, $t(55) = -2.36$, $p = .02$[12]. Indeed, there was a significant effect in the opposite direction, with those who had not engaged in any intervening task actually persisting longer.

Number of attempts made. On this measure as well, no global differences across the four conditions were found, $F(3,72) = .40$, $p = .75$. The pattern of means was similar to that found for the persistence time measure (see Table 11). The focused comparison of the depleted and non-depleted magazine interveners again indicated a lack of depletion, $t(38) = -.50$, $p = .62$. The focused comparison of those depleted participants who had completed an intervening task with those who had not again indicated a lack of replenishment, $t(54) = .26$, $p = .80$.

Mood. As in Study 1, we analyzed participants' SAM valence and arousal ratings to assess the impact of the initial tasks on mood. First, the post-task measurements were

[11] *t*-statistic *p*-values here are reported for one-tailed tests, given the hypothesized directional effect of non-depleted participants persisting longer and making more attempts.
[12] *t*-statistic *p*-values here are reported for two-tailed tests, given the uncertainty as to whether the intervening tasks would increase or decrease persistence and attempts.

evaluated via one-tailed independent samples *t*-tests using depletion condition (depleting vs. non-depleting) as the between-subjects factor. We also added time of measurement (before vs. after the depleting/non-depleting task) as a within-subjects factor, in addition to the between-subjects depletion condition factor, and ran two-way ANOVAs. The means and effect sizes are provided in Table 12.

Table 12. Mood means and effect sizes for Study 2.1

					Effect Sizes	
Mood Variable	Condition	*N*	*M*	*SD*	Depletion condition	Depletion condition x Time of measurement
SAM: Valence	Depleted: before	59	3.02	1.01	.22	
	Non-depleted: before	20	3.25	1.21		
	Depleted: after	59	3.58	1.39	.43	.58
	Non-depleted: after	20	3.00	1.30		
SAM: Arousal	Depleted: before	59	6.43	1.60	.19	
	Non-depleted: before	20	6.15	1.23		
	Depleted: after	59	5.54	1.71	.49	.70
	Non-depleted: after	20	6.35	1.50		

Note: Higher scores on valence actually indicate a more negative mood, and higher scores on arousal actually indicate a calmer mood.

SAM Valence. There was a marginally significant trend toward an effect of depletion condition (depleting vs. non-depleting) on post-task valence, $t(77) = 1.63$, $p = .054$, with those completing the non-depleting version giving higher emotional valence ratings. A significant interaction of depletion condition and time of measurement (before vs. after depleting/non-depleting task) was also found, $F(1,77) = 6.37$, $p = .014$, with the non-depleting version slightly increasing emotional valence and the depleting version negatively impacting valence (see Figure 11).

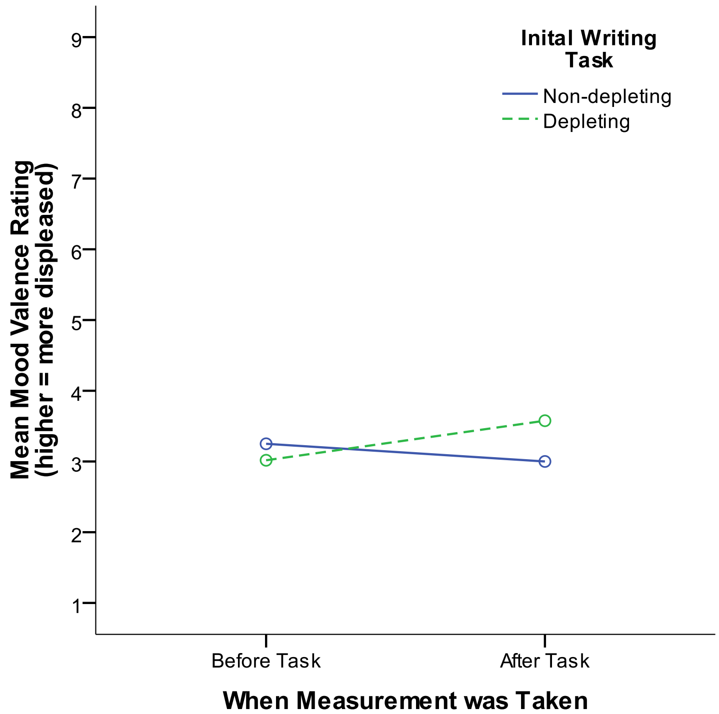

Figure 11. Mood valence by depletion condition and time of measurement in Study 2.1.

SAM Arousal. There was a significant effect of depletion condition on post-task emotional arousal, $t(77) = 1.89$, $p = .032$, with those completing the depleting versions reporting higher levels of arousal. There was also a significant interaction of depletion condition and time of measurement, $F(1,77) = 9.50$, $p = .003$, with the non-depleting version having a slight calming effect whereas the depleting version had an arousing effect (see Figure 12).

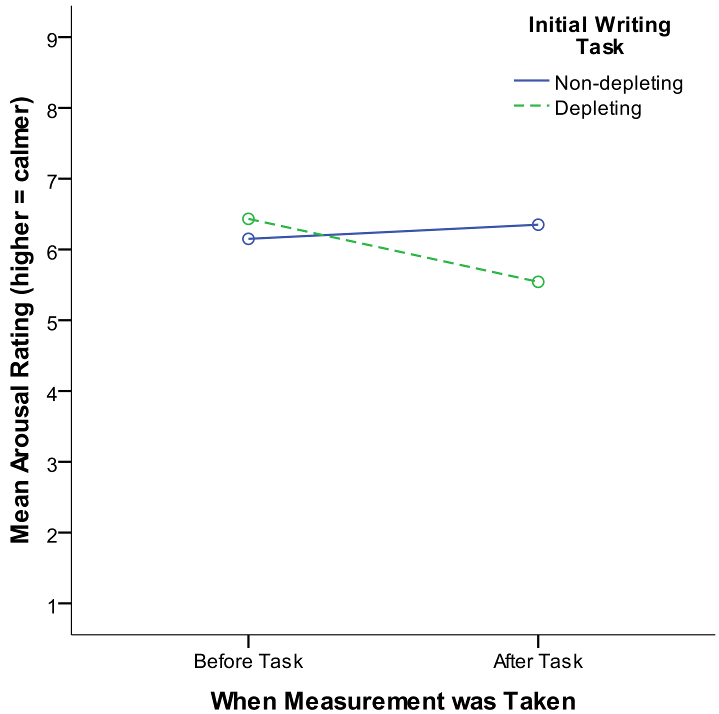

Figure 12. Mood arousal by depletion condition and time of measurement in Study 2.1.

Study 2.2

In Study 2.1, we were unable to detect any significant depletion effects on the newly adopted puzzle persistence task. We also did not see evidence of replenishment for those participants asked to exercise during the break. If anything, those in the exercise condition appeared to be more, not less, depleted, particularly as evidenced by their reduced level of persistence in attempting to complete the challenging puzzle solving task. When considering why this may have occurred, we realized that the exercise participants often appeared to be somewhat embarrassed about having to perform the designated exercises in the presence of an experimenter. This realization led us to hypothesize that one reason participants were not replenished by the exercise could be the additional monitoring elicited by feelings of embarrassment/social awkwardness, and/or by a feeling of a lack of autonomy – we hypothesized that participants did not feel they had much autonomous control over their own actions. One can imagine that while choosing to exercise could be replenishing, being asked or told to perform a number of specific movements in a lab setting with an experimenter looking on may not be. Indeed, as reviewed above, there is evidence indicating that a person's level of experienced autonomy can influence how depleting a task is (Muraven, 2008a; Muraven et al., 2008; Muraven, Rosman, & Gagné, 2007).

Therefore, we altered the procedures for our intervening tasks in Study 2.2. Some participants (as in Study 2.1) were assigned to an intervening task (e.g., exercise, magazine reading) to perform during their break. Others were given a choice to engage in any of several activities, including magazine reading, exercising, listening to music, socializing online, word finds, and drawing. Performance pressure was also reduced for all participants completing break tasks by asking them to take anything they produced during the break with them and informing them that we would not look at it. To further decrease the possibility of performance pressure during the break, the researcher left the testing room during all of the tasks in this study so that participants were not concerned about being observed. As autonomy was now being manipulated in this study, we also added two further questions to our task rating assessment (given after each task) to more fully assess perceived autonomy.

We also chose to use new dependent measures for Study 2.2. Although the figure-tracing persistence measure has been shown to be sensitive to the depletion effect in previous research (e.g., Baumeister et al., 2005; Moller et al., 2006; Muraven & Slessareva, 2003; Tice et al., 2007), when we ran participants using this task, we failed to obtain significant depletion effects. Therefore, we decided to instead use a task that collected continuous measurements in order to provide a more gradated and temporally-tied measure of depletion and replenishment – a flanker task[13] (Eriksen & Eriksen, 1974). In flanker tasks, participants are shown a row of items and asked to identify the center item. In each trial, the outer items, or "flankers," are either congruent or incongruent with the center item. The flanker task has often been viewed as an executive function measure, requiring participants to selectively attend to the central item and suppress what is often a prepotent response tendency to rely on the flankers as global indicators (e.g., Gratton, Coles, & Donchin, 1992; Purmann, Badde, Luna-Rodriguez, & Wendt, 2011; Wendt & Luna-Rodriguez, 2009). We were therefore interested in assessing the generalizablity of self-regulation depletion to this task.

We also included a fourth task – an operation span task – administered after the flanker task, in order to examine whether or not any positive effects of the intervening tasks, if present, only briefly counteracted depletion or had more lasting effects, potentially by actually replenishing self-regulatory resources. If so, it would be expected that performance for those in the replenished condition(s) on the fourth task would continue to surpass that of those in the non-replenished conditions. Like the flanker task, the operation span task provides continuous temporally-tied measurements of self-regulation across the course of the task. Additionally, and also like the flanker task, the

[13] The ability to assess performance continuously across the flanker task, as well as the operation span task, motivated our selections of these measures. However, analyses of the data by block produced nothing of consequence, and therefore will not be reported here.

operation span task is believed to tax central-executive control capacities such as working memory updating (e.g., Schmeichel, 2007; Turner & Engle, 1989).

Participants

A convenience sample of 83 participants (40 male, 43 female) from the University of Minnesota – Twin Cities area was recruited using the REP pool. Participants in this phase ranged in age from 18-30 years old (M = 20.48, SD = 2.44). Racially, they represented these categories: 80.70% Caucasian, 3.60% African or African American, 7.20% Asian or Asian American, and 7.20% who chose "Other" or chose not to respond.

Materials

The restricted and unrestricted story-writing task used in the previous studies again acted as our depletion manipulation. Additional materials retained from study 2.1 included the use of SAM measures to assess the mood (valence and arousal) of participants at various times as well as the Task Ratings forms. However, in this study, a new version of the SAM was used that provided pictograms for each of the nine levels (http://irtel.uni-mannheim.de/pxlab/demos/index_SAM.htm); this version of the SAM was adopted because it was thought that, despite instructions to utilize the space between the pictograms in the original measure, participants might have been responding in a manner more similar to a 5-point scale than the 9-point scale that was intended, reducing its sensitivity. This new version was also scored such that higher scores were actually equivalent to higher valence and more arousal. In this study, two additional questions were also added to the Task Ratings forms to further assess autonomy: (1) “Completing this task was important to me”, and (2) “I felt that I was able to master this task”.

Participants who were given a choice of intervening tasks were given an instructions page that informed them that, during the break, they could choose from among the options and could switch between activities whenever they wanted. The choices included reading a magazine, physical exercise, listening to music, socializing online, working on word finds, and drawing pictures. The magazines provided were again Sports Illustrated®, People®, and Self®. For exercising, suggestions such as stretching or walking in place were mentioned on the instructions page. For music listening, three Lifescapes® brand CDs were provided that were thought to be soothing in nature: a compilation of new age Spanish guitar music (Liebert, 2009), a compilation of modern lounging music on keyboard, guitar, and horns (Arthur, 2004), and a compilation of classical music on the piano, guitar, and lute (Freymuth, 1996). For socializing online, participants were allowed to use the lab room computer to check their email accounts, Facebook page, or the like. The word finds provided were obtained from http://www.puzzles.ca/wordsearch.html and were on the topics of photography, astronomy, and world currencies. For drawing, pieces of white 8.5” x 11” paper and a number of colored pencils in various shades were provided. Participants who were assigned to intervening tasks were given an instructions page that informed them of the task or tasks that had been assigned to complete and the materials listed above that were appropriate for the assigned task(s).

As stated previously, the new dependent measures used in this study were the flanker task and the operation span task. In the flanker task, in each trial, participants were shown a row of letters (e.g., SSSSS or SSHSS) and asked to identify the center item by pressing a button on the keyboard. The task included eight of these trials as practice items and 180 experimental trials. As is standard for this task, the outer, distracter, letters were either congruent with the center item (1st example listed above) or incongruent (2nd example). In this study, we used a smaller proportion (30%) of incongruent trials intermixed with a larger proportion (70%) of congruent trials to encourage establishing a prepotent response toward using the distracter items as a guide (Gratton et al., 1992; Purmann et al., 2011; Wendt & Luna-Rodriguez, 2009). Therefore, participants needed to exert self-regulation to overcome this prepotent response on incongruent trials. The task was presented using E-Prime software. White letters were used on a black background. Each trial consisted of a fixation cross shown for 500 ms, the stimulus shown for 200 ms, and an inter-stimulus interval of 1000 ms. Both participants' accuracy and reaction time were recorded.

The operation span task was an extended version of the task we had used to manipulate depletion in Studies 1.1–1.4. Recall that in this task, participants were required to encode and recall words while also evaluating the accuracy of interspersed mathematical equations. In the current study, the operation span task included 16 sets, ranging in size from 2 to 5 word-equation pairings per set, for a total of 56 pairings. Both participants' accuracy in word recall and accuracy in assessing the math equations, as well as their response times in assessing the math problems, were recorded.

Table 13: Timeline of Experimental Procedures for Study 2.2

Experimental Condition	Time →								
Depleted: No Choice Intervening	Mood Ratings	Depleting Version of Writing Task (6 min)	Mood Ratings and Task Ratings	Assigned Intervening Tasks (e.g., exercise, magazines, Facebook) (6 min)	Mood Ratings and Task Ratings	Flanker Task (6 min)	Mood Ratings and Task Ratings	Operation Span Task (15 min)	Mood Ratings and Task Ratings
Depleted: Choice Intervening	Mood Ratings	Depleting Version of Writing Task (6 min)	Mood Ratings and Task Ratings	Choice of Intervening Tasks (e.g., exercise, magazines, Facebook) (6 min)	Mood Ratings and Task Ratings	Flanker Task (6 min)	Mood Ratings and Task Ratings	Operation Span Task (15 min)	Mood Ratings and Task Ratings
Depleted: No Intervening	Mood Ratings	Depleting Version of Writing Task (6 min)	Mood Ratings and Task Ratings	Flanker Task (6 min)	Mood Ratings and Task Ratings	Operation Span Task (15 min)	Mood Ratings and Task Ratings		
Non-depleted: No Intervening	Mood Ratings	Non-depleting Version of Writing Task (6 min)	Mood Ratings and Task Ratings	Flanker Task (6 min)	Mood Ratings and Task Ratings	Operation Span Task (15 min)	Mood Ratings and Task Ratings		

Procedure

To aid in understanding the rather more complex experimental design of Study 2.2, a timeline of the procedures can be found in Table 13. As in the earlier studies, participants began by providing initial SAM ratings. They were then randomly assigned to one of the four conditions: (1) Non-Depleted with No Intervening task (NDNI), Depleted with No Intervening task (DNI), Depleted with a Choice of Intervening activities (DCI), and Depleted with Assigned Intervening activity(ies) (DAI). Participants then completed their assigned initial depleting or non-depleting writing task. In this study, we reduced the time participants spent on this task from seven to six minutes to more closely mirror time limits of previous research using this manipulation (Schmeichel, 2007; Schmeichel & Vohs, 2009), as there was some concern that even those in the "non-depleting" condition were having to self-regulate in order to continue writing during the seven-minute time limit we used in the earlier studies. As mentioned, the other change made in this study was that the researchers left the testing room as the participants were writing, rather than simply sitting at a nearby table. As in the earlier studies, when participants had completed their initial depleting or non-depleting writing task, they then rated that task and completed their second SAM ratings.

Participants in the NDNI and DNI conditions then proceeded directly to the flanker task (see below). Participants in the DCI and DAI conditions were told that they would take a short break at this point. Participants in the DCI condition were given the instructions page, informing them of the choices available, along with the various materials for the options. The researcher then left the room and returned 6 minutes later.

Participants in the DAI condition were given a typed page describing the task or tasks that they had been assigned to complete during the break and how long they were to engage in each of them. In order to ensure that the tasks completed by those in the DCI and DAI groups did not vary, we chose to yoke each DAI participant to a DCI participant. To accomplish this, all DCI and DAI participants were asked to report how long they engaged in which activities once the researcher had returned to the room. We used the responses of those in the DCI group to create the task assignment pages for those in the DAI group. We used the responses of those in the DAI group to ensure that they had completed the tasks as they had been assigned. Participants in both the DCI and DAI conditions also completed additional task ratings and another SAM at this point.

Participants then began the flanker task. They were given an instructions page explaining how to complete the task. It began by informing them that they should look at the fixation cross when it appeared. Then, a series of five letters would appear in a single row on the screen and consisting of either the letter H, S, or both. The participants' task was to determine whether the letter in the center was 'H' or 'S'. They were to press one button on the computer keyboard for 'H' and another for 'S,' and were to respond as quickly and as accurately they could. Before the experimental trials began, participants completed several practice trials and could ask any questions they had. As they had with the other tasks in this study, the researchers then left the room while the participants completed the flanker task. When the participant was done, the researcher returned and the participant completed another SAM rating and another Task Rating form.

Participants then completed the operation span task. Procedures were similar to when the task was used in Studies 1.1–1.4, with a few exceptions. First, as stated above, there were more trials in the current version. In addition, the researcher left the room while the participant completed the experimental trials. Therefore, rather than stating aloud their responses regarding the accuracy of each of math problems, participants were asked to press a button on the keyboard to indicate whether or not they thought the equation was correct. Participants also pressed a button themselves when they had completed a recall phase in order to advance to the next trial. Upon completion of the task, the researcher returned to the room and the participant completed their final SAM and task ratings. Finally, participants were debriefed, thanked, and compensated for their participation.

