The Limits of Limited Self-Control

Investigating the Boundaries and Conditions of Ego Depletion

Inauguraldissertation

zur Erlangung des Akademischen Grades

eines Dr. phil.,

vorgelegt dem Fachbereich 02 – Sozialwissenschaften, Medien und Sport

der Johannes Gutenberg-Universität

Mainz

von

Mario Wenzel

aus Berlin

Mainz

2016

Table of Contents

List of Tables

List of Figures

Abstract

Self-control is the individual's ability to alter cognitions, emotions, and behaviors by shielding from short-term distractions in order to obtain longer-term goals. Failures of self-control have been linked to a wide array of individual and societal problems, including substance abuse and overweight (Baumeister, 2001). To study these failures, self-control researchers focused on a dual-task setup in order to investigate performance decrements in subsequent self-control efforts that can be linked as an underlying mechanism of the wide array of problems mentioned above. Although there is a large body of research on the limitations of self-control when performing consecutive tasks, often termed *ego depletion*, evidence is still ambiguous regarding the existence and size of the ego depletion effect, as well as regarding approaches to explain those observed ego depletion effects.

Based on the ambiguous evidence and the explanatory models, I identified four current topics in ego depletion research which can help to advance research in this field. The first research question aims at comparing current models of ego depletion in order to gain a deeper understanding of which models can explain which facet of the phenomena ego depletion. The second and third current topic investigates the internal and external validity of the ego depletion evidence. The fourth topic focuses on alternative ways of capturing limited self-control performance in laboratory research aside the dual-task paradigm.

Chapter 1 provides an overview of the theoretical and empirical background underlying limited self-control performance and summarizes current hot topics in research on ego depletion. Chapter 2 – 5 contain four self-contained studies that each addresses one of the aforementioned topics in ego depletion research.

Chapter 2 examines the first topic, aiming to compare models of ego depletion. By introducing a mild induction of positive affect in between two laboratory tasks, the study was set out to test the diverging predictions of two models, the strength model of self-control

(SMSC) and the conflict-monitoring theory (CMT), with regard to the impact of affect and task characteristics on ego depletion within a dual-task paradigm. The results in Chapter 2 demonstrated that positive affect only benefitted consecutive self-control performance if response conflicts in the two tasks were different. If they were the same, positive affect impaired self-control performance. This effect cannot be explained by the SMSC, which would assume a general performance improvement under positive affect, but by the CMT, which predicts a stability-flexibility tradeoff favoring flexibility under positive affect.

Chapter 3 builds on this research in order to investigate if the flexibility-enhancing effect of positive affect on self-control can be found in everyday life as well, evidencing ecological validity (third current topic). In a 13-day diary study consisting of 297 participants, positive affect was differently associated with self-control success depending on whether the self-control strategies required stability or flexibility, in that individuals with higher positive affect were most successful when following a strategy of distraction that requires flexible disengagement from the current conflicting situation. These results reinforce the idea that positive affect is associated with both cognitive flexibility and distractibility, which may help people distract them from tempting desires, and provides evidence that the effects found in laboratory research can be found in daily life as well, demonstrating the ecological validity of the flexibility-enhancing effect of positive affect.

Although, the detrimental effect of positive affect on similar tasks that require stability and not flexibility can be explained by the CMT, the explanation only holds true if the ego depletion exists. However, one depletion condition in Chapter 2 failed to demonstrate performance differences in comparison to the control condition, which can be connected to recent research suggesting that the ego depletion effect size found in prior research (Hagger, Wood, Stiff, and Chatzisarantis, 2010) is not robust (Carter, Kofler, Forster, and McCollough, 2015; Hagger et al., 2016), which questions the internal validity of the results. The purpose of

Chapter 4 was, thus, to build on a crossover design that enables high statistical power with only a single study. The findings in Chapter 4 showed that (a) the within-subjects ego depletion effect only had small effect sizes, (b) there was substantial heterogeneity in the results depending on the outcome measure, demonstrating the influence of the researcher's choice of analysis method which may explain the heterogeneous results in prior ego depletion research, and (c) there was significant reactivity for our explicit, standardized ego depletion manipulation check, resulting in worse performance after initial performance.

The results in Chapter 4 also demonstrated a large variability of the ego depletion effect size within subjects, indicating performance decrements in some and gains in other individuals due to consecutive self-control. Thus, Chapter 5 introduced a design that presents an alternative way of capturing limited self-control: Instead of only comparing performance between the conditions, goal strength towards performance in the second task was additionally assessed between the two tasks which enables researchers to compare performance in the second task with the goal participants set themselves. Chapter 5 demonstrated that participants who initially exerted self-control set themselves a stricter instead of a more lenient goal than controls, in that they chose to eat less cookies or wanted to perform better. However, these participants could not follow through with their more ambitious intentions, whereas participants without an initial self-control task could adhere to their self-set goal.