Results

Autonomy Manipulation Check: Manipulations of choice have often been shown to create differences in autonomy in past research (e.g., Moller et al., 2006; Muraven et al., 2008; Zuckerman, Porac, Lathin, Smith, & Deci, 1978, see also Patell, Cooper, & Robinson, 2008). However, to our knowledge, no one has used our exact manipulation wherein participants are either assigned to or allowed to choose between the particular task options we provided. Therefore, we first assessed participants' ratings of the break task determine if there was any evidence that our choice (DCI) participants felt more autonomy than our assigned (DAI) participants. Indeed, the DCI participants did agree with the statement "I felt I had to do this task" (M = 5.20, SD = 2.69) to a significantly greater degree than did the DAI participants (M = 3.62, SD = 2.20), $t(39)$ = 2.07, p = .046, d = .66. However, there were not differences between the groups in terms of their agreement to two of the other statements relating to autonomy: “I felt pressure during this task” (M_{DCI} = 1.95, SD_{DCI} = 1.65, M_{DAI} = 1.95, SD_{DAI} = 1.27, $t(36) = 0.00$, p = 1.00, d = 0.00) and “I felt mastery during this task” (M_{DCI} = 8.30, SD_{DCI} = 1.69, M_{DAI} = 7.60, SD_{DAI} = 2.54, $t(38) = 1.03$, p = .31, d = .33). The DCI participants (M = 3.76, SD = 1.67) also actually agreed with the statement “Completing this task was important to me” to a significantly *lesser* degree than did the DAI participants (M = 5.70, SD = 2.66), $t(39)$ = 2.81, p = .008, d = .90. Thus, while there was some evidence that participants in the choice condition felt more autonomy than did those in the assigned condition, it was by no means unequivocal.

Depletion: Flanker. Participants’ depletion levels were first assessed via their performance on the flanker task. Participants were scored based on the accuracy and the speed of their responses. For both, their overall performance as well as the difference between their performance on the congruent and incongruent trials were analyzed. For each of these measures, we first compared all four conditions (NDNI, DNI, DCI, and DAI) using one-way ANOVAs. We then ran a number of focused pair-wise comparisons. First, performance of the NDNI participants was contrasted with that of the DNI participants to assess the amount of depletion created. Then, performance of the DNI participants was contrasted with the average across the DCI and DAI participants to determine if taking a break in general counteracted depletion. Finally, performance of the DCI participants was contrasted with that of the DAI participants to determine if participants’ autonomy condition (i.e., self-chosen versus assigned intervening task) affected the replenishing effects of the break. Means and effect sizes can be found in Table 14.

Table 14. Depletion means and effect sizes for flanker task in Study 2.2

					Effect size		
Variable	Condition	*N*	*M*	*SD*	NDNI vs. DNI	DNI vs. DAI and DCI	DCI vs. DAI
Accuracy Overall	Non-Depleted: No Intervening	19	97.95	.78	1.28		
	Depleted: No Intervening	20	95.75	2.34		-.21	
	Depleted: Choice Intervening	21	94.90	3.77			-.10
	Depleted: Assigned Intervening	20	95.25	3.73			
Accuracy Congruent-Incongruent *effect size overall = 1.61*	Non-Depleted: No Intervening	20	3.10	3.99	.10		
	Depleted: No Intervening	21	3.59	6.05		-.24	
	Depleted: Choice Intervening	19	3.98	4.60			.36
	Depleted: Assigned Intervening	20	5.70	5.32			
RT	Non-Depleted: No Intervening	19	294.74	46.23	.24		
	Depleted: No Intervening	20	306.80	56.72		.54	
	Depleted: Choice Intervening	20	274.93	55.15			.12
	Depleted: Assigned Intervening	19	281.11	53.15			
RT Incongruent-Congruent *effect size overall = 4.71*	Non-Depleted: No Intervening	17	55.94	16.64	.36		
	Depleted: No Intervening	21	63.79	26.47		-.06	
	Depleted: Choice Intervening	21	65.71	24.40			-.04
	Depleted: Assigned Intervening	19	64.79	21.59			

Flanker overall accuracy. There were significant differences across the four conditions in overall accuracy on the flanker task, $F(3, 76) = 4.18$, $p = .009$. In particular, focused comparisons showed that participants in the DNI condition were significantly less accurate than participants in the NDNI condition, $t(37) = 3.90$, $p < .001$, illustrating self-regulation depletion. However, there was no evidence that those who completed an intervening break task were able to perform better than those who had no break, $t(59) = -.75$, $p = .46$. There was also no significant difference in accuracy between those given a choice of intervening task and those where were assigned to an intervening task, $t(39) = -.30$, $p = .77$.

Flanker effect for accuracy (accuracy on congruent trials – accuracy on incongruent trials). There were strong flanker effects overall, $F(1,74) = 54.55$, $p < .001$, Cohen's $d = 1.61$. However, the strength of these flanker effects was not influenced by the condition to which the participant was assigned, $F(3,76) = 1.00$, $p = .40$. Focused comparisons showed no differences between the DNI and NDNI groups, $t(39) = .30$, $p = .76$, nor between those depleted participants who completed an intervening task and those who did not, $t(58) = -.88$, $p = .39$, nor between those who had a choice of intervening task and those assigned to one, $t(37) = 1.08$, $p = .29$.

Flanker RTs (median RT for correct answers). There were also no significant differences across the four conditions in overall reaction times on the flanker task, $F(3,74) = 1.43$, $p = .24$. Focused comparisons showed that participants in the DNI condition and participants in the NDNI condition did not differ in speed, $t(37) = .73$, $p = .47$. There was, however, a marginally significant trend toward those who were asked to complete an intervening break task being able to respond faster than those who had no break, $t(57) = 1.92$, $p = .06$. Finally, there was no significant difference in speed between those given a choice of intervening task and those who were assigned to an intervening task, $t(37) = .36$, $p = .72$.

Flanker effect RT (median RT on incongruent trials – median RT on congruent trials). Again, there were strong flanker effects overall, $F(1,77) = 474.55$, $p < .001$, Cohen's $d = 4.71$. However, again, the strength of these flanker effects was not influenced by the condition to which the participant was assigned, $F(3,74) = .68$, $p = .57$. Focused comparisons showed no depletion, with those in the DNI condition demonstrating similar flanker effects in RT to those in the NDNI condition, $t(36) = 1.06$, $p = .30$. There were also no significant differences between those depleted participants who completed an intervening task and those who did not, $t(59) = -.23$, $p = .82$, nor between those who had a choice of intervening task and those assigned to one, $t(38) = -.13$, $p = .90$.

Depletion: Operation Span. Participants' depletion levels were also scored based on their performance on the Operation Span task that was administered after the flanker task. Participants were scored based on their recall performance, in terms of the number of words they recalled correctly, the number of sets in which they recalled all of the words correctly, the size of the biggest set they were able to recall completely, and the number of intrusions they made – that is, words written down at recall that were not shown in that set. Participants' accuracy and the speed of their responses to the math questions were scored as well. Means and effect sizes can be found in Tables 15 and 16.

As with the flanker task, we first compared all four conditions (NDNI, DNI, DCI, and DAI), and then ran focused pair-wise comparisons to test for depletion, overall counteracting of the depletion, and differences in replenishment due to autonomy condition (self-chosen vs. assigned intervening task).

Number of Words Recalled (out of 56 possible). There were no significant differences in the number of words recalled across the four conditions, $F(3,76) = 1.31$, $p = .28$. Focused comparisons showed no difference between DNI and NDNI participants, $t(40) = .86$, $p = .40$. There were also no significant differences between those depleted participants who completed an intervening task and those who did not, $t(57) = .41$, $p = .68$. There was a trend toward a difference between those who had a choice of intervening task and those assigned to one, $t(36) = -1.73$, $p = .09$; however, this reflected a tendency for those assigned to an intervening task to recall more words.

Number of Full Sets of Words Recalled (out of 16 possible). The conditions did not different significantly overall on this measure either, $F(3,78) = .41$, $p = .75$, nor were DNI and NDNI participants significantly different, $t(40) = .70$, $p = .49$. There were no significant differences between those depleted participants who completed an intervening task and those who did not, $t(59) = -.19$, $p = .85$, nor between those with a choice of intervening task and those assigned to one, $t(38) = -.47$, $p = .64$.

Size of Biggest Word Set Fully Recalled (out of 5 total). Again, there were no differences overall between conditions, $F(3,77) = .26$, $p = .85$. DNI participants did not differ from NDNI participants, $t(40) = .00$, $p = 1.00$. There were also no significant differences between those depleted participants who completed an intervening task and those who did not, $t(58) = -.60$, $p = .55$, nor between those with a choice of intervening task and those assigned to one, $t(37) = -.53$, $p = .60$.

Table 15. Depletion means and effect sizes for OpSpan word scores in Study 2.2

					Effect size		
Variable	Condition	*N*	*M*	*SD*	NDNI vs. DNI	DNI vs. DAI and DCI	DCI vs. DAI
Number of Words Recalled	Non-Depleted: No Intervening	21	47.60	5.43	.27		
	Depleted: No Intervening	21	46.12	5.74		.11	
	Depleted: Choice Intervening	21	45.26	7.62			-.58
	Depleted: Assigned Intervening	17	48.71	3.38			
Number of Full Word Sets Recalled	Non-Depleted: No Intervening	21	10.48	3.06	.22		
	Depleted: No Intervening	21	9.81	3.08		-.05	
	Depleted: Choice Intervening	21	9.43	3.68			-.15
	Depleted: Assigned Intervening	19	9.89	2.36			
Biggest Set of Words Fully Recalled	Non-Depleted: No Intervening	21	4.48	.75	.00		
	Depleted: No Intervening	21	4.48	.75		-.17	
	Depleted: Choice Intervening	20	4.30	.80			-.17
	Depleted: Assigned Intervening	19	4.42	.61			
Number of Word Intrusions	Non-Depleted: No Intervening	20	1.83	1.09	.35		
	Depleted: No Intervening	21	2.31	1.69		.42	
	Depleted: Choice Intervening	21	2.24	1.79			-.97
	Depleted: Assigned Intervening	17	.92	.67			

Number of Intrusions. Here, there was a significant difference between the conditions overall, $F(3,75) = 3.73$, $p = .015$. DNI and NDNI participants did not differ significantly, $t(39) = 1.08$, $p = .29$. The differences between those depleted participants who completed an intervening task and those who did not also did not reach significance, $t(57) = 1.53$, $p = .13$. However, those with a choice of intervening task and those assigned to one did differ significantly, $t(36) = -2.89$, $p = .007$, with those assigned to an intervening task making fewer intrusions.

Table 16. Depletion means and effect sizes for OpSpan math scores in Study 2.2

					Effect size		
Variable	Condition	*N*	*M*	*SD*	NDNI vs. DNI	DNI vs. DAI and DCI	DCI vs. DAI
Math Accuracy	Non-Depleted: No Intervening	20	94.11	2.32	.85		
	Depleted: No Intervening	20	91.91	2.91		.66	
	Depleted: Choice Intervening	16	93.75	1.30			.11
	Depleted: Assigned Intervening	20	93.48	3.08			
Math RTs	Non-Depleted: No Intervening	21	5084.83	1386.05	.10		
	Depleted: No Intervening	20	5204.63	1206.27		-.03	
	Depleted: Choice Intervening	21	5033.64	1011.71			.40
	Depleted: Assigned Intervening	20	5461.85	1202.90			

Math Accuracy Percentage. The comparison of conditions overall did show significance, $F(3,72) = 2.83$, $p = .044$. Furthermore, there was evidence of significant depletion, with DNI participants scoring significantly lower than NDNI participants, $t(38) = 2.63$, $p = .012$. There were also significant differences between those depleted participants who completed an intervening task and those who did not, $t(54) = 2.32$, $p = .024$, with higher performance among those who completed an intervening task supporting the restorative effects of taking a break. There was little evidence that mood mediated this effect. There was no significant correlation between participants' mood valence ($r = .14$, $p = .314$) or arousal level scores ($r = .03$, $p = .836$) before the operation span task and their accuracy. With participants' moods before the operation span task added as a covariate, restorative effects were only slightly reduced and remained significant: mood valence added: $p = .030$, $d = .63$; mood arousal added: $p = .026$, $d = .65$. Finally, there were no significant differences between those with a choice of intervening task and those assigned to one, $t(34) = .32$, $p = .75$.

Math RTs (medians). The four conditions did not differ overall with regards to their speed on the math problems, $F(3,78) = .51$, $p = .68$. DNI and NDNI participants did not differ significantly, $t(39) = .30$, $p = .77$. There were also no significant differences between those depleted participants who completed an intervening task and those who did not, $t(59) = -.12$, $p = .90$. Those with a choice of intervening task and those assigned to one also did not perform significantly differently, $t(39) = 1.24$, $p = .22$.

Mood. As in the previous studies, participants' SAM valence and arousal ratings were analyzed to assess the impact of the initial tasks on mood. First, the pre-task ratings of participants were compared via one-tailed independent samples *t*-tests to assess whether (as hoped) the groups were equivalent on mood at the beginning of the experiment. Then, the post-task measurements were compared. As before, we also added time of measurement (before vs. after the depleting/non-depleting task) as a within-subjects factor and ran two-way ANOVAs. The means and effect sizes are given in Table 17.

Table 17. Mood means and effect sizes for Study 2.2

					Effect Sizes	
Mood Variable	Condition	*N*	*M*	*SD*	Depletion condition	Depletion condition x Time of measurement
SAM: Valence	Depleted: before	60	6.57	1.17	.04	
	Non-depleted: before	17	6.59	0.62		
	Depleted: after	60	6.30	1.39	.29	.35
	Non-depleted: after	17	6.65	.70		
SAM: Arousal	Depleted: before	62	4.15	1.56	.18	
	Non-depleted: before	20	3.65	1.39		
	Depleted: after	62	4.68	1.52	.98	.74
	Non-depleted: after	20	3.15	1.76		

Note: Using the new 9-point SAM, higher scores on valence do indicate a more positive mood, and higher scores on arousal indicate a more aroused mood.

SAM Valence. Participants who were assigned to complete the depleting and non-depleting versions of the initial task did not differ significantly in mood valence at the beginning of the study, $t(76) = 0.15$, $p = .88$. They did, however, show a trend toward differing significantly on their post-task valence ratings when using a one-tailed t-test, $t(76) = 1.05$, $p = .15$, as well as a trend toward an interaction of depletion condition and time of measurement (before vs. after depleting/non-depleting task) akin to those we saw in earlier studies, $F(1,75) = 2.26$, $p = .137$, with the non-depleting version of the initial task having little impact on valence and the depleting version somewhat negatively impacting valence (see Figure 13 – note the new DV direction).

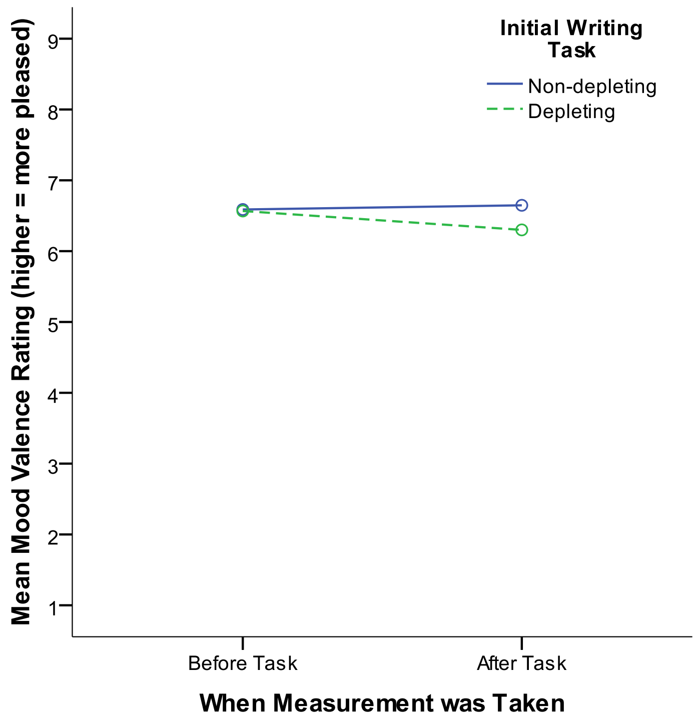

Figure 13. Mood valence by depletion condition and time of measurement in Study 2.2.

SAM Arousal. Participants who were assigned to complete the depleting and non-depleting versions of the initial task did not differ significantly in arousal at the beginning of the study, $t(81) = 0.72$, $p = .47$. However, there was a significant effect of depletion condition on their post-task arousal ratings, $t(80) = 3.76$, $p < .001$, with those completing the depleting version of the initial task reporting higher arousal. As done in earlier studies, we then conducted analyses to determine if these mood effects mediated the depletion effects found, in this case for overall accuracy on the flanker task and accuracy on the math component of the operation span task. Post-task arousal was not strongly or significantly correlated with accuracy on the flanker task, *r* (36) = -.11, *p* = .50, nor with accuracy in assessing the math problems on the operation span task, *r* (37) = -.10, *p* = .55. Not surprisingly, therefore, when post-task mood valence was added as a covariate, depletion effects were not reduced and remained significant for both flanker task accuracy and operation span math assessment accuracy (*p* < .001, *d* = 1.32, and *p* = .005, *d* = .98, respectively).

There was also a significant interaction of depletion condition and time of measurement, *F*(1,80) = 11.07, *p* = .001, with the non-depleting version of the initial task having a calming effect whereas the depleting version had an arousing effect (see Figure 14 – note the new DV direction). As before, we assessed the mediating potential of these mood effects. Change in arousal was not significantly correlated with flanker accuracy, *r* (36) = -.11, *p* = .527, nor with operation span math accuracy, *r* (37) = -.19, *p* = .257. When change in arousal was added as a covariate, depletion effects again remained strong and significant for both flanker task accuracy and operation span math assessment accuracy (*p* = .001, *d* = 1.28 and *p* = .011, *d* = .89, respectively).

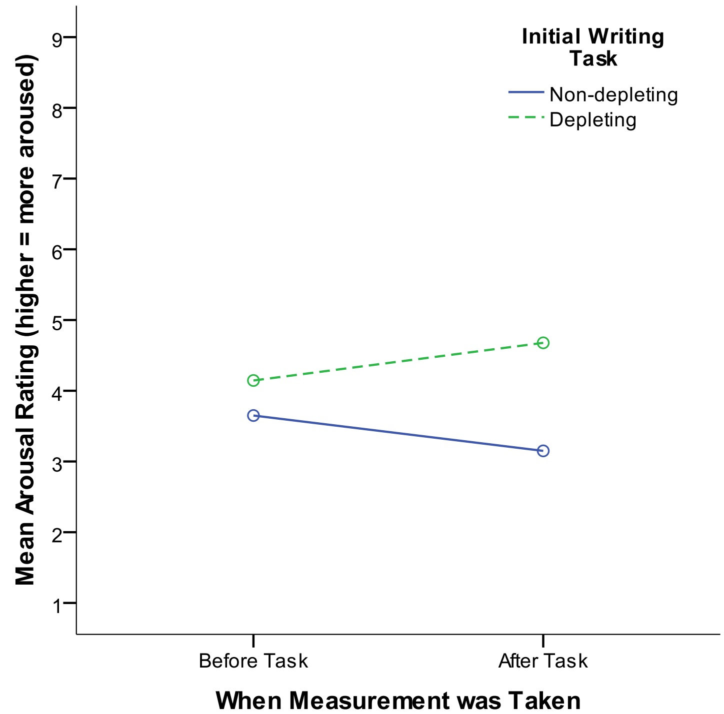

Figure 14. Mood arousal by depletion condition and time of measurement in Study 2.2.

Study 2 Discussion

As in Study 1, our most consistent finding in Study 2 was the impact of the initial self-regulation manipulations on mood. Differences in post-task arousal levels as well as differential changes in arousal from pre- to post-task were the most consistently seen. Thus, Study 2 provides further evidence that self-regulation does affect mood, and additionally suggests that arousal levels may be the more likely of the two aspects of mood to be influenced. However, Study 2 provided no evidence that either post-task mood or changes in mood mediated the depletion effects found. In Study 2.1, there were no significant depletion effects, negating the possibility of exploring mediation effects. In Study 2.2, however, there were significant depletion effects and significant arousal effects, but no evidence of the arousal mediating the depletion. One can only speculate as to why this was the case. One possible explanation may lie in the fact that in Study 2.2 we only found increases in arousal, not mood valence. In line with Carver's (2004) theory of mood and self-regulation, a small degree of failure in self-regulation on the initial task could heighten arousal, encouraging participants to reengage and to try harder. This would likely counteract depletion. Only when greater degrees of failure on the initial task are experienced is sadness experienced and disengagement (often a measurement of depletion) triggered.

Depletion effects on the primary self-regulation measures also remained inconsistent, as was previously also observed in Study 1. No significant depletion was detected in Study 2.1 and only two of the ten dependent measures in Study 2.2 demonstrated depletion. This may perhaps be explained in Study 2.2 by our use of the new dependent measures. However, both the flanker and operation span tasks have been used in a number of studies testing executive function and response-inhibition (flanker – e.g., Martin & Kerns, 2010; Rowe, Hirsh, & Anderson, 2006; operation span – e.g., Aslan & Karl-Heinz, 2011; Schmeichel, 2007; Turner & Engle, 1989), and the operation span task has been used specifically to measure self-regulation depletion previously (Schmeichel, 2007). Furthermore, Study 2.1 involved materials and procedures for both the initial self-regulation manipulation phase and the subsequent self-regulation measurement phase that had been used successfully in the past (initial writing task – Schmeichel, 2007; Schmeichel & Vohs, 2009; subsequent puzzles task – Baumeister et al., 2005; Moller et al., 2006; Muraven & Slessareva, 2003; Tice et al., 2007). This gives further support to the argument that self-regulation depletion is not as ubiquitous a phenomenon as it perhaps appears to be from the published literature.

Finally, with Study 2, we attempted to examine various ways to restore self-regulation abilities. We found only one potential restoration effect – the effect of taking a break on participants' math accuracy on the operation span task in Study 2.2. This result could suggest that, in addition to the increases in persistence seen by Tyler and Burns (2008) following a break, periods of rest may heighten one's mental acuity as well. However, this result should be interpreted cautiously. Given the number of analyses required in Study 2.2, as well as the lack of similar restorative effects on the other measures, the potential for this effect being due to Type I error cannot be ruled out.

Study 3

As we saw earlier in Study 1, evidence of depletion was not consistently found in Study 2. Furthermore, we found little support for the possibility that the intervening tasks used in Study 2 could counteract depletion. Effects of the initial tasks on participants' mood, however, continued to be quite consistent. Therefore, we wished focus mainly on exploring these mood effects in greater detail in Study 3. In particular, we wished to further investigate the impact of the initial self-regulation tasks on participants' moods using new biological and implicit measurements of mood. One motivation for using these more indirect measures of mood was the possibility that direct or explicit measures of mood themselves may encourage attempts at the self-regulation and/or monitoring of mood (e.g., Caruso & Shafir, 2006; Tice et al., 2001; cf. Clore & Hutsinger, 2007; Schwartz, 2001, for arguments that moods have less impact on subsequent tasks when they can be attributed to a salient cause). Therefore, it is important to determine if there are mood effects of self-regulatory tasks using less direct measures.

We also wanted to continue exploring other ways to manipulate demands on self-regulation and to measure subsequent self-regulation abilities. In particular, we were curious in this final study to investigate whether depletion would be evident if the initial task and the subsequent task were similar to each other. Most previous research on the self-regulatory depletion theory has used tasks in quite different domains to cause and measure the depletion (e.g., *inducing* depletion via requiring mood regulation but *assessing* depletion via the ability to resist tempting foods) in order to demonstrate that depletion effects were not task-specific but a generalized effect on self-regulation. However, we were interested in investigating the phenomenon of self-regulation depletion using measurements that were quite similar to the original depleting task. According to the limited resources model of self-regulation, having to complete any self-regulation task should tax this resource, leaving less remaining for any subsequent self-regulation task. However, one would expect that repeated completions of similar self-regulation tasks might lead to practice effects rather than depletion effects. Indeed, training in any domain requiring self-regulation or executive functions (e.g., potty training children, or learning to play an instrument or complex game) relies on an assumption that practice should produce improvement in performance. There is also empirical evidence that executive functions such as working memory updating (Dahlin, Bächman, Stigsdotter Neely, & Nyberg, 2009) and executive attention (Shalev, Tsal, & Mevorach, 2007) can be improved through training. Some researchers' conceptualizations of self-regulation do acknowledge the potential for such training effects (e.g., Baumeister et al., 2007; Muraven, Baumeister, & Tice, 1999); however, they are usually confined to long-term practice effects. For example, Baumeister and colleagues (2007; Muraven & Baumeister, 2000; Schmeichel & Baumeister, 2004) have argued that self-regulation is akin to a muscle that tires when exercised (i.e., becomes depleted), but can become stronger over time with repeated use (i.e., can be strengthened through training). Evidence indicating that more immediate practice effects can surpass depletion effects would, therefore, raise questions about this analogy.

Finally, Study 3 involved two phases with slightly different versions of the initial high and low self-regulation tasks. Combining the data across these phases allowed us to explore depletion effects beyond the simple pair-wise comparison used in our earlier studies and the bulk of previous research. That is, whereas earlier studies have almost always manipulated initial self-regulation as a dichotomous variable with only two levels of initial self-regulation requirements (depleted vs. non-depleted; cf. Vohs et al., 2008), in this final study, we compared four levels of initial self-regulation.

Participants

Phase 1. A convenience sample of 83 undergraduate students (27 male, 54 female, 2 opted not to report gender) were recruited from the University of Minnesota – Twin Cities psychology REP pool. These participants ranged in age from 18-30 years old (M = 20.19, SD = 2.70) and the sample was racially composed of 68.70% Caucasians, 9.60% African or African Americans, 10.80% Asian or Asian Americans, 4.8%

Hispanic/Latino, and 6.00% who reported their race as mixed, “other” or chose not to report the information.

Phase 2. Another convenience sample of 80 undergraduate students (36 male, 42 female, 2 opted not to report gender) were recruited from the University of Minnesota – Twin Cities psychology REP pool for this phase. Participants’ ages ranged from 18-30 years old (*M* = 19.82, *SD* = 1.84). Racially, 86.30% identified as Caucasians, 2.50% as African or African Americans, 3.80% as Asian or Asian Americans, and 7.60% who reported their race as mixed, “other” or chose not to report the information.

Materials

Phase 1. For the biological indicators of mood, we collected pulse (heart rate, or HR) and diastolic and systolic blood pressure (BP) data from participants using a simple cuff measure that is minimally invasive and provides rapid readings (the Omron Ultra Premium Wrist Blood Pressure Monitor with advanced positioning sensor, model HEM-670IT). Similar wrist monitors have recently been used in empirical research assessing the relationship between cardiovascular responses and mood (Ilies, Dimotakis, & Watson, 2010) as well as assessing mental fatigue using physiological responses (Stewart et al., 2009).

For the implicit behavioral measurement of mood, we designed brief word completion tasks based on some previous implicit mood measures (e.g., deWall & Baumeister, 2007; cf. Koole, Smeets, van Knippenberg, & Dijksterhuis, 2009; Quirin, Kazén, & Kuhl, 2009). Participants were shown fragments of words and asked to fill in one letter in order to complete each of words. Many of the fragments could be completed with words that had positive or negative mood connotations (e.g., _OWN could be the negative “frown” or the neutral “flown”; P _ ACE the positive “peace”, or the neutral “place”). The logic behind this task was that participants would be more likely to respond with words that fit their current mood state. Norms from previous studies wherein participants rated the valence (i.e., the associated positive or negative emotions) of words (Bradley & Lang, 1999; Hoversten & Koutstaal, 2011) were used to guide the selection of word fragments to include on the tasks. Two of these word completion tasks were created – one for participants to be complete before the initial self-regulation task and one for them to complete afterward. Each task contained 24 fragmented words, including 17 fragments that could be completed as positive, negative, and/or neutral words and 7 filler fragments that could only be completed with mood-neutral words (e.g., KN_T).

As mentioned previously, we also utilized a new initial task in this study in order to examine the generalizablity of our previous mood findings. We particularly wished to utilize a task that would give us more temporally continuous measurements of performance in order to provide the potential to assess changes in the performance of participants over the course of the task for the depleting and non-depleting versions. The task chosen was the Sustained Attention to Response Task (SART), modeled after that used by Robertson, Manly, Andrade, Baddeley, and Yiend (1997; see also Manly, Robertson, Galloway, & Hawkins, 1999). In this Number SART, participants saw single-digit numbers flash on the screen. Their task was to press a button each time they saw a number with one exception: they were told not to press anything if the number was a 3.

Participants needed to exert self-regulation to sustain attention to the task and to overcome the prepotent response of responding on the exception trials. For the task stimuli, the digits 1-9 were created in white text on black backgrounds in five font sizes: Times New Roman 48, 72, 94, 100, and 120 points, which corresponded to displays between approximately 17-44 mm when presented on the computer screen using E-Prime software. These stimuli were presented one at a time in the center of the screen, and in a semi-randomized order, for 250 ms, followed by a 900 ms mask between trials. The task consisted of 225 trials total, 25 of which were exception trials (i.e., the number 3). A less depleting version of this task was also created wherein participants were asked to press the button for all of the numbers shown with no restrictions.

As we wished to explore the effects of an initial self-regulation task on a similar subsequent self-regulation task, we also used a version of the SART for our dependent measurement phase. This new dependent measurement was a letter version of the SART task, wherein participants saw letters of the alphabet flash on the screen. Their task was to press a button for each letter with one exception: they were told not to press anything if the letter was a ‘D’. As with the Number SART, participants needed to exert self-regulation to sustain attention to the task and to overcome a prepotent response of responding on the exception trials. For the task stimuli, the consonant letters between and including the letters ‘B’ through ‘L’ were created in white text on black backgrounds in the same five font sizes as for the Number SART stimuli; all the letters were presented in uppercase. The number of response items and the number of exception items, as well as the speed of presentation, were also exactly the same as in the Number SART task described above. Participants' accuracy of responding/not responding and their reaction times to response item trials acted as our new dependent measures.

Phase 2. We utilized the same explicit, implicit, and biological indicators of mood in this phase as we did in Phase 1. We also again used versions of the Number SART task to manipulate depletion, but with slight alterations. The new depleting version contained 50 exception trials (rather than the 25 in the earlier phase), in order to increase the instances of self-regulation required of the participants. A further change was made to the non-depleting task. The new non-depleting version asked participants to simply watch the numbers on the screen, with no button pressing required. However, to allow across-phase comparisons, the final dependent measure was held constant; that is, the same 25-exception trials Letter SART task that was used in Phase 1 was retained as our measure of subsequent self-regulation.

Procedure

Phase 1. Participants were randomly assigned to one of the four conditions in a 2 x 2 between-subjects factorial design: (1) Non-depleted with explicit SAM mood measurements, (2) Depleted with explicit SAM mood measurements, (3) Non-depleted with implicit word completion mood measurements, and (4) Depleted with implicit word completion mood measurements. All participants began by putting on the wrist cuff and providing HR and BP measurements. Participants then completed their initial SAM ratings or their first word completion task, dependent upon the condition to which they were assigned. HR and BP measurements were then taken again for all participants.

Participants then completed their assigned initial depleting or non-depleting Number SART task. Participants assigned to complete the depleting version were told:

> In this task, you will see single digits from 1 through 9 flash in the center of the computer screen, one at a time. Your job is to press the "B" key each time a digit appears EXCEPT for the digit 3. That is, when you see a 3 on the screen, DO NOT press any key. Please use the index finger of your dominant hand and respond as quickly and as accurately as possible. Once the experimenter starts the computer program, you will complete a small set of practice trials, after which you can ask any additional questions you might have. You will then complete the experimental trials, which will take about 5 minutes.

Participants assigned to the non-depletion were given instructions that said the same thing, with the exception that the information about not pressing for a 3 was removed. After completion of this task, HR and BP were taken again. Participants then completed their second SAM or word completion task, and HR and BP were taken a fourth time.

All participants then began the Letter SART task. They were given an instructions page exactly like the depleting version of the Number SART task, except that all references to digits were replaced with references to letters, and references to the digit 3 were replaced with the letter D. After participants had completed the Letter SART, HR and BP were measured a final time. Finally, participants were asked to complete Task Ratings for both the Number SART and the Letter SART. Upon completion, they were debriefed and compensated.

Phase 2. Procedures in this phase were all identical to those in Phase 1, with the exception of new instructions for those completing the non-depleting Number SART stating that the participant’s job was simply to watch the numbers on the screen.

Results

Manipulation check. In our previous studies, we used initial self-regulation tasks that had been shown to produce depletion in previously published research. However, here in Study 3, we utilized versions of the SART task, which to our knowledge have not been used in self-regulation depletion research to date. Therefore, we first analyzed performance on the initial Number SART (see Table 18) and participants' perceptions of it (see Table 19) to determine if the "depleting" versions were indeed more challenging. There is some evidence to suggest that they were. Depletion condition had a significant effect on reaction times, $F(2, 112) = 68.55$, $p < .001$, with Tukey HSD post-hoc tests indicating that participants needed to take longer on average per trial on the depleting versions than on the non-depleting version. However, the participants completing the non-depleting version in which they pressed for all items actually missed a higher percentage of trials on average than those completing either of the depleting versions, $F(2, 110) = 21.54$, $p < .001$, suggesting they may have had trouble keeping their attention on the task. Another interesting difference to note was that participants who completed the depleting version with 25 exception trials actually had a higher percentage of false positives than those with 50 exception trials, $F(1, 80) = 8.31$, $p = .005$, $d = .65$, suggesting that the version with fewer required instances of self-regulation was actually more difficult in some respects. Even when comparing the number of false positives participants made on their first 25 exception trials – i.e., all trials for the SART 25 participants and the first 25 for the SART 50 participants – this significant difference remained, $F(1, 78) = 9.66$, $p = .003$, $d = .70$. Number SART 50 participants (M = 7.85, SD = 4.92) made significantly fewer false positives on their first 25 exception trials than the Number SART 25 participants (M = 11.24, SD = 4.86) did in their entire task. This task version difference, found even after equating the number of exception trials that are considered, is likely due to differences in the task-induced prepotency of responding. Suppression responses were required, on average, more frequently for the SART 50 (overall proportion of required suppression responses of .20, or 50/250 trials) than for the SART 25 (overall proportion of required suppression responses of .11, or 25/225 trials).

Table 18. Performance on the initial Number SART task in Study 3

	Median RT		Percent Presses Missed		Percent False Alarms	
Condition	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*
Non-depleting version - pressing for all	208.45	39.07	3.9	3.79	na	na
Depleting version with 50 exceptions	311.76	46.53	0.78	1.17	32.95	18.21
Depleting version with 25 exceptions	298.00	39.47	0.79	1.02	44.95	19.43

Participants' task ratings also illustrated results indicative of the depleting versions being more likely to tax self-regulation resources. Participants' ratings of the difficulty of the Number SART did depend on the version they completed, $F(3, 152)$ = 85.07, $p < .001$. As expected, Tukey HSD post-hoc comparisons indicated that participants found the non-depleting versions easier than the depleting versions, with the no pressing version thought to be the easier of the two non-depleting versions. Interestingly, and in line with their performance, a focused comparison indicated that participants found the depleting version with 25 exception trials to be more difficult than the depleting version with 50 exception trials, *F*(1, 87) = 6.75, *p* = .011, *d* = .59. When rating how frustrating the task was, participants also found the non-depleting versions less frustrating than the depleting versions, *F*(3, 155) = 35.72, *p* < .001; the differences within subsets did not reach significance. Participants completing the non-depleting versions also found them somewhat more pleasant, *F*(3, 154) = 3.51, *p* = .017, and felt less pressure during the task, *F*(3, 159) = 11.36, *p* < .001, than did those completing the depleting versions. Finally, participants in the non-depleting conditions felt they were able to master the task to a higher degree than those in the depletion conditions, *F*(3, 157) = 30.73, *p* < .001. Again in line with their performance, a focused comparison revealed that those completing the depleting version with 50 exceptions reported greater mastery than those completing the depleting version with 25 exceptions, *F*(1, 80) = 10.62, *p* = .002, *d* = .73.

Table 19. Ratings of the initial Number SART task in Study 3

	Difficulty		Frustration		Pleasantness		Pressure		Mastery	
Condition	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*
Non-depleting - no pressing	1.21	0.47	2.73	2.3	5.85	2.37	2.58	2.53	8.03	2.52
Non-depleting - all pressing	2.37	1.53	2.38	1.48	5.98	2.03	3.41	2.18	7.27	2.15
Depleting - 50 exceptions	5.38	2.2	5.33	2.19	4.95	1.23	4.78	2.21	5.25	2.46
Depleting - 25 exceptions	6.60	2.01	6.5	2.34	5.05	1.15	5.21	2.29	3.67	1.92

Depletion. Next, participants' performance on the subsequent Letter SART task was analyzed to explore whether depletion actually occurred (see Table 20). First, one-way ANOVAs were conducted to compare performance by the four initial task conditions (i.e., Number SART versions). Then, mood measure was added as an additional factor in 2-way ANOVAs to explore potential interactions.