To sum up, the present dissertation shows that ego depletion can be influenced by positive affect, is not as robust as suggest in prior research, and, thus, may be better captured as the inability to follow through with one's own intentions. Chapter 6 discusses the implications of the presented set of studies and discusses prospects for future research. Specifically, Chapter 6 shows how research on ego depletion could (a) benefit from connected lines of work such as the literature on mental fatigue, (b) explain the large variability of the ego depletion effect found in Chapter 4, and (c) provide more evidence of ecological validity of self-control limitations by further using ambulatory assessment as complementary tool to laboratory research.

Chapter 1:

General Introduction

People often believe that they could improve their lives if only they had more of that elusive quality known as self-control. Life would radically change then: People would eat healthy, exercise regularly, stop procrastinating, save money, and avoid drugs. However, as tempting as that bowl of cookies might seem, lack of self-control is the major reason why participants do not follow through with such changes according to the Stress in America 2011 survey (American Psychological Association, 2012), in that 27 percent of the participants in that survey responded that lack of self-control was the most significant barrier to change. However, although many people seem to blame limited self-control for their limited success in changing their thoughts and behaviors, the Stress in America 2011 survey also demonstrate that participant still have hope: 71 percent of the respondents believe that self-control can be learned and that it can be done by improving their ability to make changes and by helping them to feel better about themselves.

Interestingly, these rather optimistic views of improving self-control deviate from the focus on limitations of self-control that has been given in psychological research. Here, *self-control* is often defined as one's ability to adaptively alter dominant responses (Baumeister, 2001) by directing one's thoughts and actions to control emotions, cognitions, and behaviours (Vohs & Baumeister, 2004). Without the ability to exert self-control, people would merely stick to dominant behaviours and pursue "old habits" (Baumeister, Vohs, & Tice, 2007), without any freedom to choose (Baumeister, Masicampo, & DeWall, 2009; Rigoni, Kühn, Gaudino, Sartori, & Brass, 2012). Although research on self-control has been conducted for a long time (e.g., Mischel, Ebbesen, & Raskoff Zeiss, 1972), interest in self-control sparked in recent years. Figure

1 presents the number of articles on self-control published since 1986 and it shows that self-control research was increasingly popular since the late 1990s.