Table 20. Letter SART depletion means and effect sizes for Study 3

					Partial Eta²	
Variable	Condition	*N*	*M*	*SD*	Initial task	Initial task x mood measure
Number of False Alarms	Non-depleting - no pressing	38	9.45	4.27	.11	.02
	Non-depleting - all pressing	38	13.74	4.81		
	Depleting - 50 exceptions	40	9.40	5.78		
	Depleting - 25 exceptions	41	11.07	5.60		
Percentage Missed	Non-depleting - no pressing	37	1.22	1.75	.09	.16
	Explicit SAM	20	1.80	2.07		
	Implicit Word Completion	17	0.53	0.94		
	Non-depleting - all pressing	37	2.89	3.94		
	Explicit SAM	19	5.00	4.50		
	Implicit Word Completion	18	0.67	1.13		
	Depleting - 50 exceptions	34	0.71	1.19		
	Explicit SAM	18	1.11	1.49		
	Implicit Word Completion	16	0.25	0.45		
	Depleting - 25 exceptions	40	1.50	2.25		
	Explicit SAM	20	1.00	1.45		
	Implicit Word Completion	20	2.00	2.79		
Reaction Time	Non-depleting - no pressing	38	321.16	63.10	.13	.06
	Explicit SAM	19	295.74	41.55		
	Implicit Word Completion	19	346.58	71.37		
	Non-depleting - all pressing	37	272.05	43.32		
	Explicit SAM	18	255.78	54.21		
	Implicit Word Completion	19	287.47	21.50		
	Depleting - 50 exceptions	38	332.13	83.13		
	Explicit SAM	20	342.90	100.64		
	Implicit Word Completion	18	320.17	58.61		
	Depleting - 25 exceptions	39	292.82	51.89		
	Explicit SAM	19	298.42	49.68		
	Implicit Word Completion	20	287.50	54.63		

False alarms. First, participants were scored on their number of false alarms, i.e., the number of times they pressed the button when they should not have because the item appearing was a letter 'D' (possible range = 0-25). There was a significant effect of initial task version on this measure, $F(3, 153) = 5.96$, $p = .001$. Post-hoc Tukey HSD comparisons indicated that participants who completed the non-depleting Number SART wherein they only watched the numbers *and* those who earlier completed the depleting version with 50 exception trials made significantly fewer false alarm responses on the Letter SART than did those who pressed for all trials on the Number SART. Those who had 25 exception trials on the Number SART fell in the middle and did not differ significantly from either subset. Finally, adding mood measure as an additional factor did not produce a significant interaction, $F(3, 149) = .82$, $p = .49$.

Percentage missed. Next, participants were scored on the percent of trials that they missed, i.e., the percentage of time they failed to press the button when they should have done so because the item appearing was any letter other than 'D.' We saw a significant effect of initial task version here as well, $F(3, 144) = 4.97$, $p = .003$. Tukey HSD post-hoc comparisons again showed that participants who completed the non-depleting Number SART where they watched the numbers *and* those who earlier completed the depleting version with 50 exception trials missed significantly *fewer* trials on the Letter SART than those who pressed for all trials on the initial Number SART. As with false alarms, those who had 25 exception trials did not differ significantly from either subset.

Here, however, adding mood measure type as an additional factor did produce a significant interaction, $F(3, 140) = 9.17, p < .001$. As seen in Figure 15, for participants who completed SAMs, those pressing for all items in the Number SART missed more than those in the other conditions. However, for those who completed the implicit word completion tasks, participants' performances were more consistent, and those who completed the Number SART with 25 exception trials missed the most trials on the Letter SART.

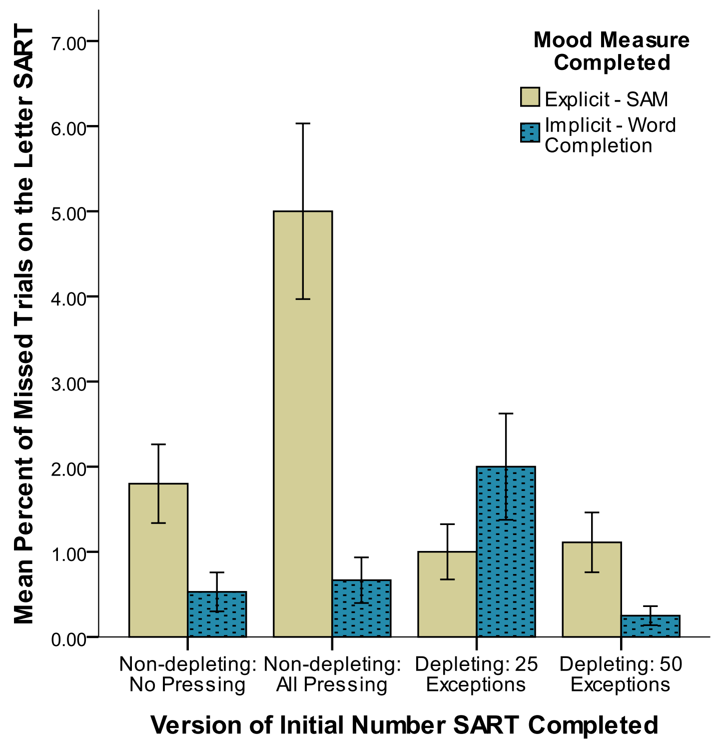

Figure 15. Percentage of trials missed on the Letter SART task by depletion condition and mood measurement in Study 3. Error bars extend ± one standard error.

Reaction Time. Finally, we analyzed how quickly participants responded to trials on the Letter SART using their median reaction times. There was again a significant effect of initial task version, $F(3, 148) = 7.22, p < .001$. Participants who earlier completed the non-depleting Number SART wherein they pressed for all trials were fastest on the Letter SART, followed by those who completed the depleting version with 25 exceptions, then those with the non-depleting version with no pressing, and finally, those with the depleting version with 50 exceptions. Adding mood measure type as an additional factor again produced a significant interaction, $F(3, 144) = 3.10$, $p = .029$; see Figure 16. Whereas there was comparatively little difference in reaction times for the depleted participants who were given the implicit versus explicit measure of mood, for the non-depleted participants, responding tended to be faster if they had completed the explicit rather than the implicit assessment of mood.

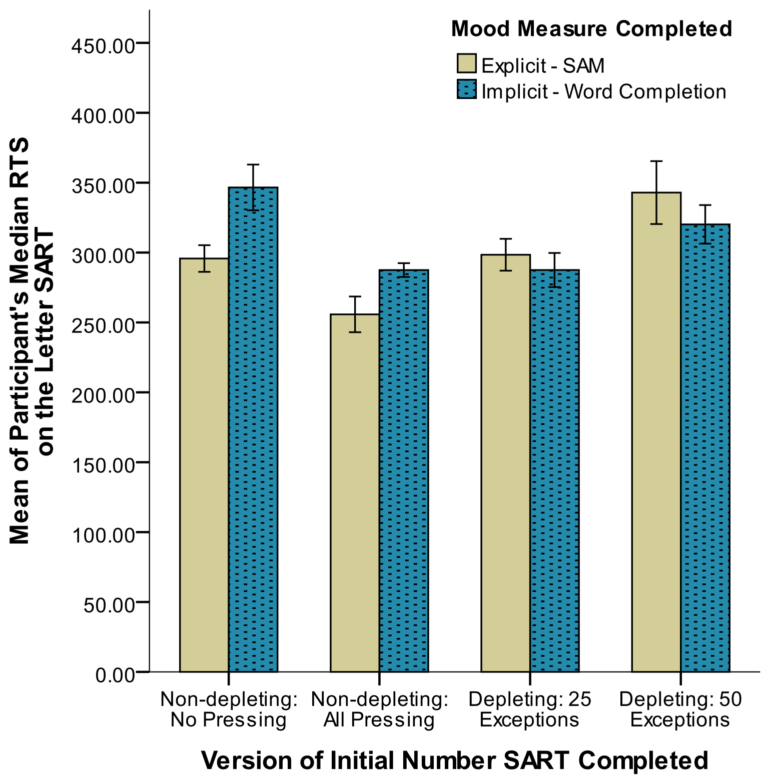

Figure 16. Mean median reaction time to trials on the Letter SART task by depletion condition and mood measurement in Study 3. Error bars extend ± one standard error.

Mood – self report. Participants provided self-report data on their mood before and after the initial Number SART task either explicitly via SAM ratings or implicitly via the word completion tasks. As in previous studies, we first assessed the effects of the initial task on participants' post-task mood ratings, this time using one-way ANOVAs given that we had four versions of the initial task. We then assessed the effects of the initial task on participants' changes in mood from pre- to post-task using two-way ANOVAs. Tables 21 and 22 present the means and effect sizes for these analyses.

SAMs. In this study, there was no significant effect of depletion condition on participants' mood valence ratings after the Number SART task, $F(3, 77) = .12$, $p = .95$. Unlike many of our previous studies, we also did not see a significant time of measurement x depletion condition interaction on participant's mood valence, $F(3,76) = .03$, $p = .99$. As seen in Figure 17, all of the Number SART task variations lead equally to modest decreases in mood valence.

There was, however, a marginally significant effect of depletion condition on participants' mood arousal ratings after the Number SART task, $F(3, 76) = 2.66$, $p = .054$. Tukey HSD post-hoc comparisons showed that those who completed the non-depleting version requiring no pressing had significantly lower arousal rates than those completing the depleting version with 50 exception trials. Those completing the non-depleting version requiring presses for all items and the depleting version with 25 exception trials fell in the middle and did not differ significantly from either extreme.

We also found a significant time of measurement x depletion condition interaction on participant's mood arousal, $F(3, 75) = 6.98$, $p < .001$, akin to those found previously.

Table 21. SAM mood means and effect sizes for Study 3

						Partial Eta²	
Mood Variable	When	Condition	*N*	*M*	*SD*	Depletion condition	Depletion condition x Time of measurement
SAM: Valence	Before task	Depleted – 25 exceptions	20	6.55	0.95	.003	.001
		Depleted – 50 exceptions	20	6.60	1.19		
		Non-depleted – all presses	20	6.50	0.95		
		Non-depleted – no presses	20	6.65	0.99		
	After task	Depleted – 25 exceptions	20	6.00	1.34	.004	
		Depleted – 50 exceptions	20	6.00	1.26		
		Non-depleted – all presses	20	5.95	1.10		
		Non-depleted – no presses	20	6.15	0.99		
SAM: Arousal	Before task	Depleted – 25 exceptions	21	3.90	1.64	.07	.22
		Depleted – 50 exceptions	19	4.79	0.98		
		Non-depleted – all presses	19	4.47	1.17		
		Non-depleted – no presses	20	4.95	1.96		
	After task	Depleted – 25 exceptions	21	4.29	1.49	.10	
		Depleted – 50 exceptions	19	5.16	1.26		
		Non-depleted – all presses	19	4.11	1.33		
		Non-depleted – no presses	20	3.90	1.77		

Note: Again, higher scores on valence do indicate a more positive mood, and higher scores on arousal indicate a more aroused mood.

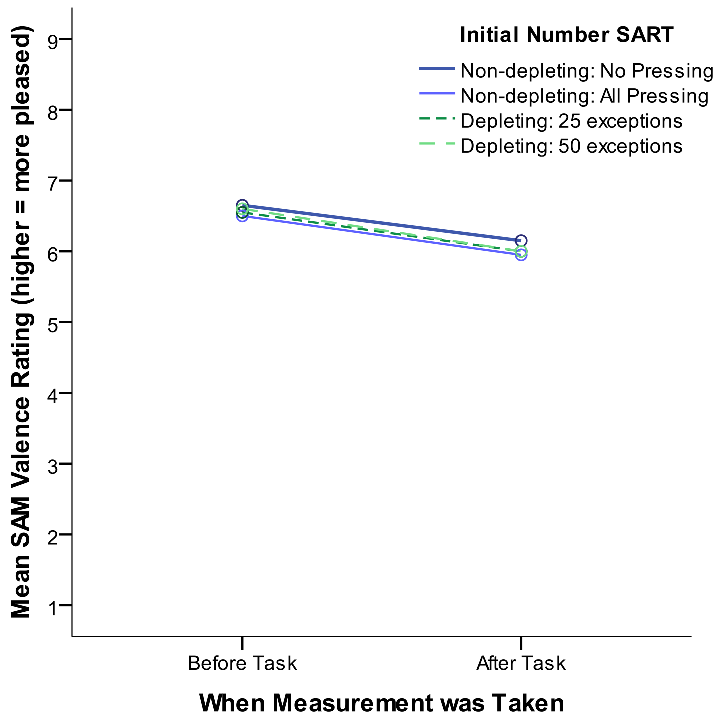

Figure 17. Mean SAM valence rating depletion condition and time of measurement in Study 3.

We also found a significant time of measurement x depletion condition interaction on participant's mood arousal, *F*(3, 75) = 6.98, *p* <.001, similar to those found in We also found a significant time of measurement x depletion condition interaction on participant's mood arousal, *F*(3, 75) = 6.98, *p* <.001, similar to those found in previous studies. Figure 18 illustrates that while the non-depleting versions of the Number SART task left participants feeling somewhat calmer, the depleting versions of the task somewhat increased participants' arousal levels.

Word Completions. Participants' word completion tasks were initially scored in a basic manner, awarding +1 for each fragment the participant completed with a word with positive connotations, -1 for each word with negative connotations, and zero for neutral words. However, significant effects were not detected using this scoring.

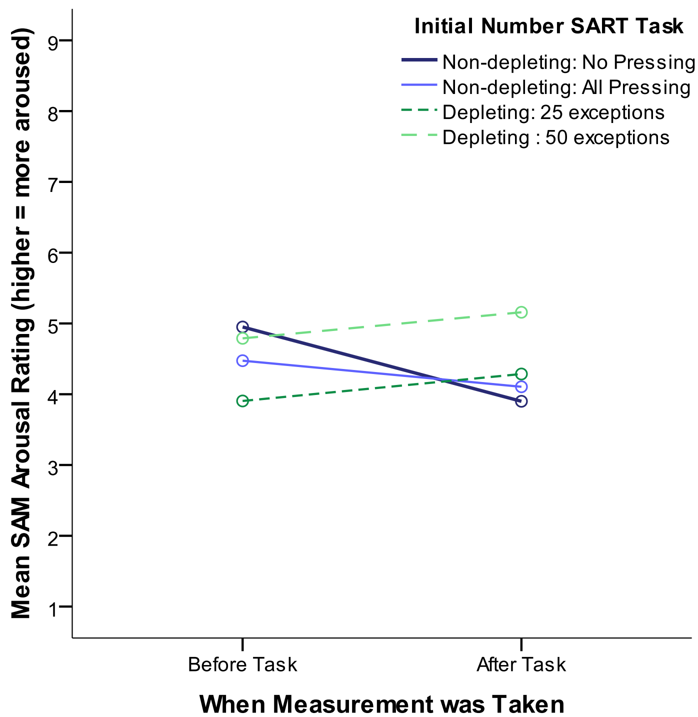

Figure 18. Mean SAM arousal rating by depletion condition and time of measurement in Study 3.

Because of the possibility that the lack of effects was due to an insensitivity of our measure, we conducted a norming study (see Appendix A) to determine the exact degree of valence and arousal connotations associated with each of the possible completions of each word fragment. We then used these normative data to re-calculate participants' implicit mood scores. For each participant, the average valence and average arousal for each word completion task was determined.[14]

[14] Averages were used here rather than totals to avoid the necessity of assigning a value to an item when the participant completed the fragment with a non-word.

Table 22. Word completion mood means and effect sizes for Study 3

						Partial Eta²	
Mood Variable	When	Condition	*N*	*M*	*SD*	Depletion condition	Depletion condition x Time of measurement
Word Completion: Valence	Before task	Depleted – 25 exceptions	19	5.26	0.26	.04	.01
		Depleted – 50 exceptions	19	5.40	0.27		
		Non-depleted – all presses	21	5.29	0.27		
		Non-depleted – no presses	20	5.27	0.27		
	After task	Depleted – 25 exceptions	19	5.49	0.26	.05	
		Depleted – 50 exceptions	19	5.61	0.30		
		Non-depleted – all presses	21	5.41	0.34		
		Non-depleted – no presses	20	5.47	0.34		
Word Completion: Arousal	Before task	Depleted – 25 exceptions	19	4.69	0.17	.04	.02
		Depleted – 50 exceptions	20	4.67	0.20		
		Non-depleted – all presses	21	4.58	0.19		
		Non-depleted – no presses	20	4.64	0.25		
	After task	Depleted – 25 exceptions	19	4.43	0.15	.02	
		Depleted – 50 exceptions	20	4.44	0.14		
		Non-depleted – all presses	21	4.39	0.17		
		Non-depleted – no presses	20	4.41	0.13		

Note: Again, higher scores on valence do indicate a more positive mood, and higher scores on arousal indicate a more aroused mood.

There was not a significant effect of depletion condition on participants' implicit mood valence scores after the Number SART task, $F(3, 75) = 1.35$, $p = .27$. We also found no significant time of measurement x depletion condition interaction on participants' implicit mood valence scores, $F(3, 75) = .29$, $p = .83$. Figure 19 illustrates participants' scores on these measures; from the graph, it appears that all of the Number SART tasks were associated with a small "increase" in mood valence. However, remember that we used different word completion stimulus sets before and after the tasks; thus, the overall "increases" here may more likely be attributed to the differences in the word stems we provided on the two versions than to the Number SART tasks actually increasing mood valence.

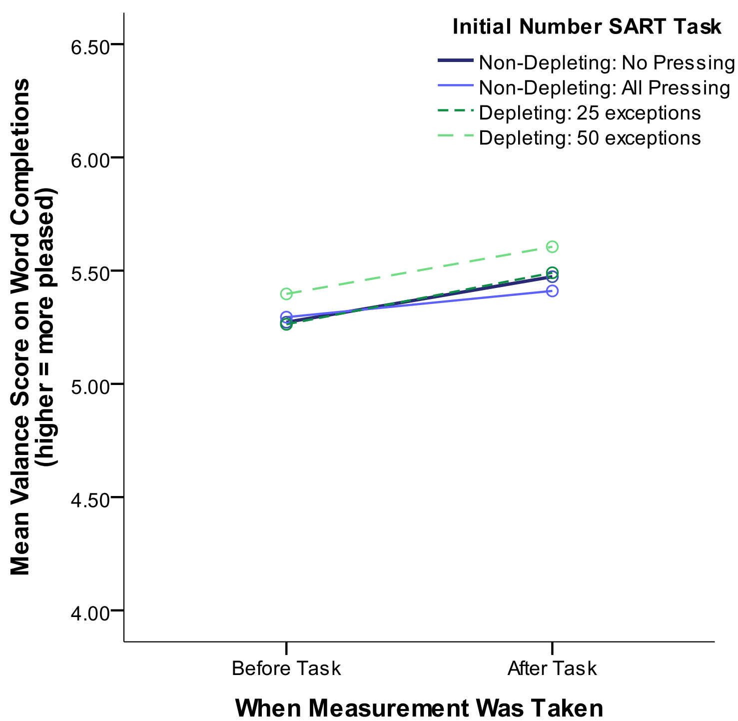

Figure 19. Mean Word Completion valence score by depletion condition and time of measurement in Study 3.

There was also not a significant effect of depletion condition on participants' implicit arousal scores after the Number SART task, $F(3, 78) = .47$, $p = .71$. We also found no significant time of measurement x depletion condition interaction on participants' implicit arousal scores, $F(3, 76) = .42$, $p = .74$. As seen in Figure 20, it appears that all of the Number SART tasks were associated with a small "decrease" in arousal. However, again, as these scores were obtained using two different word completion stimulus sets, the "decreases" here may be ascribed to the differences in the word stems provided.

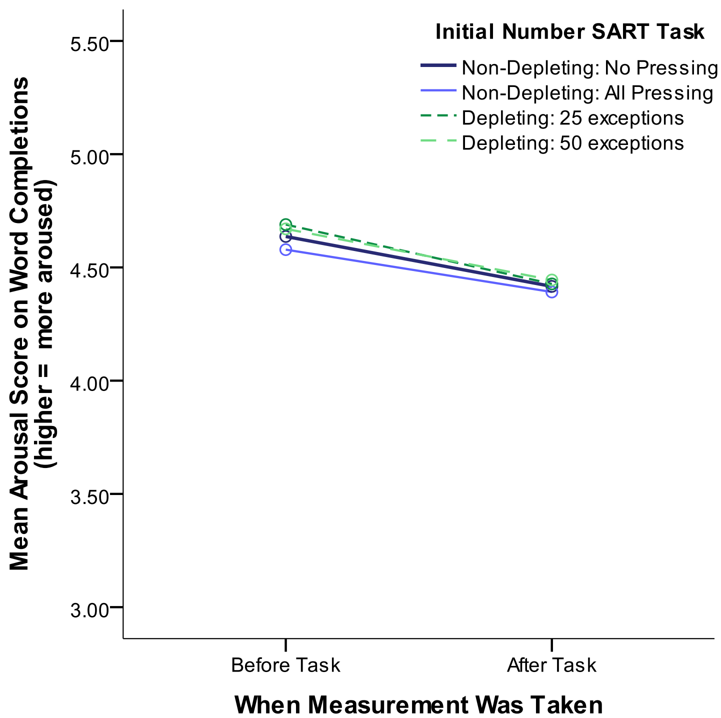

Figure 20. Mean Word Completion arousal score by depletion condition and time of measurement in Study 3.

Mood – biological indices. Participants also provided biological indices of mood arousal throughout the study via HR and BP measures. As before, we first assessed the effects of the initial task on their post-task physiological readings using one-way ANOVAs. We then assessed the effects of the initial task on their changes from pre- to post-task physiological reading using two-way ANOVAs. Tables 23, 24, and 25 have the means and effect sizes for these analyses.

Table 23. Heart rate means and effect sizes for Study 3

					Partial Eta²	
When	Condition	*N*	*M*	*SD*	Depletion condition	Depletion condition x Time of measurement
Before task	Depleted – 25 exceptions	41	77.07	11.86	.02	.002
	Depleted – 50 exceptions	40	79.20	14.38		
	Non-depleted – all presses	40	74.73	12.07		
	Non-depleted – no presses	39	74.56	11.12		
After task	Depleted – 25 exceptions	41	76.22	11.09	.02	
	Depleted – 50 exceptions	40	78.83	14.43		
	Non-depleted – all presses	40	74.70	10.46		
	Non-depleted – no presses	39	74.44	11.88		

Heart rate. We found no significant effect of depletion condition on post-task measurements of heart rate, $F(3, 156) = 1.11$, $p = .35$. Additionally, although the linear trend was in the expected direction, with numerically lower heart rates observed for the two non-depleted than the two depleted conditions, the two depleted groups also began (before the task) with higher heart rates. The time of measurement x depletion condition interaction on participant's heart rate was also non-significant, $F(3, 156) = .12$, $p = .95$, see Figure 21. Neither was there any three way-interaction found when adding mood measure as an additional factor, $F(3, 152) = .74$, $p = .53$.

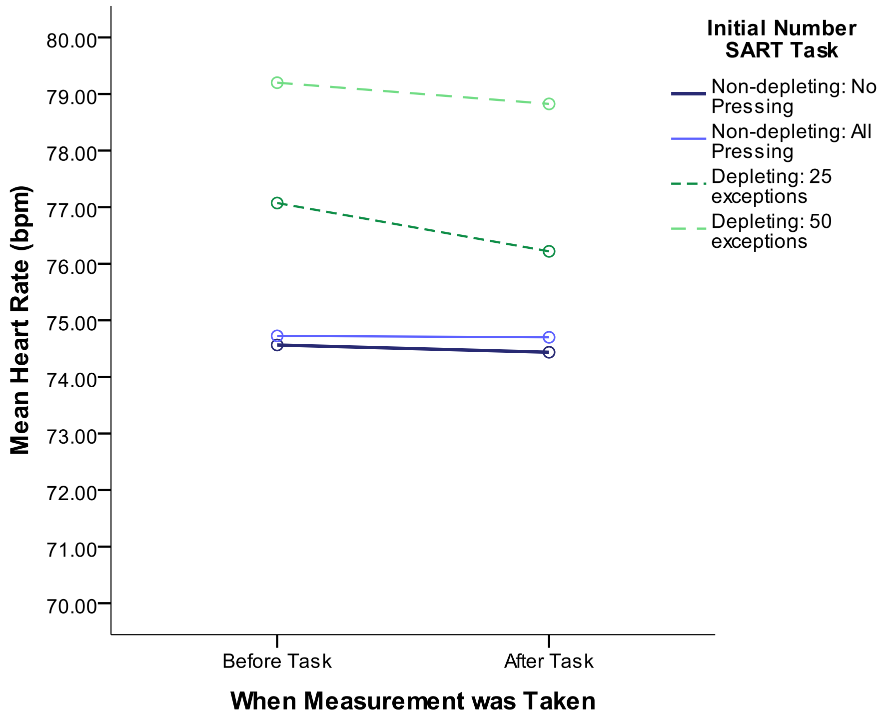

Figure 21. Mean heart rate reading by depletion condition and time of measurement in Study 3.

BP – Systolic. We saw similar null effects in the blood pressure systolic readings. There was no significant effect of depletion condition on post-task systolic BP measurements, $F(3, 155) = .64$, $p = .59$. The interaction of depletion condition and time of measurement was also non-significant, $F(3, 151) = .85$, $p = .47$, with trends for participants' systolic BP to decrease from pre- to post-task measurements for participants completing all initial task versions (see Figure 22). Neither was there any three way-interaction found when adding mood measure type as an additional factor, $F(3, 147) = 1.22$, $p = .30$.

Table 24. Systolic blood pressure means and effect sizes for Study 3

					Partial Eta^2	
When	Condition	*N*	*M*	*SD*	Depletion condition	Depletion condition x Time of measurement
Before task	Depleted – 25 exceptions	40	110.53	9.97	.06	.02
	Depleted – 50 exceptions	40	112.15	10.76		
	Non-depleted – all presses	37	106.30	9.26		
	Non-depleted – no presses	38	112.47	9.00		
After task	Depleted – 25 exceptions	40	107.80	9.90	.01	
	Depleted – 50 exceptions	40	108.90	10.14		
	Non-depleted – all presses	37	104.78	7.99		
	Non-depleted – no presses	38	108.32	9.37		

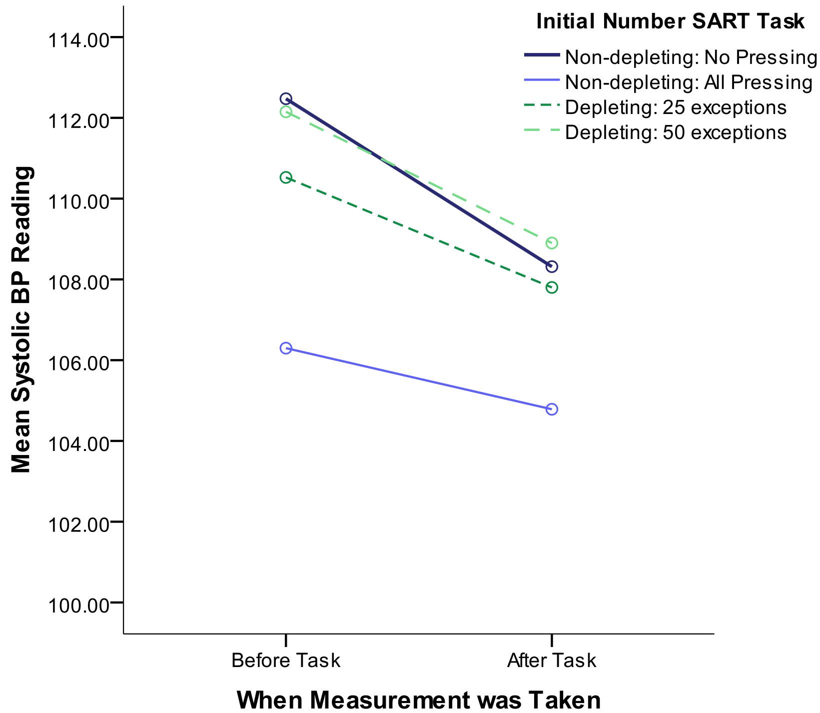

Figure 22. Mean systolic blood pressure reading by depletion condition and time of measurement in Study 3.

BP – Diastolic. Similar trends continued for the diastolic BP readings. There was not a significant effect of depletion condition on post-task diastolic BP measurements, $F(3, 152) = .91$, $p = .44$. Again, the interaction of time of measurement and depletion condition did not reach significance, $F(3, 152) = 1.96$, $p = .122$, with unexpected trends (e.g., those completing the depleting version with 25 exceptions demonstrated the largest reduction in BP from pre- to post-task measurements). There was, however, a significant three way-interaction found when adding mood measure type as an additional factor, $F(3, 148) = 3.81$, $p = .011$, partial eta^2 = .08. As seen in Figure 23, only those participants who completed the non-depleting version that required presses for all trials and the word completion tasks showed an increase in diastolic blood pressure. All other groups showed varying degrees of reductions in diastolic blood pressure.

Table 25. Diastolic blood pressure means and effect sizes for Study 3

					Partial Eta^2	
When	Condition	*N*	*M*	*SD*	Depletion condition	Depletion condition x Time of measurement
Before task	Depleted – 25 exceptions	39	70.00	9.17	.06	.04
	SAMs	20	67.50	7.37		
	Word Completions	19	72.63	10.28		
	Depleted – 50 exceptions	40	68.63	8.68		
	SAMs	20	70.40	7.45		
	Word Completions	20	66.85	9.61		
	Non-depleted – all presses	40	63.63	9.22		
	SAMs	20	63.85	6.74		
	Word Completions	20	63.40	11.36		
	Non-depleted – no presses	37	68.43	9.88		
	SAMs	19	68.79	10.66		
	Word Completions	18	68.06	9.27		
After task	Depleted – 25 exceptions	39	65.21	7.66	.02	
	SAMs	20	66.45	7.69		
	Word Completions	19	63.89	7.62		
	Depleted – 50 exceptions	40	65.95	9.23		
	SAMs	20	67.40	9.20		
	Word Completions	20	64.50	9.27		
	Non-depleted – all presses	40	63.00	8.41		
	SAMs	20	61.20	8.15		
	Word Completions	20	64.80	8.48		
	Non-depleted – no presses	37	64.46	7.90		
	SAMs	19	64.89	8.68		
	Word Completions	18	64.00	7.21		

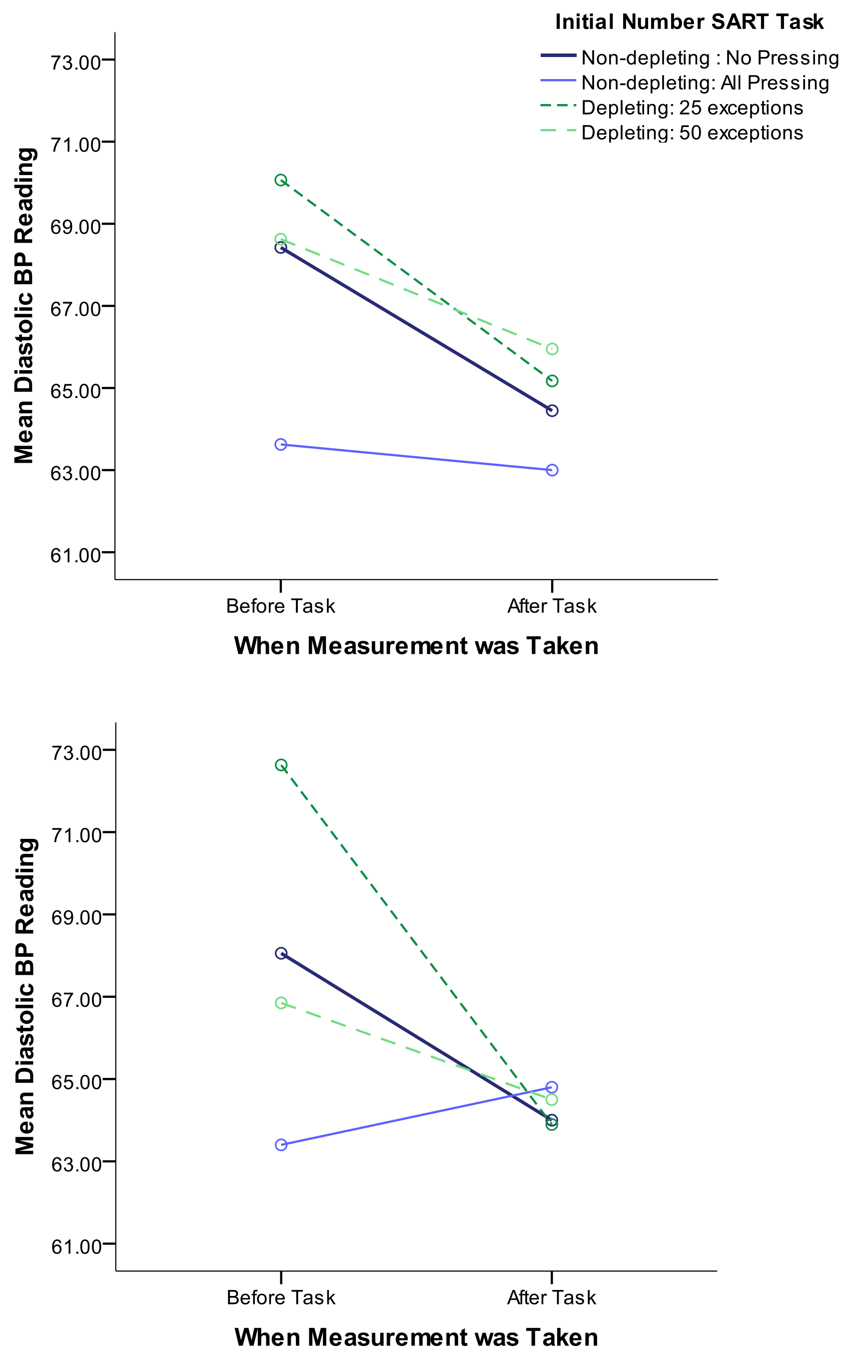

Figure 23. Mean diastolic blood pressure reading by depletion condition and time of measurement for those completing SAMs (top) and word completions (bottom) in Study 3.

Study 3 Discussion

The results of this final study have provided several new and important findings regarding the conditions that evoke self-regulation depletion and the effects of self-regulation depletion. We focus here on: *(1)* the evidence that completing different versions of the SART task required differing amounts of initial self-regulation, *(2)* the effects these different tasks had on subsequent self-regulation, and *(3)* the variability in the effects of self-regulation on mood as measured by the explicit, implicit and biological indices.

First, the different versions of the initial Number SART task did appear to differentially require self-regulation. This was reflected both by participants' ratings of the tasks and by their performance on them. As intended, the non-depleting versions were rated as easier, and performance was typically better on them. When contrasting the two depleting versions, interestingly, performance was worse on the SART 25 than on the SART 50 and participants' ratings of the two were consistent with this. That is, those completing the SART 50 had to carry out *more* instances of self-regulation, yet they felt the task was easier and performed better. This seems in conflict with the likely prediction of the self-regulation depletion theory. One would expect that more self-regulation would be more challenging and more likely to induce errors. However, a more broad interpretation is needed here, incorporating the strength of the habit being overcome. In the SART 50, the proportion of exception to non-exception trials was higher, weakening the habit of pressing that participants needed to surmount. This interpretation is supported by the greater number of false positive responses made by participants in their first 25 exception trials in the SART 50 as compared to the SART 25.

In congruence with the differences in performance and ratings of the initial Number SARTs, we also saw effects of our initial task manipulations on each of our three subsequent self-regulation measurements on the Letter SART: false alarms, misses, and reaction time. In light of the inconsistencies in depletion in Studies 1 and 2, this was exciting. However, the effects seen here in Study 3 would not be best characterized as illustrating depletion. Those who completed the non-depleting Number SART who pressed for all items were “able” to respond most quickly to items on the Letter SART, suggesting the possibility of a depletion effect. However, those who completed the other non-depleting Number SART with no required pressing actually had slower reaction times on the Letter SART than those how had completed the depleting Number SART 25. Furthermore, taken together with the accuracy data, the more valid conclusion appears to be that those who completed the non-depleting constant button pressing Number SART had a harder time switching over to the Letter SART in a way that allowed them to balance their speed with accuracy.

Regarding our mood findings, when examining participants’ explicitly reported mood, we found no evidence of an effect of the initial self-regulation task on participants’ post-task or change from pre- to post-task mood valence. However, as seen in most of our previous studies, we again found effects on participants’ arousal ratings, with the depleting versions heightening arousal and the non-depleting versions calming participants. Unsurprisingly, the most calming task was the non-depleting no-pressing condition.

This pattern of mood findings was not consistently supported by our biological mood indices. On one hand, blood pressure, which has been associated with negative affect (Ilies et al., 2010) and our direct measure of mood valence both illustrated non-significant effects of the initial task. However, effects on heart rate, which has been associated with arousal (Ilies et al., 2010) did not mirror the effects on arousal as measured by the SAM. One possible reason for our lack of clear findings for our biological measurements might be the likely noisiness of the data due to the various other factors contributing to HR at any given time. For instance, while Ilies et al. (2010) found HR to be significantly correlated with within-subject variability in positive affect and negative affect (both of which involve arousal), the size of the relationships were by no means large (rs = .08 and .05 respectively).

We also did not find evidence of significant mood effects using our new implicit word completions measurement. Again, while the lack of effect on the valence dimension was in accord with our findings from the explicit measure, the lack of effect on arousal was in conflict with the significant effects for arousal as measured by the SAM. This inconsistency may be explained by the design of our new word completions measure. In particular, the word fragments included in these tasks were selected in order to ensure the potential for variability on the valence aspect of emotion. The same consideration was not given to ensuring variability on the arousal aspect. Standard deviations from the valence (M = .29) and arousal dimensions (M = .18) do suggest a possible restriction of range on the arousal scale. We were also unable to directly test the convergent validity of this new measure, as no participants completed both the word completion tasks and the explicit measures using the SAM. Development and validation of sensitive implicit assessments of both affective valence and affective arousal remains a fruitful direction for future research. Whereas a number of ways to behaviorally measure implicit mood valence have been developed (e.g., deWall & Baumeister, 2007; Koole et al., 1999; Quirin et al., 2009; Sheppes & Meira, 2007) and various biological recordings have been used to implicitly measure arousal (e.g., Lang, Greenwald, Bradley, & Hamm, 2007), few, if any, measures have been designed in order to implicitly measure *both* aspects of mood.