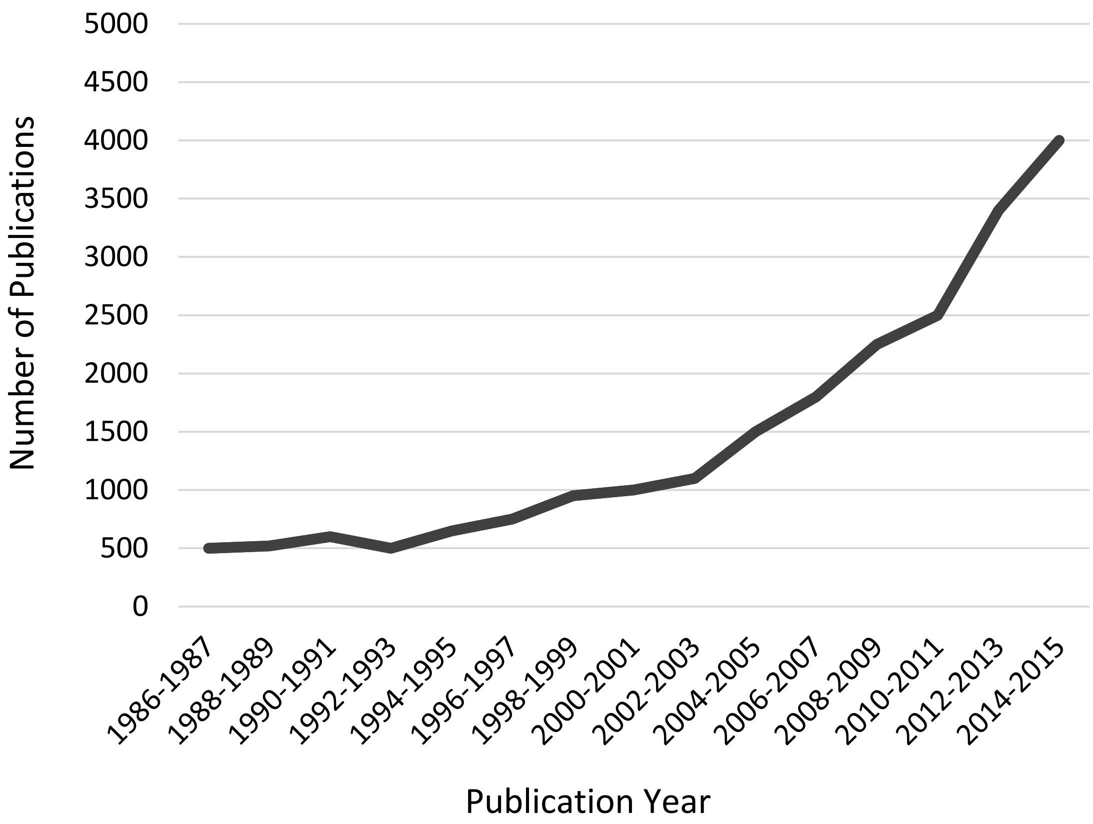

Figure 1. Publication trend in research on self-control.

In line with the Stress in America 2011 survey (American Psychological Association, 2012), research found that measures of self-control correlate with a number of important outcomes such as weight gain or loss (Crescioni et al., 2011), substance abuse (Moffitt et al., 2011), self-esteem (Tangney, Baumeister, & Boone, 2004), or academic achievement (Cleary, 2006). Specific examples of situations or acts where self-control seems to be limited are impulse buying, drug use, gambling, unprotected sex, binge eating, procrastination, and violence (Baumeister & Heatherton, 1996).

The timing of the spike in interest in self-control research coincided with the development of the *dual-task paradigm* by Roy Baumeister and his colleagues (e.g., Baumeister et al., 1998) as a new approach to investigate the state of self-control failure in a controlled laboratory environment. In the dual-task paradigm, participants complete two tasks where the first task differs in self-control demands between an experimental and a control condition: Participants in the experimental condition perform a first task that requires self-control (e.g., thought suppression) while the control group participants perform a first task that does not require self-control. The second task requires self-control in both conditions (e.g., resisting to eat cookies). The performance in the second task can then be compared between participants in the experimental and control condition. A plethora of studies have demonstrated that performance after initial self-control is reduced compared to the control condition, termed *ego depletion* effect (Baumeister et al., 1998), as evidenced by a medium-to-large pooled effect size of d = .62 across various spheres of self-control in a meta-analysis of 83 studies by Hagger and colleagues (2010).

This dissertation builds on and advances this dual-task paradigm to deepen the understanding of the limits of self-control or ego depletion and its influencing factors. In the present chapter, I outline the theoretical background, provide an overview of the relevant literature, summarize prior research findings, introduce the models explaining the ego depletion effect, and point to current topics in the research on limited self-control. Building on these current topics, Chapter 2 to 5 present four empirical studies examining open questions regarding alternative explanations of the ego depletion effect (Chapter 2), the external validity of the moderating influence of positive affect on self-control performance in daily life (Chapter 3), the robustness of the ego depletion effect (Chapter 4), and an alternative way of assessing limited self-control in laboratory research (Chapter 5). In the final Chapter 6, I will summarize and integrate the main findings, discuss the limitations and implications for the study of ego depletion based on the presented four studies, and provide prospects for future research.

1.1 Models Explaining Ego Depletion

1.1.1 Strength Model of Self-Control

To explain the observed ego depletion effects in the dual-task paradigm, Baumeister and his colleagues formulated the strength model of self-control (SMSC; Baumeister, Vohs, & Tice, 2007) which is based on five main assumptions (Muraven & Baumeister, 2000). The first assumption is that acts of self-control draw from a self-control resource called self-control strength, which is, second, limited, in that individuals have a finite capacity for self-control. This means that self-control strength can get depleted by consecutive acts of self-control, for example by performing the two tasks in the dual-task paradigm. Third, all acts of self-control require the same self-control strength and, thus, executing self-control in the first (e.g., regulating emotions) should diminish self-control strength available for self-control in any other sphere in the second task (e.g., performing the Stroop task). Fourth, performance in self-control tasks, or more general the success of self-control, depends on the individual's level of self-control strength. Thus, the existence of depletion of the limited resource is (indirectly) inferred from the deteriorating performance in the experimental condition in the second task of the dual-task paradigm compared to the control condition. Fifth, executing self-control task directly depletes self-control strength which diminishes the level of self-control strength available for subsequent acts of self-control efforts and, thus, must be replenished first before the full measure is available again.