General Discussion

Overall, this series of experiments has shed new light on the theory of self-regulation depletion. The implications of our findings fall into three broad categories, outlined below. First, conclusions about the phenomenon of self-regulation depletion itself are summarized. Then, we discuss the implications of our results regarding ways to counteract depletion. Next, a summary of the mood effects we found and their meaning is provided. Finally, we provide a summary of our main conclusions.

Self-Regulation Depletion

The studies reported here, overall, lead to the conclusion that self-regulation depletion is less omnipresent than one might assume from a reading of the literature. We did obtain depletion on some measures in some of the studies, but never in all measures and not in all studies. The contrast between our findings and those in the published literature likely is due, at least in part, to what is referred to as the "file drawer" problem (Rosenthal, 1979). As such, null results are often filed away, whereas significant effects are submitted and published. This is not to say that the frequency of self-regulation depletion that is reported is solely the result of publication bias and not a true effect – the volume of research in support of depletion clearly indicates otherwise. Our findings simply act as a caution against an over-reliance on the power of the phenomenon, and point to the potential contribution of many contextual factors that may influence when, and if, self-regulatory depletion effects are observed. These conclusions are in accord with the results of the meta-analysis by Haggar et al. (2010), illustrating that published self-regulation depletion effects have an average effect size of $d = .62$, but also that various factors do appear to moderate the magnitude of observed depletion effects, as shown by significant heterogeneity of effect sizes as well as direct tests for moderation. The factors found to significantly moderate depletion effects included the duration of the initial self-regulation task, the amount of time elapsing between the initial and subsequent self-regulation tasks, and the nature of the subsequent self-regulation task.

While our many null results regarding self-regulatory depletion were disappointing to us as researchers hoping, in part, to explore replenishment options, they are actually reassuring when considering practical, real-world implications. It is comforting to know that engaging in self-regulation for such brief periods (e.g., 6 minutes) does not guarantee that your future performance on additional tasks requiring self-regulation will suffer. In light of this, more research is needed in applied academic and industry settings, akin to the experiment by Goldberg and Grandey (2007) examining emotional regulation in call centers and its impact on task performance. These researchers observed that being required to regulate one's emotions in order to always provide "service with smile" led to higher rates of errors on order forms. More studies of this type would help to establish exactly what real-life tasks lead to depletion, how large these depletion effects are, and which subsequent tasks requiring self-regulation are most likely to be impacted. Such studies could also provide information regarding the extent to which, and the conditions under which, self-regulatory depletion emerges when the broader rationale for engaging in particular tasks is neither as arbitrary nor as potentially unconnected with an individual's ongoing goals and pursuits as may be the case in many laboratory-based studies that typically involve tasks with little inherent or longer-term meaning for participants beyond satisfying an experimental requirement (cf. Vohs & Heatherton, 2000; Vohs et al., 2008).

Another potentially heartening aspect to consider is that while depletion may be relatively common in many laboratory settings, it may be less common or slower to develop in everyday life. As seen in our Study 1.1 and Study 3, even when our participants were completing a relatively simple task (watching a video of waves or pressing a button for all items), they appeared to need to self-regulate somewhat in order to sustain their attention to a task they were not much interested in. Another way of saying this is that participants are often extrinsically motivated to engage in the required tasks in laboratory studies. In contrast, people are often (though certainly not always!) more intrinsically motivated to regulate themselves in their daily lives. As shown by Muraven and colleagues, depletion is heightened when autonomy is low – i.e., when participants feel obligated to complete the initial task for a reward (Muraven et al., 2007) and mitigated when autonomy is high, i.e., when participants feel that they have more of a choice in the matter (Muraven et al., 2008). We found similar trends in Study 2.1 when attempting to explore exercise as a replenisher – those who were required to complete exercises tended to actually perform worse on the final self-regulation task.

Taken together, these findings emphasize the importance of encouraging autonomy and intrinsic motivations whenever self-regulation or executive functioning is essential. Note, however, that there are different degrees of autonomy or self-determination associated with extrinsic motivation. In their self-determination theory, Ryan and Deci (2000 a, b) define four levels of extrinsic motivation: external, introjected, identified, and integrated. The first two – external and introjected – are low in self-determination/autonomy, involving self-regulation for the sole purpose of gaining some concrete reward or avoiding feelings of guilt or anxiety. At the identified regulation level, acts of regulation become more autonomous. People engage in acts of identified and integrated regulation because they further some goal that the person has come to value and to hold as personally relevant. Such autonomous extrinsic motivation appears to have the potential to enhance academic performance (Burton, Lydon, D'Alessandro, & Koestner, 2006) as well as creativity (Kasof, Chen, Himsel, & Greenberger, 2007). Indeed, Burton et al. (2006) argue that it is identified motivation that encourages performance, whereas intrinsic motivation enhances feelings of well-being. Thus, it may not be how intrinsically enjoyable a controlled act is, but how closely it is tied to important self-held goals that matters most for successful self-regulation.

The results of our first experiment (Study 1.1), showing trends toward anti-depletion are also interesting to note. Recent evidence from Converse and DeShon (2009) provides one potential explanation. They were interested in resolving two conflicting bodies of research. One, the self-regulation depletion theory, we have covered in great depth already. The other, often referred to as learned industriousness, holds that participants can learn that high levels of effort will be required from initial self-regulation tasks, and therefore, actually perform *better* on subsequent tasks than participants who had lower demands placed on initial self-regulation. Converse and DeShon particularly hypothesized that the differences in effects could be explained by differences in the number of initial tasks participants were asked to complete in studies examining the two models. When a two-task paradigm (one initial task, one subsequent task) is used, depletion is the typical result. However, studies on learned industriousness often use more than one initial task and find heightened self-regulation as a result. Converse and DeShon's results were in line with their hypothesis – "anti-depletion" was found when participants completed two initial high-self-regulation tasks, whereas depletion was found when they completed only one. Relating this to our Study 1.1 findings, our "depleted" participants might be thought of as having completing two initial high self-regulation tasks – the waves video pre-task and the restricted writing or operation span task. As such, one might have expected them to perform as they did – better on the final task than those who had completed only one initial high self-regulation task (waves video) and one initial low self-regulation task (unrestricted writing or simple memory and math). We need to be cautious in drawing such conclusions, given the nonsignificance of the trends. However, if such effects could be verified and replicated, the results would have strong implications for potential means of counteracting depletion (see the section below on counteracting depletion for more on this idea).

From the instances where we did find significant depletion, other inferences can be drawn. For instance, the way in which depletion manifests varied depending on the task. In Studies 1.2 and 1.4, non-depleted participants showed more speed and/or persistence by answering more items within the time limit, and in doing so, therefore answered more items correctly. However, their percentage correct did not differ. In contrast, in Study 2.2, we did not see any depletion effects when analyzing speed (i.e., reaction time) data, but instead found significant effects on the accuracy measures for the flanker task and the math aspect of the operation span task. Similar contrasts are present in the published literature on self-regulation depletion. For example, self-regulation depletion appears to manifest itself sometimes as an increase in activity, as evidenced by the larger amounts of ice cream eaten by dieters in Vohs and Heatherton (2000) and, at other times, as a decrease in activity, as in the tendency for depleted participants to passively watch a broken film as reported in Vohs et al. (2008). In the future, researchers may want to more systematically examine the factors that influence the way in which self-regulation depletion manifests (e.g., aspects of the depleting task, aspects of the subsequent task, as well as the consistency of depletion across different outcome measures within a particular subsequent task).

The results of Study 3 in particular demonstrate interesting new findings regarding depletion. Chiefly, they reveal that depletion effects can be surmounted by practice effects if the two self-regulation tasks are similar in nature. Those who completed 50 exception trials on the initial Number SART task were able to perform as well on the subsequent Letter SART task as those who were only asked to watch the numbers on the screen. This clearly contrasts with previous research wherein more initial self-regulation produces poorer performance (i.e., depletion) on the subsequent task.

This novel finding likely stems from differences in methodology. Research on self-regulation depletion has almost exclusively utilized paradigms wherein the initial and subsequent self-regulation tasks are dissimilar in order to demonstrate that the self-regulation resource is global rather than task-specific. In contrast, in Study 3 we purposely sought to explore the consequences of using initial and subsequent tasks that were quite similar to each other. Undoubtedly, more work is needed along similar lines. Questions remain as to how similar the two tasks need to be and how much practice is needed in order to overcome depletion. Our findings suggest that instances of practice may be more important than the exact similarity in tasks. Those who completed the initial Number SART 25, which was most similar to the subsequent Letter SART (in that both had 25 exception trials) tended to perform more poorly than those who completed the initial Number SART 50, which was less similar but allowed for more practice at inhibiting responses to exception trials.

More broadly, these outcomes may also provide evidence for the potential benefits of arranging successive tasks such that initial tasks provide more opportunities for *successful* self-regulation—that is, both opportunities for practice, and for positive feedback regarding one's self-regulation efforts. This then may further suggest connections between our practice-related findings and recent longer-term training regimes aimed at extending particular aspects of executive capacity, such as working memory. For instance, in work with both young adults (Jaeggi, Buschkuehl, Jonides, & Perrig, 2008; Jaeggi, Studer-Luethi, Buschkuehl, Su, Jonides, & Perrig, 2010) and children (Jaeggi, Buschkuehl, Jonides, & Shah, 2011) researchers have demonstrated that executive capacities may be enhanced through systematic practice at working memory tasks if the individual is asked to continuously work at or near the "outer reaches" of their current capability.

Counteracting Depletion

Our examinations of potential replenishing activities yielded little in the way of positive findings. Only one significant effect was found – the effect of taking a break on participants' math accuracy on the operation span task in Study 2.2. As stated in our discussion earlier, caution is needed in interpreting this result, given the number of analyses performed and the fact that this was the only one of multiple effects examined to reach significance.

Our results highlight the difficulties researchers face in experimentally testing potential replenishing tasks. Even in Study 2.2, when participants were allowed to choose their own activities during a break, there was still little in the way of evidence for replenishment. One explanation for these results could be the nature of the experimental setting. In Study 2.2, we controlled for one such factor by removing the experimenter from the room during task completions – including the break tasks. This eliminated the potential for participants to feel self-conscious about being watched while engaging in their tasks. However, other factors likely remained. Participants recruited from introductory psychology classes, as ours were, often seem to suspect deception of some sort. As such, they may be trying to figure out what the "real" purpose of the study is. This problem solving in itself is likely expending cognitive resources. In addition, such ponderings likely add a self-monitoring aspect to what otherwise would be a low self-regulation task. For instance, our participants who were given a choice of break tasks to engage in may have wondered how their choices would be analyzed and how that would reflect on them. Similarly, participants who are asked to complete very basic tasks (e.g., watching a video) may wonder why they are doing so. Will they be tested on it later? These various additional musings of participants surely make what would otherwise be a cognitively light task more self-regulation heavy, reducing the likelihood of them producing replenishment in the lab. Researchers can try to minimize these risks through reassuring instructions, but given participants' penchant for seeing things as potentially deceptive, these may not be effective.

Other difficulties in examining replenishing tasks also arise. For example, consider Study 2.2 wherein we attempted to investigate the restorative properties of autonomy by manipulating choice. Although providing several activity options from which to choose could be construed as increasing participants' potential sense of autonomy, if the "choices" from the participant's point of view seemed to be just another (additional) experimental requirement, then the need to choose – rather than simply being assigned – an interim task to perform might itself place higher demands on self-regulation (e.g., Vohs et al., 2008; cf. Moller et al., 2006.). Concerns such as these imply that investigators wishing to further explore the self-regulation-rejuvenating properties of various tasks may be better served using field research.

Further consideration should also be given as to what the ideal self-regulation-restoring task would look like. Our assumption had been that completing some task requiring little self-regulation task in between the two high self-regulation tasks would most likely to be replenishing. We were unable to find much support for this hypothesis with our results. However, given the caveats above, this may not be so surprising.

What were unexpected were the implications of our Study 1.1 results wherein participants watched the waves pre-task. Assuming that the initial waves video task required significant amounts of self-regulation, our “non-depleted” participants in Study 1.1 would have engaged in what many would think of as a potentially replenishing series of tasks – high self-regulation (waves video), followed by low self-regulation (unrestricted writing or easy memory and math), followed by a high self-regulation measurement task. Yet, their performance was trending toward being *worse* than the “depleted” participants who had completed two initial high self-regulation tasks. This, taken together with the results of Converse and DeShon (2009) illustrating the occurrence of learned industrious rather than depletion when a three-task paradigm is used, suggests that using a low self-regulation task as a break may actually be counterproductive.

Assuming, however, that low self-regulation tasks can be restoring, thought should be given as to what the ultimate low or no self-regulation task could be. Many have argued for sleep as a way to restore self-regulation resources (e.g., Baumeister, 2002; Schmeichel & Baumeister, 2004), citing evidence that people are often better at self-regulation in the morning. Anecdotal evidence from citizens where afternoon siestas are common also suggests that these naps allow them to think more clearly and to stay more on-task in the afternoon. Research evidence further links sleep to higher order mental abilities. For example, Nilsson et al. (2005) found that as little as one night of sleep deprivation had significant detrimental effects on participants' executive functioning abilities. Sleep deprivation also has been found to impair executive functioning by leading participants to make more risky choices on the Iowa Gambling Task (Killgore, Balkin, & Wesensten, 2006). The monitoring aspect of self-regulation may be weakened by sleep deprivation as well; novelty detection as measured by the novelty P3 ERP component – important to controlled processing and linked to the frontal lobes – can be significantly impaired following sleep deprivation and restored following sleep (Gosselin, De Koninck, & Campbell, 2005). However, Baranski (2007) found that the accuracy of participants' confidence assessments was not compromised after 28 hours of sleep deprivation. Overall, the evidence suggests that sleep is likely to be a good way to replenish depleted self-regulation resources. On the other hand, practical limitations (e.g., finding a place to sleep, being able to fall asleep) may make it untenable as a way to counteract fluctuations in self-regulation abilities throughout the day.

Others have suggested that having one's attention spontaneously and effortlessly *drawn* to something – in contrast to having to purposefully and intentionally *direct* one's attention towards an item or task – can boost self-regulation. This is the essence of attention restoration theory (ART; Kaplan, 1995), which has especially focused on natural scenes as a "bottom-up" attention-drawing object (although other activities and stimuli, such as films or stories are also recognized as potential sources of comparatively restorative and bottom-up "fascination"). A number of studies have now been completed in support of the theory. For instance, Tennessen and Cimprich (1995) found that students with views of natural scenes out their windows performed better on a battery of directed attention tasks. Similar findings have been demonstrated when measuring mental fatigue, aggression, and violence among inner-city residents (Kou & Sullivan, 2001), inner-city children's abilities in self-regulation abilities (Taylor, Kou, & Sullivan, 2002), as well as when using photographs of nature scenes versus control stimuli in experimental assessments of attention and short-term memory (Berman, Jonides, & Kaplan, 2008; Berto, 2005). As such, viewing nature appears to be a simple and easy way to boost waning self-regulation. Indeed, the notion that spending time attending to what is inherently interesting or fascinating may also tie back to the notions of autonomy and intrinsic/identified motivation discussed earlier. One would expect that freely attending to things that one finds intriguing or interesting would be akin to intrinsic motivation. Additional studies examining the potential for other attention-drawing stimuli such as videos to have similar replenishing effects would be highly informative.

Mood

Overall, our most consistent finding was that self-regulation does impact mood, with greater self-regulation typically producing more negatively valenced emotions and higher arousal. In Studies 1.1–1.3, where participants' moods were equated beforehand via a preliminary task, this often manifested as significant post-task differences in mood. In the remaining studies when no preliminary task was used, this was more often evidenced by significant differences in participants' changes in mood from pre- to post-task. In addition to being significant, these interaction effects tended to be large (Cohen's effect size *d* for valence between .35 and 1.17, $M = .68$; effect size *d* for arousal between .28 and .87, $M = .62$). Assuming that mood is determined, at least in part, by the meta-monitoring processes outlined by Carver and Scheier (1990), one would expect such results. Based on their model, negative feelings are the result of less-than-expected progress in self-regulation. Assuming that participants expect to do rather well on the initial task, the challenging nature of many depleting initial tasks should be likely to induce frustration if not sadness.

The mood effects found also tended to mediate depletion effects when found, but only to a small degree. Thus, as stated earlier, depletion cannot be explained as solely a mood effect. Instead, mood likely plays a more nuanced part in a complicated process. For instance, participants' changes in mood from before to after the initial task appear to only have a minor impact on later performance. However, depletion effects may be better explained by a confluence of a lack of self-regulation resources and mood *during* the subsequent task. That is, depleted participants diminish their resources accomplishing the initial task, which causes them to perform more poorly on the subsequent task than expected. In turn, the meta-monitoring loop senses that they are not doing as well as they think they should be on the measurement task, prompting negative feelings. If their mood is sufficiently negative, this may cause them to disengage, and thus illustrate depletion by deprioritizing their performance on the task or even giving up.

Research measuring mood during rather than before and after the various tasks would be needed to test this hypothesis.

Conclusion

In sum, our research indicates that self-regulation depletion, while a common phenomenon, is not inevitable when one has completed an earlier self-regulation task. We also found little evidence that completing various, purportedly low self-regulation tasks such as exercise, magazine reading, and drawing, between the initial and final self-regulation tasks replenished participants' self-regulation resources. However, we argue that the problems in assessing the restorative benefits of such break tasks in laboratory experiments may provide an explanation of this lack of effects. We recommend greater work on self-regulatory depletion effects in ecologically valid contexts, where all tasks – potentially depleting and potentially restorative as well as tasks that might encourage "learned industriousness" – may be more closely interconnected with the longer-term goals and motivation of participants. We also recommend more direct explorations of sleep and of stimuli that may elicit more "spontaneous" and effortless modes of attention, such as those described by Attention Restoration Theory, as possible means of replenishment. Furthermore, we provide clear evidence that mood, both in terms of mood valence and arousal levels, is impacted by self-regulation. While the impact of these effects on depletion is small, we recommend further investigations of the potential mediating role of during-task mood in mediating depletion.

References

Alberts, H. J. E. M., Martijn, C., Greb, J., Merckelbach, H., & de Vries, N. K. (2007). Carrying on or giving in: The role of automatic processes in overcoming ego depletion. *British Journal of Social Psychology, 46,* 383–399. doi: 10.1348/014466606X130111

Arthur, R. (Producer). (2004). *Simply hip* [CD]. Toronto, Ontario, Canada: Compass Productions.

Aslan, A., & Baulm, K.-H. B. (2011). Individual differences in working memory capacity predict retrieval-induced forgetting. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 37,* 264–269. doi: 10.1037/a0021324

Badre, D., & Wagner, A. D. (2004). Selection, integration, and conflict monitoring: Assessing the nature and generality of prefrontal cognitive control mechanisms. *Neuron, 41,* 473–487.

Barch, D. M., Braver, T. S., Nystrom, L. E., Forman, S. D., Noll, D. C., & Cohen, J. D. (1997). Dissociating working memory from task difficulty in human prefrontal cortex. *Neuropsychologia, 35,* 1373–1380. doi: 10.1016/S0028-3932%2897%2900072-9

Baranski, J. V. (2007). Fatigue, sleep loss and confidence in judgment. *Journal of Experimental Psychology: Applied, 13,* 182–196. doi: 10.1037/1076-898X.13.4.182

Barros, R. M., Silver, E. J., & Stein, R. E. K. (2009). School recess and group classroom behavior. *Pediatrics, 123,* 431–436. doi: 10.1542/peds.2007-2825

Baumeister, R. F. (2002). Ego depletion and self-control failure: An energy model of the self's executive function. *Self and Identity, 1,* 129–136. doi: 10.1080/152988602317319302

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*, 1252–1265. doi: 10.1037/0022-3514.74.5.1252

Baumeister, R. F., DeWall, C. N., Ciarocco, N. J. , & Twenge, J. M. (2005). Social exclusion impairs self-regulation. *Attitudes and Social Cognition, 88,* 589–604. doi: 10.1037/0022-3514.88.4.589

Baumeister, R. F., Galliott, M., DeWall, N. C., & Oaten, M. (2006). Self-regulation and personality: How interventions increase regulatory success, and how depletion moderates the effects of traits on behavior. *Journal of Personality, 74*, 1773–1801. doi: 10.1111/j.1467-6494.2006.00428.x

Baumeister, R. F., Vohs, K., D., & Tice, D. M. (2007). The strength model of self-control. *Current Directions in Psychological Science, 16,* 351–355. doi: 10.1111/j.1467-8721.2007.00534.x

Berman, M. G., Jonides, J., & Kaplan, S. (2008). The cognitive benefits of interacting with nature. *Psychological Science, 19*, 1207–1212. doi: 10.1111/j.1467-9280.2008.02225.x

Berto, R. (2005). Exposure to restorative environments helps restore attention capacity. *Journal of Environmental Psychology, 25*, 249–259. doi:10.1016/j.jenvp.2005.07.001

Boreham, C., & Riddoch, C. (2001). The physical activity, fitness and health of children *Journal of Sports Sciences, 19,* 915–929. doi: 10.1080/026404101317108426

Bradley, M. M., & Lang, P. J. (1994). Measuring emotion: The Self-Assessment Manikin and the sematic differential. *Journal of Behavioral Therapy and Experimental Psychiatry, 25,* 49–59. doi:10.1016/j.jrp.2006.11.002

Bradley, M.M., & Lang, P.J. (1999). *Affective norms for English words (ANEW): Instruction manual and affective ratings*. Technical Report C-1, The Center for Research in Psychophysiology, University of Florida.

Brass, M., Derrfuss, J., Forstmann, B., & von Cramon, D. Y. (2005). The role of the inferior frontal junction area in cognitive control. *Trends in Cognitive Sciences, 9,* 314–316. doi: 10.1016/j.tics.2005.05.001

Braver, T. S., Reynolds, J. R., & Donaldson, D. I. (2003). Neural mechanisms of transient and sustained cognitive control during task switching. *Neuron, 39,* 713–726.

Burton, K. D., Lydon, J. E., D'Alessandro, D. U., & Koestner, R. (2006). The differential effects of intrinsic and identified motivation on well-being and performance: Prospective, experimental, and implicit approaches to self-determination theory. *Journal of Personality and Social Psychology, 91,* 750–762. doi: 10.1037/0022-3514.91.4.750

Carter, C. S., & VanVeen, V. (2007). Anterior cingulate cortex and conflict detection: An update of theory and data. *Cognitive, Affective, & Behavioral Neuroscience, 7,* 367–379. doi: 10.3758/CABN.7.4.367

Caruso, E. M., & Shafir, E. (2006). Now that I think about it, I'm in the mood for laughs: Decisions focused on mood. *Journal of Behavioral Decision Making, 19,* 155-169. doi: 10.1002/bdm.506

Carver, C. S. (2004). Negative affects deriving from the behavioral approach system. *Emotion, 4,* 3–22. doi: 10.1037/1528-3542.4.1.3

Carver, C. S. (2006). Approach, avoidance, and the self-regulation of affect and action. *Motivation and Emotion, 30,* 105–110. doi: 10.1007/s11031-006-9044-7

Carver, C. S., & Scheier, M. F. (1990). Origins and functions of positive and negative affect: A control-process view. *Psychological Review, 97,* 19–35. doi: 10.1037/0033-295X.97.1.19

Clore, G. L., & Huntsigner J. R. (2007). How emotions inform judgment and regulate thought. *Trends in Cognitive Sciences, 11,* 393-399. doi: 10.1016/j.tics.2007.08.005

Colcombe, S., & Kramer, A. F. (2003). Fitness effects on the cognitive function of older adults: A meta-analytic study. *Psychological Science, 14,* 125–130. doi: 10.1111/1467-9280.t01-1-01430

Colcombe, S. J., Kramer, A. F., Erickson, K. I., Scalf, P., McAuley, E., Cohen, N. J., Webb, A., Jerome, G. J., Marquez, D. X., & Elavsky, S. (2004). Cardiovascular fitness, cortical plasticity, and aging. *Proceedings of the National Academy of Sciences, 101,* 3316–3321. doi: 10.1073/pnas.0400266101

Converse, B. A., Lin, S., Keysar, B., & Epley, N. (2008). In the mood to get over yourself: Mood affects theory-of-mind use. *Emotion, 8,* 725–730. doi: 10.1037/a0013283

Converse, P. D., & DeShon, R. P. (2009). A tale of two tasks: Reversing the self-regulatory resource depletion effect. *Journal of Applied Psychology, 94,* 1318–1324. doi: 10.1037/a0014604

Dahlin, E., Bäckman, L., Stigsdotter Neely, A., & Nyberg, L. (2009). Training of the executive component of working memory: Subcortical areas mediate transfer effects. *Restorative Neurology and Neuroscience, 27,* 405–419. doi: 10.3233/RNN-2009-0492

Denson, T. F., von Hippel, W., Kemp, R. I., & Teo, L. S. (2010). Glucose consumption decreases impulsive aggression in response to provocation in aggressive individuals. *Journal of Experimental Social Psychology, 46,* 1023–1028. doi:10.1016/j.jesp.2010.05.023

deWall, C. N., & Baumeister, R. F. (2007). From terror to joy: Automatic tuning to positive affective information following mortality salience. *Psychological Science, 18,* 984–990. doi: 10.1111/j.1467-9280.2007.02013.x

droolpurr (Poster). (2007, July 2). Ocean waves [video]. Retrieved from http://www.youtube.com/watch?v=XWZAz9Qbzos&feature=related

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

Elsbach, K. D., & Hargadon, A. B. (2006). Enhancing creativity through “mindless” work: A framework for workday design. *Organization Science*, *17*, 470–483. doi: 10.1287/orsc.1060.0193

Finkel, E. J., Campbell, W. K., Brunnell, A. B., Dalton, A. N., Scarbeck, S. J., & Chartrand, T. L. (2006). High maintenance interaction: Inefficient social coordination impairs self-regulation. *Journal of Personality and Social Psychology, 91*, 456–475. doi: 10.1037/0022-3514.91.3.456

Fischer, P., Greitemeyer, T., & Frey, D. (2007). Ego depletion and positive illusions: Does the construction of positivity require regulatory resources? *Personality and Social Psychology Bulletin, 33,* 1306–1321. doi: 10.1177/0146167207303025

Fischer, P., Greitemeyer, T., & Frey, D. (2008). Self-regulation and selective exposure: The impact of depleted self-regulation resources on confirmatory information processing? *Journal of Personality and Social Psychology, 94,* 382–395. doi: 10.1037/0022-3514.94.3.382

Fitzsimons, G. M., & Bargh, J. A. (2004). Automatic self-regulation. In R. F. Baumeister & K. D. Vohs (Eds.), *Handbook of self-regulation: Research, theory and applications* (pp. 151 –170). New York: Guilford Press.

Freymuth, D. (Producer). (1996). *Pachelbel* [CD]. Toronto, Ontario, Canada: Compass Productions.

Forgas, J. P. (1998). On being happy and mistaken: Mood effects on the fundamental attribution error. *Journal of Personality and Social Psychology, 75*, 318–331. doi: 10.1037/0022-3514.75.2.318

Gaillot, M. T. (2008). Unlocking the energy dynamics of executive functioning: Linking executive functioning to brain glycogen. *Perspectives in Psychological Science, 3*, 245–263. doi: 10.1111/j.1745-6924.2008.00077.x

Gailliott, M. T., & Baumeister, R. F. (2007). The physiology of willpower: Linking blood glucose to self-control. *Personality and Social Psychology Review, 11,* 303–327. doi: 10.1177/1088868307303030

Gailliott, M. T., Baumeister, R. F., DeWall, C. N., Maner, J. K., Plant, A. E., Tice, D. M., Brewer, L. E., & Schmeichel, B. J. (2007). Self-control relies on glucose as a limited energy source: Willpower is more than just a metaphor. *Journal of Personality and Social Psychology, 92*, 325–336. doi: 10.1037/0022-3514.92.2.325

Gasper, K. (2004). Do you see what I see: Affect and visual information processing. *Cognition and Emotion, 18,* 405–421. doi: 10.1080/02699930341000068

Goldberg, L. S., & Grandey, A. A. (2007). Display rules versus display autonomy: Emotion regulation, emotional exhaustion, and task performance in a call center simulation. *Journal of Occupational Health Psychology, 12*, 301–318. doi: 10.1037/1076-8998.12.3.301

Goodyear, L. J., & Kahn, B. B. (1998). Exercise, glucose transport, and insulin sensitivity. *Annual Review of Medicine, 49*, 235–261.

Gosselin, A., De Koninck, J., & Campbell, K. B. (2005). Total sleep deprivation and novelty processing: Implications for frontal lobe functioning. *Clinical Neurophysiology, 116,* 211–222. doi: 10.1016/j.clinph.2004.07.033

Gratton, G., Coles, M. G. H., & Donchin, E. (1992). Optimizing the use of information: Strategic control of activation of responses. *Journal of Experimental Psychology, 121,* 480–506. doi: 10.1037/0096-3445.121.4.480

Haggar, 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,* 495–525. doi: 10.1037/a0019486

Hedgecock, W. M., Vohs, K.D., & Rao, A. R. (manuscript under review). Neural correlates of self-regulatory resource depletion. *Journal of Consumer Psychology*.

Hodes, R., Cook, E. W. III, & Lang, P. J. (1985). Individual differences in autonomic response: Conditioned association or conditioned fear? *Psychophysiology, 22,* 545-560. doi: 10.1111/j.1469-8986.1985.tb01649.x

Hoversten, S., & Koutstaal, W. (2011). Unpublished data.

Ignarro, L. J., Balestrieri, M. L., & Napoli, C. (2007). Nutrition, physical activity, and cardiovascular disease: An update. *Cardiovascular Research, 73*, 326–340. doi: 10.1016/j.cardiores.2006.06.030

Ilies, R., Dimotakis, N., & Watson, D. (2010). Mood, blood pressure, and heart rate at work: An experience-sampling study. *Journal of Occupational Health Psychology, 15*, 120–130. doi: 10.1037/a0018350

Inzlicht, M., & Gutsell, J. N. (2007). Running on empty: Neural signals for self-control failure. *Psychological Science, 18,* 933–937. doi: 10.1111/j.1467-9280.2007.02004.x

Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. *Proceedings of the National Academy of Sciences, 105,* 6829–6833. doi: 10.1073/pnas.0801268105

Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Shah, P. (2011). Short- and long-term benefits of cognitive training. *Proceedings of the National Academy of Sciences, 108,* 10081–10086. doi: 10.1073/pnas.1103228108

Jaeggi, S. M., Studer-Luethi, B., Buschkuehl, M., Su, Y.-F., Jonides, J., & Perrig, W. J. (2010). The relationship between n-back performance and matrix reasoning – Implications for training and transfer. *Intelligence, 38,* 625–635. doi: 10.1016/j.intell.2010.09.001

Kahneman, D. (2003). A perspective on judgment and choice: Mapping bounded rationality. *American Psychologist, 58*, 697–720. doi: 10.1037/0003-066X.58.9.697

Kaplan, S. (1995). The restorative benefits of nature: Towards an integrative framework. *Journal of Environmental Psychology, 15,* 169–182. doi: 10.1016/0272-4944%2895%2990001-2

Kasof, J., Chen, C., Himsel, A., & Greenberger, E. (2007). Values and creativity. *Creativity Research Journal, 19,* 105–122.

Kennedy, M. R. T. & Coelho, C. (2005). Self-regulation after traumatic brain injury: A framework for intervention of memory and problem solving. *Seminars in Speech and Language, 26*, 242–255. doi: 10.1055/s-2005-922103

Killgore, W. D. S., Balkin, T. J., & Wesensten, N. J. (2006). Impaired decision-making following 49 h of sleep deprivation. *Journal of Sleep Research, 15,* 7–13. doi: 10.1111/j.1365-2869.2006.00487.x

Koole, S. L., Smeets, K., van Knippenberg, A., & Dijksterhuis, A. (2009). The cessation of rumination through self-affirmation. *Personality Processes and Individual Differences, 77,* 111–125. doi: 10.1037/0022-3514.77.1.111

Kuo, F. E., & Sullivan, W. C. (2001). Aggression and violence in the inner city: Effects of environment via mental fatigue. *Environment and Behavior, 33,* 543–571. doi: 10.1177/00139160121973124

Lang, P. J. (1980). Behavioral treatment and bio-behavioral assessment: Computer applications. In J. B. Sidowski, J. H. Johnson, & T. A. Williams (Eds.), *Technology in mental health care delivery systems* (pp. 119–l 37). Norwood, NJ: Ablex.

Liebert, O. (Producer). (2009). *Spanish sun* [CD]. Toronto, Ontario, Canada: Compass Productions.

Lungu, O. V., Liu, T., Waechter, T., Willingham, D. T., & Ashe, J. (2007). Strategic modulation of cognitive control. *Journal of Cognitive Neuroscience, 19,* 1302–1315.

MacDonald, A. W. III, Cohen, J. D., Stenger, V. A., & Carter, C. S. (2000). Dissociating the role of the dorsolateral prefrontal cortex and anterior cingulate cortex in cognitive control. *Science, 288,* 1835–1838.

Manly, T., Robertson, I. H., Galloway, M., & Hawkins, K. (1999). The absent mind: Further investigations of sustained attention to response. *Neuropsychologia, 37,* 661–670. doi:10.1016/S0028-3932(98)00127-4

Mansouri, F. A., Buckley, M. J., & Tanaka, K. (2007). Mnemonic function of the dorsolateral prefrontal cortex in conflict-induced behavioral adjustment. *Science, 318,* 987–990. doi: 10.1126/science.1146384

Martijn, C., Alberts, H. J. E. M., Merckelbach, H. Haermans, R., Huijts, A., & de Vries, N. K. (2007). Overcoming ego depletion: The influence of exemplar priming on self-control performance. *European Journal of Social Psychology, 37,* 231–238. doi: 10.1002/ejsp.350

Martin, E. A., & Kerns, J. G. (2010). The influence of positive mood on different aspects of cognitive control. *Cognition and Emotion, 25,* 265–279. soi: 10.1080/02699931.2010.491652

Metcalfe, J., & Mischel, W. (1999). A hot/cool system analysis of delay gratification: Dynamics of willpower. *Psychological Review, 106,* 3–19.

Miller, H. C., Pattison, K. F., deWall, N., Rayburn-Reeves, R., & Zentall, T. R. (2010). Self-control without a "self"? Common self-control processes in humans and dogs. *Psychological Science, 21,* 534-538. doi: 10.1177/0956797610364968

Mischel, W. & Ayduk, O. (2004). Willpower in a cognitive-affective processing system: The dynamics of delay of gratification. In R. F. Baumeister & K. D. Vohs (Eds.), *Handbook of self-regulation: Research, theory and applications* (pp. 99–129). New York: Guilford Press.

Moller, A. C., Deci, E. L., & Ryan, R. M. (2006). Choice and ego-depletion: The moderating role of autonomy. *Personality and Social Psychology Bulletin, 32,* 1024–1036. doi: 10.1177/0146167206288008

Moors, A., & De Houwer, J. (2006). Automaticity: A theoretical and conceptual analysis. *Psychological Bulletin, 132,* 297–326. doi: 10.1037/0033-2909.132.2.297

Muraven, M. (2008a). Autonomous self-control is less depleting. *Journal of Research in Personality, 42,* 763–770. doi: 10.1016/j.jrp.2007.08.002

Muraven, M. (2008b). Prejudice as self-control failure. *Journal of Applied Social Psychology, 38,* 314–333.