These assumptions are often illustrated in terms of a muscle analogy: If an individual uses a muscle for some period of time, this muscle becomes fatigued by exertion, leading to a state of reduced capacity (e.g., Muraven et al., 1998). If this individual, then, uses the same muscle immediately again for another effort, executing that act is more difficult and less effective if the muscle had been at rest. To the best of my knowledge, there is only one study that tested how long such a break between the first and second task in the dual-task paradigm must be until the limited resource is replenished: By manipulating the time of the break (i.e., 1, 3, or 10 minutes), Tyler and Burns (2008) found that performance in the experimental condition in the second task was significantly worse in the 1-minute condition and 3-min condition compared to the control condition but did not differ in the 10-minute condition. Tyler and Burns (2008) further reported in a second experiment that a 3-minute break, where participants were encouraged to actively relax, also diminished the ego depletion effect, suggesting that short breaks should be enough for relatively short acts of self-control in the first task.

Although the evidence for the ego depletion effect is impressive (Hagger et al., 2010), there are a large number of moderators: positive affect (e.g., Tice, Baumeister, Shmueli, & Muraven, 2007; Nealis, van Allen, & Zelenski, 2016), motivation (e.g., Muraven & Slessereva, 2003), autonomy (e.g., Muraven, Gagné, & Rosman, 2008; Moller, Deci, & Ryan, 2006), implementation intention (e.g., Webb & Sheeran, 2003), priming (e.g., Alberts, Martijn, Greb, Merkelbach, & de Vries, 2007; Martijn et al., 2007), distraction (e.g., Alberts, Martijn, Nievelstein, Jansen, & de Vries, 2008), self-affirmation (e.g., Schmeichel & Vohs, 2009), trait self-control (e.g., Imhoff, Schmidt, & Gerstenberg, 2014), lay theories about self-control (e.g., Job, Dweck, & Walton, 2010), or age of participants (Dahm et al., 2011), among others. This rather long list demonstrates the large variability of the ego depletion effect which is supported by the substantial degree of heterogeneity in the effect size across the studies in the meta-analysis by Hagger and colleagues (2010). Moreover, the evidence regarding moderating influences results in the need for auxiliary assumptions of the SMSC in order to keep the model in line with the evidence. Thus, Baumeister and Vohs (2007) published an extension of the SMSC where they added two auxiliary assumptions. First, people do not deplete their self-control strength completely but conserve it for more important self-control efforts. Second, individuals with high motivation can overcome the effect of depletion to some extend by compensating for the reduced ability to self-control in a depleted state.

However, at least three problems arise from the modification. First, the additional assumptions can only explain some moderators such as motivation but not others such as distraction. Second, extending models via auxiliary assumptions comes at the risk of immunizing the theory which increase the difficulty to falsify it. This problem is even further aggravated by the main assumption of the SMSC that the level of the limited resource cannot, up to this date, be directly assessed but only indirectly inferred by the deteriorating performance, which is the very thing the model tries to explain. Third, large evidence regarding the influence of motivation on self-control performance in the dual-task paradigm is hard to reconcile with the assumption of a limited resource in the SMSC. Early evidence of the importance of motivation was reported by Muraven and Slessareva (2003) who found that participants could maintain high levels of self-control if they were offered internal or external incentives. In the same vein, the ego depletion effect disappeared when temptations were reframed as tests of self-control (Magen & Gross, 2007) or when an unrelated reward was given in the break between the two tasks of the dual-task paradigm such as smoking cigarettes (Heckman, Ditre, & Brandon, 2012) or receiving a gift (Tice et al., 2007). By adding motivation to consume or conserve self-control strength as an auxiliary assumption to the SMSC, which then explains why participants in the experimental condition perform worse in the second task, the question arises whether the limited self-control resource is a necessary condition in order to explain the reduced self-control performance in the experimental condition. If individuals have high self-control strength, self-control performance deteriorates because individuals rate the tasks sufficiently important or interesting and, thus, are not willing to apply enough effort. In turn, when they rate it as interesting, they are willing to exert self-control. Thus, it is an open question if the limited resource is more than a metaphor that is necessary to explain the limits of self-control observed in the dual-task paradigm.

1.1.2 Motivational-Attentional Process Model

In recent years, several alternative approaches have been developed to explain ego depletion and its moderating influences without necessarily relying on a limited self-control resource. In the next sections, I will outline the three most important alternative explanations: the motivational-attentional process model (Inzlicht & Schmeichel, 2012), the opportunity cost model (Kurzban, Duckworth, Kable, & Myers, 2013), and the conflict-monitoring theory (Botvinick, Braver, Barch, Carter, & Cohen, 2001; Dewitte, Bruyneel, & Geyskens, 2009).

Inzlicht and Schmeichel (2012) proposed a prominent re-formulation of the original ego depletion model by adopting a process perspective on ego depletion (Inzlicht, Schmeichel, & Macrae, 2014): In their *motivational-attentional process model* (MAPM), they argue that shifts in motivation and attention offer a suitable alternative explanation for ego depletion that does not necessarily require the recourse to a limited resource. The *motivational process* describes changes in motivation that occur after an individual exerts self-control. Since practicing self-control is effortful, motivation is thought to shift from self-control motives or 'have-to'-tasks towards reward motives or 'want-to' tasks. 'Have-to' tasks are seen to be carried out due to a sense of duty, whereas 'want-to' tasks are seen to be carried out since they are inherently meaningful and enjoyable (Inzlicht, Schmeichel, & Macrae, 2014). Thus, after the first, effortful task in the dual-task paradigm, individuals, instead of wanting to control themselves any longer, want something rewarding for their efforts, which is often absent in experimental setups. The shift in motivation towards reward triggers the *attentional process*: People are, then, more attentive for reward cues after initial self-control and less attentive for cues signaling further self-control, leading to worse performance in the second task.

The MAPM has several advantages over the SMSC. First, it accounts for the existing evidence of the ego depletion effect without reference to the problematic resource concept (Navon, 1984). The notion of a limited self-control resource is of little use when explaining why short times of exerting self-control (e.g., one minute: Halali, Bereby-Meyer, & Meiran, 2014), which should not impact the self-control capacity substantially, still lead to observable declines in performance. Second, it can account for phenomena that cannot be explained by the SMSC. For example, research found that tasks that should not specifically require self-control such as memory updating (Schmeichel, 2007) or the expressed willingness to help a stranger in a hypothetical scenario (Gailliot et al., 2007) also lead to observable declines in consecutive task performance in participants that performed a more difficult version of the first task compared to participants with an easy version. Third, since ego depletion is seen as the outcome of a process of shifting priorities, it enables research to investigate the conditions at which ego depletion occur in more detail, leading to a deeper understand that can be used for future interventions of self-control limitations.

A model that is closely connected to the MAPM is the justification-based account (de Witt Huberts, Evers, & de Ridder, 2012; de Witt Huberts, Evers, & de Ridder, 2014; Prinsen, Evers, & de Ridder, 2016), which ties in why and how people switch from ‘have-to’ to ‘want-to’ goals. Here, decreases in subsequent self-control performance are seen as the result of self-licensing: When individuals experience a self-regulation dilemma between immediate impulses (e.g., discontinue a boring, effortful task) and long-term intentions (e.g., getting partial course credit or being an agreeable person), they resolve the conflict by developing and employing justifications that allow violations of the long-term goal they endorse (the ‘have-to’ goal) in order to follow the immediate impulse (the ‘want-to’ goal).

1.1.3 Opportunity Cost Model

In the *opportunity cost model* (OCM; Kurzban et al., 2013), which integrates from multiple lines of research on cognitive control or mental fatigue, self-control is understood as a decision to engage in one course of action over another. For effective self-control, individuals must balance when to engage in a task, when to persevere, and when to disengage. If this balance is not achieved, individuals may give up prematurely or spend too much effort for too little reward. This view is based on the assumption that natural selection would favor adaptations which minimize opportunity costs due to poor decisions regarding whether to engage or to disengage from a task (Kurzban et al., 2013). Thus, such adaptation has inherent disutility which makes self-control aversive (Kool et al., 2010). Consequently, individuals constantly weighing up the costs associated with task engagement and disengagement and the inherent disutility of continuing self-control accumulates the more individuals have already engaged in self-control (Kool & Botvinick, 2014). Taken together, in this view, performance decreases in subsequent self-control efforts are not due to some finite resource being exhausted but due to the change in cost they associate with different options and, thus, in individuals' preferences and priorities.

Comparing to the MAPM, both the OCM and the MAPM share striking similarities. Both models are based on modern theories of fatigue that do not rely on resources or energy but motivation (Hockey, 1997; Hockey & Earle, 2006) and both models explain self-control decrements by shifts in the valuation effort versus reward. A strength of the OCM compared to the MAPM is that it captures a broader range of inputs that influence a decision such as information, opportunity costs, task importance, task framing, and effortfulness which allows for accommodating findings that are more difficult to account for with a resource model (e.g., Clarkson, Hirt, Jua, & Alexander, 2010; Dang, Dewitte, Mao, Xiao, & Shi, 2013; Heckman et al., 2012).