Muraven, M., & Baumeister, R. F. (2000). Self-regulation and depletion of limited resources: Does self-control resemble a muscle? *Psychological Bulletin, 126,* 247–259. doi: 10.1037/0033-2909.126.2.247

Muraven, M., Baumeister, R. F., & Tice, D. M. (1999). Longitudinal improvement of self-regulation through practice: Building self-control through repeated exercise. *Journal of Social Psychology, 139,* 446–457. doi: 10.1080/00224549909598404

Muraven, M., Rosman, H., & Gagné, M. (2007). Lack of autonomy and self-control: Performance contingent rewards lead to greater depletion. *Motivation and Emotion, 31,* 322–330. doi: 10.1007/s11031-007-9073-x

Muraven, M., Gagné, M., & Rosman, H. (2008). Helpful self-control: Autonomy support, vitality, and depletion. *Journal of Experimental Social Psychology, 44,* 573–585. doi: 10.1016/j.jesp.2007.10.008

Muraven, M., Shmueli, D., & Burkley, E. (2006). Conserving self-control strength. *Journal of Personality and Social Psychology, 91*, 524–537. doi: 10.1037/0022-3514.91.3.524

Muraven, M., & Slessareva, E. (2003). Mechanisms of self-control failure: Motivation and limited resources. *Personality and Social Psychology Bulletin, 29,* 894–906. doi: 10.1177/0146167203029007008

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*, 774–789. doi: 10.1037/0022-3514.74.3.774

Nilsson, J. P., Söderström, M., Karlsson, A. U., Lekander, M., Åkerstedt, T., Erixon Lindroth, N., & Axelsson, J. (2005). Less effective executive functioning after one night's sleep deprivation. *Journal of Sleep Research, 14*, 1–6. doi: 10.1111/j.1365-2869.2005.00442.x

Oaten, M., Williams, K. D., Jones, A., & Zadro, L. (2008). The effects of ostracism on self-regulation in the socially anxious. *Journal of Social and Clinical Psychology, 27,* 471–504. doi: 10.1521/jscp.2008.27.5.471

Passamonti, L., Fera, F., Magariello, A., Cerasa, A., Gioia, M. C., Muglia, M., Nicoletti, G., Gallo, O., Provinciali, L., & Quattrone, A. (2006). Monoamine oxidase-A genetic variations influence brain activity associated with inhibitory control: New insight into the neural correlates of impulsivity. *Biological Psychiatry, 59*, 334–340. doi: 10.1016/j.biopsych.2005.07.027

Patell, E. A., Cooper, H., & Robinson, J. C. (2008). The effects of choice on intrinsic motivation and related outcomes: A meta-analysis of research findings. *Psychological Bulletin, 134,* 270–300. doi: 10.1037/0033-2909.134.2.270

Purmann, S., Badde, S., Luna-Rodriguez, A., & Wendt, M. (2011). Adaptation to frequent conflict in the Eriksen flanker task: An ERP study. *Journal of Psychophysiology, 25,* 50–59. doi: 10.1027/0269-8803/a000041

Quirin, M., Kazén, M., & Kuhl, J. (2009). When nonsense sounds happy or helpless: The Implicit Positive and Negative Affect Test (IPANAT). *Personality Processes and Individual Differences, 97,* 500–516. doi: 10.1037/a0016063

Richeson, J. A., Baird, A. A., Gordon, H. L., Heatherton, T. F., Wyland, C. L., Trawalter, S., & Shelton, J. N. (2003). An fMRI investigation of the impact of interracial contact on executive function. *Nature Neuroscience, 6,* 1323–1328. doi: 10.1038/nn1156

Robertson, I. H., Manly, T., Andrade, J., Baddeley, B. T., & Yiend, J. (1997). “Opps”: Performance correlates of everyday attentional failures in traumatic brain injured and normal subjects. *Neuropsychologia, 35,* 747–758. doi: 10.1016/S0028-3932%2897%2900015-8

Rosenthal, R. (1979). The “file drawer problem” and tolerance for null results. *Psychological Bulletin, 86,* 638–641. doi: 10.1037/0033-2909.86.3.638

Rowe, G., Hirsh, J. B., & Anderson, A. K. (2006). Positive affect increases the breadth of attentional selection. *Proceedings of the National Acadmeny of Sciences of the United States of America, 104,* 383–388. doi: 10.1073/pnas.0605198104

Ryan, R. M. (1982). Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory. *Journal of Personality and Social Psychology, 43,* 450-461. doi: 10.1037/0022-3514.43.3.450

Ryan, R. M., & Deci, E. L. (2000a). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, *25,* 54–67. doi:10.1006/ceps.1999.1020

Ryan, R. M., & Deci, E. L. (2000b). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist, 55,* 68–78. doi: 10.1037110003-066X.55.1.68

Schmeichel, B. J. (2007). Attentional control, memory updating, and emotional regulation temporarily reduce the capacity for executive control. *Journal of Experimental Psychology: General*, *136,* 241–255. doi: 10.1037/0096-3445.136.2.241

Schmeichel, B. J., & Baumeister, R. F. (2004). Self-regulatory strength. In R. F. Baumeister & K. D. Vohs (Eds.), *Handbook of self-regulation: Research, theory and applications* (pp. 84–98). New York: Guilford Press.

Schmeichel, B. J., & Vohs, K. D. (2009). Self-affirmation and self-control: Affirming core values counteracts ego depletion. *Journal of Personality and Social Psychology*, *96*, 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,* 33–46. doi: 10.1037/0022-3514.85.1.33

Schmeichel, B. J., Vohs, K. D., & Duke, S. C. (in press). Self-control at high and low levels of mental construal. *Social Psychological and Personality Science*. doi: 10.1177/1948550610385955

Schneider, W., & Chein, J. M. (2003). Controlled and automatic processing: Behavior, theory, and biological mechanisms. *Cognitive Science*, *27*, 525–559. doi: 10.1016/S0364-0213%2803%2900011-9

Schwarz, J. C., & Pollack, P. R. (1977). Affect and delay of gratification. *Journal of Research in Personality, 11,* 147–164. doi: 10.1016/0092-6566%2877%2990013-7

Schwarz, N. (1990). Feelings as information: Informational and motivational functions of affective states. In R. M. Sorrentino & E. T. Higgins (Eds), *Handbook of motivation and cognition: Foundations of social behavior* (pp. 527-561). New York: Guilford Press.

Schwarz, N. (2001). Feelings as information: Implications for affective influences on information processing. In L. L. Martin & G. L. Clore (Eds), *Theories of mood and cognition: A user's handbook* (pp. 159-176). Mahwah, New Jersey: Lawrence Erlbaum and Associates, Inc.

Shalev, L., Tsal, Y., & Mevorach, C. (2007). Computerized Progressive Attentional Training (CPAT) program: Effective direct intervention for children with ADHD. *Child Neuropsychology, 13,* 382–388. doi: 10.1080/09297040600770787

Sheppes, G., & Meiran, N. (2007). Better late than never? On the dynamics of online regulation of sadness using distraction and cognitive reappraisal. *Personality and Social Psychology Bulletin, 33*, 1518–1532. DOI: 10.1177/0146167207305537

Simmonds, D. J., Pekar, J. J., & Mostofsky, S. H. (2008). Meta-analysis of Go/No-go tasks demonstrating that fMRI activation associated with response inhibition is task-dependent. *Neuropsychologia, 46*, 224–232. doi: 10.1016/j.neuropsychologia.2007.07.015

Stewart, C. C., Wright, R. A., Hui, S.-K. A., & Simmons, A. (2009). Outcome expectancy as a moderator of mental fatigue influence on cardiovascular response. *Psychophysiology, 46*, 1141–1149. doi: 10.1111/j.1469-8986.2009.00862.x

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

Taylor, A. F., Kuo, F. E., & Sullivan, W. C. (2002). Views of nature and self-discipline: Evidence from inner-city children. *Journal of Environmental Psychology*, *22,* 49–63. doi:10.1006/jevp.2001.0241

Taylor, A. H., Cable, N. T., Faulkner, G., Hillsdon, M., Narici, M., & van der Bij, A. K. (2004). Physical activity and older adults: a review of health benefits and the effectiveness of interventions. *Journal of Sports Sciences*, *22,* 703–725. doi: 10.1080/02640410410001712421

Tennessen, C. M., & Cimprich, B. (1995). Views to nature: Effects on attention. *Journal of Environmental Psychology, 15,* 77–85. doi: 10.1016/0272-4944%2895%2990016-0

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,* 379–384. doi:10.1016/j.jesp.2006.05.007

Tice, D. M., Bratslavsky, E., & Baumeister, R. F. (2001). Emotional distress regulation takes precedence over impulse control: If you feel bad, do it! *Journal of Personality and Social Psychology, 80,* 53–67. doi: 10.1037//0022-3514.80.1.53

Turner, M. L., & Engle, R. W. (1989). Is working memory task dependent? *Journal of Memory and Language, 28,* 127–154. doi: 10.1016/0749-596X%2889%2990040-5

Tyler, J. M., & Burns, K C. (2008). After depletion: The replenishment of the self's regulatory resources. *Self and Identity, 7,* 305–321. doi: 10.1080/15298860701799997

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

Vohs, K. D., & Baumeister, R. F. (Eds.). (2004a). *Handbook of self-regulation: Research, theory and applications* (pp. 1–9). New York: Guilford Press.

Vohs, K. D., & Baumeister, R. F. (2004b). Understanding self-regulation. In R. F. Baumeister & K. D. Vohs (Eds.), *Handbook of self-regulation: Research, theory* *and applications* (pp. 1–9). New York: Guilford Press.

Vohs, K. D., Baumeister, R. F., & Ciarocco, N. J. (2005). Self-regulation and self-presentation: Regulatory resource depletion impairs impression management and effortful self-presentation depletes regulatory resources. *Journal of Personality and Social Psychology, 88*, 632–657. doi: 10.1037/0022-3514.88.4.632

Vohs, K. D., Baumeister, R. F., Schmeichel, B. J., Twenge, J. M., Nelson, N. M., & Tice, D. M. (2008). Making choices impairs subsequent self-control: A limited resource account of decision-making, self-regulation, and active initiative. *Journal of Personality and Social Psychology, 94,* 883–898. doi: 10.1037/0022-3514.94.5.883

Vohs, K. D., & Ciarocco, N. J. (2004). Interpersonal functioning requires self-regulation. In R. F. Baumeister & K. D. Vohs (Eds.), *Handbook of self-regulation: Research, theory and applications* (pp. 392–407). New York: Guilford Press.

Vohs, K. D., & Heatherton, T. F. (2000). Self-regulatory failure: A resource-depletion approach. *Psychological Science, 11*, 249–254. doi: 10.1111/1467-9280.00250

Wagner, A. D., Maril, A., Bjork, R. A., & Schacter, D. L. (2001). Prefrontal contributions to executive control: fMRI evidence for functional distinctions within lateral prefrontal cortex. *NeuroImage, 14,* 1337–1347. doi: 10.1006/nimg.2001.0936

Wendt, M., & Luna-Rodriguez, A. (2009). Conflict frequency affects flanker interference: Role of stimulus-ensemble-specific practice and flanker response contingencies. *Experimental Psychology, 56,* 206–217. doi: 10.1027/1618-

3169.56.3.206

Wyland, C. L., & Forgas, J. P. (2010). Here's looking at you kid: Mood effects on processing of eye gaze as a heuristic cue. *Social Cognition, 28,* 133–144. doi: 10.1521/soco.2010.28.1.133

Zuckerman, M., Porac, J., Lathin, D., Smith, R., & Deci, E. L. (1978). On the importance of self-determination for intrinsically-motivated behavior. *Personality and Social Psychology Bulletin, 4,* 443–446. doi: 10.1177/014616727800400317

Appendix A: Word Ratings Study

Participants

The participants for the word ratings phase were recruited from the same REP pool from which we obtained the participants in Studies 3.1 and 3.2. As in all of our other studies, participants were required to be native English speakers and to have normal or corrected-to-normal vision. Our convenience sample consisted of 38 undergraduate students (13 male, 25 female). The participants ranged in age from 18-28 years (M = 19.95, SD = 2.18) and the sample was racially composed of 65.79% Caucasians, 10.53% African or African Americans, 15.79% Asian or Asian Americans, 2.63% Hispanic/Latino, and 5.26% who reported their race as mixed or "American".

Materials and Procedure

We compiled a list of all of the possible completions of all of the word fragments from the word completion task, based on researcher knowledge and the Psycholinguistic Database Version 2.00 available online (http://www.psy.uwa.edu.au/mrcdatabase/uwa_mrc.htm). In all, the list contained 222 words, ranging from common words such as "have" to obscure words such as "malm". See Table 26 below for the complete list. Participants were shown the words, one at a time, and asked to make their ratings using E-Prime software. For each word, participants were asked to rate how they felt when reading the word on two separate dimensions: mood valence and arousal level. Participants were asked to use the same 9-point SAM ratings scale as was used in our previous studies. If the participant was certain they did not know the meaning of the word, they were asked not to rate it, but to press X instead. The ratings were self-paced.

Table 26. Full list of words and their average valence and arousal ratings.

Word	Valence	Arousal	Word	Valence	Arousal
ALLAY	4.87	2.93	COVE	5.56	3.69
ALLEY	3.89	4.22	CRIME	2.11	6.03
ALLOY	5.42	3.24	CURD	4.71	2.97
ANGEL	7.53	4.87	CURSE	2.55	5.11
ANGER	1.97	6.61	CURTAIN	4.97	2.66
ARIGHT	5.35	3.95	CUTE	7.71	6.11
BALD	3.58	3.47	DEAD	1.39	5.89
BALM	5.08	2.75	DEAF	3.26	3.92
BASIC	5.21	2.79	DEAL	6.11	4.61
BASIL	5.57	3.46	DEAN	5.21	3.95
BASIN	4.85	2.62	DEAR	6.95	4.97
BASIS	5.05	2.68	DEAS	5.00	3.00
BERRY	6.65	4.70	DERRY	5.19	3.31
BORROW	4.82	3.61	DOLLY	5.78	3.53
BOTHER	3.39	5.26	DONKEY	5.50	3.16
BOUL	3.94	2.00	DOVE	6.89	4.05
BOVE	4.73	2.18	DOWN	3.37	3.55
BRIEF	4.84	2.79	FAB	6.24	4.74
BRIGHT	7.39	5.68	FAD	5.49	3.83
BUILT	6.26	4.58	FAG	2.46	4.70
BURSE	4.25	2.81	FAH	5.00	2.38
CALM	6.68	2.13	FAN	5.87	4.21
CARD	5.76	3.70	FAR	3.53	4.29
CARROT	5.42	2.82	FAT	2.47	3.95
CATE	4.60	2.80	FAX	4.62	2.74
CERTAIN	6.92	4.43	FAY	4.79	2.95
CETE	4.40	1.79	FEN	4.84	2.58
CITE	4.17	3.39	FERRY	5.78	3.67
CLACK	5.03	3.31	FIN	5.38	3.78
CLAD	5.11	3.19	FLOWN	6.31	4.78
CLASS	5.00	4.50	FOLLY	4.60	3.44
CLECK	4.50	2.79	FON	4.86	3.14
CLICK	4.87	2.95	FOOD	7.45	6.53
CLOCK	5.11	3.79	FOOL	3.66	4.08
CLUCK	5.47	3.37	FOOT	5.03	2.79
COLLY	5.61	3.72	FOTHER	5.50	2.83
CORD	2.62	4.89	FOUL	5.05	2.62
COTE	2.88	4.71	FRIGHT	6.55	3.00

Word	Valence	Arousal	Word	Valence	Arousal
FROWN	2.21	3.63	HUE	5.60	3.66
FUN	8.24	7.47	HUG	8.18	6.89
GARROT	4.60	2.27	HUM	6.22	3.32
GERRY	4.84	3.26	HUMOR	8.13	6.68
GLAD	7.66	5.50	HUN	6.38	4.31
GLASS	5.16	3.21	HUP	5.08	3.00
GOLD	6.92	5.79	HUT	4.76	3.03
GOLF	5.79	3.61	JERRY	5.00	3.25
GOLL	5.36	2.79	JOB	5.55	6.05
GOLLY	6.19	4.16	JOD	4.79	3.07
GOWN	6.47	5.00	JOE	5.75	3.92
GRIEF	1.92	4.84	JOG	6.45	5.32
GRIME	2.78	3.72	JOLLY	7.89	5.97
GUILT	1.95	5.79	JOT	5.09	3.13
HADE	4.29	2.94	JOUL	4.53	2.53
HAKE	4.57	2.62	JOVE	5.44	4.50
HALE	5.07	3.59	JOY	8.16	6.55
HALM	4.50	2.25	KERRY	5.15	2.35
HAME	4.31	2.69	KNIT	5.53	2.92
HARD	3.66	5.11	KNOT	4.34	3.58
HARE	5.39	3.31	KNUT	4.90	3.00
HARK	5.32	3.48	LOLLY	5.97	4.47
HARL	4.62	3.00	LOVE	8.39	7.74
HARM	1.73	6.19	LOWN	5.14	2.57
HARP	6.35	3.65	MAB	5.36	3.45
HART	5.64	3.48	MAC	5.50	3.97
HATE	1.68	6.26	MAD	2.11	6.37
HAVE	6.03	3.47	MAG	5.42	3.67
HAZE	3.92	4.39	MALM	5.36	2.21
HOLE	4.29	4.47	MAM	5.28	2.76
HOLLY	5.88	3.76	MAN	6.18	4.71
HOME	7.58	5.87	MAP	5.66	3.50
HONE	5.21	3.10	MAR	4.23	3.62
HOPE	8.00	6.24	MAS	5.56	3.11
HORE	4.26	4.37	MAT	5.00	2.71
HOSE	5.00	3.13	MAW	4.35	3.35
HOTE	5.10	3.40	MAX	6.11	4.58
HOVE	4.80	2.40	MAY	7.05	5.84
HUB	5.54	3.80	MERRY	7.92	6.05
HUD	5.44	2.38	MOLLY	5.71	3.61

Word	Valence	Arousal	Word	Valence	Arousal
MONKEY	7.03	4.89	SLANT	4.76	3.29
MORROW	4.65	2.90	SOD	4.65	1.97
MOTHER	7.89	5.87	SORROW	1.86	4.89
MOUL	4.69	3.46	SOUL	6.97	5.50
MOVE	5.00	4.95	SOWN	5.31	2.76
MOWN	4.71	2.88	SPILL	3.29	4.79
NOTHER	4.24	2.88	STILL	5.03	2.39
NURSE	6.21	4.58	STOOD	4.82	2.89
PALM	5.79	3.58	STOOK	5.07	2.00
PARROT	6.34	4.55	STOOL	4.76	2.82
PEACE	7.97	4.08	STOOP	4.70	3.03
PERRY	5.40	3.48	SWILL	4.89	2.39
PLACE	5.45	3.13	TACK	4.06	3.72
PLAID	5.89	3.17	TALK	6.82	5.29
PLAIN	4.55	2.32	TANK	4.55	5.03
PLAIT	4.87	2.87	TASK	4.95	4.55
PLANT	6.08	3.45	TERRY	5.88	3.77
POLLY	5.28	3.28	TOTHER	4.81	2.75
POTHER	4.64	3.00	TOWN	5.68	3.76
PRIME	6.42	4.58	TUMOR	1.61	5.92
PURSE	5.95	3.71	VERRY	5.59	3.94
QUILT	5.79	3.50	WIG	4.71	3.50
ROTHER	4.75	3.33	WIN	8.26	7.53
ROVE	5.23	2.77	WIS	4.94	4.00
RUMOR	2.89	6.08	WIT	6.58	4.97
SAD	1.68	4.84	WOVE	5.14	3.46
SCARED	1.97	7.13	WRIGHT	5.26	3.78
SCORED	6.87	6.42			
SEAH	4.38	2.23			
SEAL	5.97	3.39			
SEAM	5.11	2.19			
SEAN	5.47	3.87			
SEAR	4.36	3.70			
SEAS	5.76	4.18			
SEAT	5.34	2.87			
SERRY	5.08	2.46			
SHILL	4.54	3.00			
SID	4.92	3.25			
SKILL	6.79	5.34			