1.1.4 Conflict Monitoring Theory

However, although both the MAPM and the OCM demonstrates that a self-control resource is not necessary in explaining the evidence regarding the ego depletion effect, both still have difficulties in explaining why performing one short task results in performance declines. For instance, one may assume it to be rather unlikely that participants who exert self-control for as short as one minute (Halali et al., 2014) should already experience a decrease in motivation to a degree that hampers subsequent self-control effort. If true, this would render task engagement and, consequently, self-control surprisingly volatile. Moreover, the MAPM draws on evidence from research on mental fatigue such as performing a Stroop task for 90 minutes (Marcora, Staiano, & Manning, 2009) to explain shifts in motivation. It is unclear, however, if this translates to substantially shorter task. Furthermore, there is little direct evidence for changes in motivation or goals that mediate ego depletion (Baumeister, 2014; see 1.2.1 for more information). Given the limited and ambiguous evidence on the role of task motivation in ego depletion, self-control tasks in laboratory research may either be extraordinarily boring and not engaging at all, leading to a prompt reevaluation of the given options, or there may be another mechanism in place that mitigates self-control decrements.

Drawing on the *conflict-monitoring theory* (CMT; Botvinick, Braver, Barch, Carter, & Cohen, 2001; Miller & Cohen, 2001), self-control decrements after a short initial task can be explained by *switch costs*, in that switching from a first task that is too short for individuals to completely adapt to its self-control demands leads to perseveration to outdated self-control cues from the first task when performing the second task. In this view, a self-control situation is characterized by a *response conflict* where competing responses are simultaneously activated, such as the response to name the color or the word in the Stroop task. Furthermore, cognitive control is guided by two distinct putative neural systems, a conflict-monitoring system and a regulatory system. The *conflict-monitoring system* should detect conflicts between intended and actual responses, which are then passed on to the *regulatory system* for further processing. In case of different consecutive self-control demands, the response conflict in the first task could still linger in the regulatory system, leading to difficulties to adapt to the novel response conflict and, thus, to self-control decrements. Thus, ego depletion can be understood as switch costs resulting in perseveration to outdated self-control cues. First evidence for switch costs was reported by Inzlicht and Gutsell (2007) who conducted an electroencephalography study in order to assess event-related potentials during consecutive self-control. As a measure of activity in the conflict-monitoring system, a part of the event-related potential – the error-related negativity – was used, which is evoked when people make errors on reaction time tasks (Kerns, Cohen, MacDonald, Stenger, & Carter, 2004; Heathterton & Wagner, 2011). Inzlicht and Gutsell (2007) found that participants who exerted initial self-control performed worse than participant without initial self-control efforts and that this difference was mediated by a weaker error-related negativity signal. This finding indicates that initial self-control weakens monitoring of subsequent self-control demands by diverting attention away from (novel) self-control cues towards, presumably, rewarding stimuli.

1.2 Current Topics in Research on Ego Depletion

To date, there has been much theoretical and empirical advancement in understanding when and why self-control seems to be limited in laboratory research. In the following paragraphs, I will outline several limitations of prior research and point to open questions that provided the basis for the studies in this dissertation.

1.2.1 Limited Evidence for Alternative Explanations

Since alternative explanations for the ego depletion have only recently been developed, there is little direct evidence regarding motivational and attentional processes underlying ego depletion compared to the large body of evidence for the demonstration of the ego depletion effect itself. Regarding the role of motivation, both the MAPM and the OCM make strong claims about motivational shifts towards reward mediating ego depletion. First evidence for motivational influences on ego depletion was provided by Muraven and Slessareva (2003) who found that exerting initial self-control did not reduce subsequent self-control performance when a monetary incentive was offered to the participants. Moreover, Schmeichel, Harmon-Jones, and Harmon-Jones (2010) reported that engaging in self-control in the first task increased approach-motivated impulses, in that depleted participants gambled more often for low-stakes winnings than controls. Other studies found that rewarding activities such as watching television (Derrick, 2012), smoking (Heckman et al., 2012), or affirming core values (Schmeichel & Vohs, 2009) also prevent ego depletion effects. Moreover, participants after initial self-control efforts appraised their goals as less important and worthwhile as well as of less priority (vanDellen, Shea, Davisson, Koval, & Fitzsimons, 2014) and reported lower levels of commitment to their self-control goals (Walsh, 2014).

Evidence for the CMT is similarly scarce. Beside the first evidence provided by Inzlicht and Gutsell (2007), switch costs have been shown within tasks on a trial-to-trial basis (Kiesel et al., 2010), but also between consecutive tasks under certain conditions: Dewitte, Bruyneel, and Geyskens (2009) only found ego depletion if both tasks in a dual-task experiment were dissimilar and, therefore, required significant switching of responses for completing the second task. When both tasks were similar, no ego depletion effect was found. The role of switch costs was further supported by a recent study by Dang, Dewitte, Mao, Xiao, and Shi

(2013) who found that performance in the second task of two dissimilar self-control tasks was only impaired when the first task was too short to adapt to the self-control demands. If participants performed the first task (a Stroop task) for a longer period of time that allowed for adapting to the self-control demands (reflected in a reduced Stroop effect), no ego depletion effect was found. Finally, evidence for adaptation effects have been found outside laboratory research as well: Instead of increasing the risk of smoking relapse, the number of recently resisted temptations to smoke decreased the chance of smoking relapse in a sample with 11,176 smoking temptation episodes (O'Connell, Schwartz, & Shiffman, 2008). This finding demonstrates that prior instances of resistance to smoking did not invoke but instead protected people from indulging in current smoking temptations.

Although more research is accumulating to suggest that initial self-control does indeed influences individual's appraisal of their goals and motives or that it is associated with switch costs, evidence, however, is still preliminary and, thus, more research is needed in order to gain a deeper understanding of the underlying cognitive and motivational mechanisms of ego depletion.

1.2.2 Limited Internal Validity

Over the past two decades, researchers have collected a large body of evidence for the ego depletion effect: Hagger and colleagues (2010) conducted a meta-analysis of 83 studies and reported a medium-to-large pooled effect size of $d = 0.62$ for the ego depletion effect. Although an impressive amount of evidence has been amassed so far, the notion of ego depletion and its supporting evidence have recently been challenged. Carter and McCollough (2013; 2014) reanalyzed the studies used in the meta-analysis by Hagger and colleagues (2010) and found evidence for small-study effects such as publication bias. By applying two bias-correction methods, they found average effect sizes of $d = 0.25$ and $d = -0.10$, respectively. In addition, Carter and colleagues (2015) conducted a meta-analysis in which they included non-published effect sizes (40% from the total of included effect sizes) and excluded studies with tasks that did not directly relate to self-control (e.g., imagining cheating on a hypothetical romantic partner). They reported a pooled effect size of Hedge's *g* = 0.43, 95% CI [0.34, 0.52]. This pooled effect was reduced to a Hedge's *g* = 0.24, 95% CI [0.13, 0.34] when 25% extra studies were imputed in order to counterbalance the funnel plot. However, although the trim-and-fill procedure is widely used in meta-analytical research, it is also often criticized (e.g., Moreno et al., 2009; Simonsohn, Nelson, & Simmons, 2014). Carter and colleagues (2015), thus, employed two novel bias-correction techniques – which are sparingly used in economic research (e.g., Stanley, 2008) – that reduced the pooled effect size to zero. The two metaregression techniques, the Precision Effect Test (PET) and the Precision Effect Estimation with Standard Error (PEESE), rely on weighted least squares regression in order to taking into account the association between the effect size and the standard error. Since the association should not dependent on the sample size – which is strongly correlated to the standard error –, a significant association is seen as evidence for publication bias and, in turn, can be used to bias-correcting the meta-analytical results. While this assumption is not uncontroversial (Borenstein, Hedges, Higgins, & Rothstein, 2009), there is evidence that this association is not apparent in large studies (Kühberger, Fritz, & Scherndl, 2014).

However, research has shown that both PET and PEESE (as well as the trim-and-fill procedure) perform poorly in case of heterogeneous effects (Moreno et al., 2009; Reed, Florax, & Poot, 2015) which were found in both meta-analyses of the ego depletion effect (Hagger et al. 2010; Carter et al., 2015). Inzlicht, Gervais, and Berkman (2015) simulated 40,000 meta-analyses with the presence of heterogeneous effects in order to compare several bias-correction techniques including PET and PEESE. The authors found that none of the techniques performed constantly adequately and that the bias-corrected estimates did not

converge. Thus, although bias-correction is of use in ego depletion research, it is still of limited use in order to decide whether the ego depletion is different from zero or not.
An alternative approach to facilitate this decision is a large-scale approach with high statistical power. In 2016, the results of a pre-registered multi-lab replication study of the ego-depletion effect including 23 labs with a total of 2141 participants were published (Hagger et al., 2016). Out of the 23 participating labs, three labs found significant effect sizes, with one effect size in the opposite direction (i.e., better performance in the experimental compared to the control condition), converging to a non-significant pooled effect size of Cohen's d =0.04, 95% CI [-0.07, 0.15]. Unsurprisingly, the replication project was criticized by the authors of the original study (Sripada, Kessler, & Jonides, 2014) by providing evidence that the results were moderated by whether English was the first language or not, in that the effect was slightly larger in the predicted direction, d = 0.14, 95% CI [-0.02, 0.30] (Sripada, Kessler, Jonides, 2016). Although this effect is still not significant considering the commonly used threshold of p = .05, the authors conclude that the tasks in ego depletion research vary considerably across studies and may depend on different underlying mechanisms, leading to heterogeneous results. In order to test this assumption, single studies with high power are needed to compare the various tasks and investigate underlying mechanisms.

1.2.3 Limited Ecological Validity

Initially, research on ego depletion was set out to investigate problems associated with low self-control (Muraven et al., 1998). For example, low levels of self-control have long been associated with criminal behavior (Gottfredson & Hirschi, 1990), dieting (Herman & Polivy, 1975), smoking (Russel, 1971), risk of divorce (Kelly & Conley, 1987), aggression and lower (Baumeister, 1997; Funder & Block, 1989; Funder, Block, & Block, 1983), stress (Shoda, Mischel, & Peake, 1990), and obsessive or ruminative thoughts (Martin & Tesser, 1989;

Wegner, Schneider, Carter, & White, 1987). Researcher, thus, argued that failure of self-control is negatively related to success in many aspects of life (Muraven et al., 1998) and that many of the individual and societal difficulties, ranging from unprotected sexual behavior to addiction to school underachievement, involve failures of self-control (Baumeister, Heatherton, & Tice, 1994).

However, it is striking that although research on ego depletion aimed at providing a deeper understanding of the boundaries and conditions of those self-control breakdowns, it rarely investigated the behavior where and when it happened. Instead, research focused on more or less artificial tasks such as the Stroop task or attention control tasks without investigating or discussing whether these tasks and experimental designs are appropriate to examine the broad range of problematic behavior the research was set out to explain. Thus, like the problems associated with the internal validity discussed in the Section 1.2.2, it is unclear how ecologically valid the results found in laboratory research on ego depletion are. In recent years, ambulatory assessment emerged as a necessary complementary strategy to examine ecological validity (Kubiak & Stone, 2012). However, to the best of my knowledge, only a few studies up to date expand on these approaches to emphasize the importance of everyday behavior in self-control (e.g., Berkman, Falk, & Lieberman, 2011; Hofmann, Baumeister, Förster, & Vohs, 2012; Quinn, Pascoe, Wood, & Neal, 2010). An impediment of ambulatory self-control research was that self-report measures of self-control were mainly focusing on trait instead of state self-control which renders it difficult to investigate self-control related processes in daily life. Consequently, it remains an important, open empirical question how the accrued evidence can be transferred to the large body of self-control related problems in everyday life.

1.2.4 Scarcity of Alternative Designs

As outline before, the SMSC has recently been criticized. A particular problematic notion is the fourth assumption in the model that self-control success is tied to the level of individual's self-control strength (Muraven & Baumeister, 2000). Following, self-control performance decrements in the second task are explained by reduced levels of self-control strength. However, in order to empirically test this assumption, research must assess self-control capacity independent of self-control performance in order to show that differences in self-control capacity can explain differences in performance. Up to date, only glucose has been investigated as a candidate for the limited self-control resource: Drawing on evidence from research in diabetes, Gailliot and Baumeister (2007) posit that differences in the ability to metabolize glucose should reflect differences in the in the level of individual's self-control strength. Thus, actual decreases in blood glucose levels after exerting self-control may, then, mirror performance decrements in subsequent self-control efforts (Gailliot & Baumeister, 2007). Initial evidence for this assumption by a series of laboratory studies (Gailliot et al., 2007) was confirmed by some studies (DeWall, Baumeister, Gailliot, & Maner, 2008; Dvorak & Simons, 2009; Gailliot et al., 2007; Gailliot, Peruche, Plant, & Baumeister, 2009) but also challenged by others (Denson, von Hippel, Kemp, & Teo, 2010; Hagger & Chatzisarantis, 2013; Job, Walton, Bernecker, & Dweck, 2013; Lange & Eggert, 2014; Molden et al., 2012; Sanders, Shirk, Burgin, & Martin, 2012). The initial findings of Galliot & Baumeister (2007) may have been difficult to replicate due to, on the one hand, methodological shortcomings (e.g., insufficient fasting periods) and, on the other hand, conceptual problems: Although it seems intriguing that cognitive control processes involved in self-control consume more glucose in the brain since normal cognitive functioning relies on a constant supply of glucose (McCall, 2004), variation in peripheral glucose levels might not be a good proxy brain glucose consumption: Peripheral blood glucose levels in healthy participants are tightly regulated which renders it unlikely that increased brain metabolism considerably influences peripheral glucose levels in normal physiological states (Messier, 2004). Moreover, previous research has shown that cognitive demand in general is only slightly associated with increases brain glucose uptake (Messier, 2004) and there is little evidence for specific effects of cognitive processes on brain glucose uptake (Kurzban et al., 2013).

Taken together, glucose does not seem to be a good proxy for the self-control resource, neither from a theoretical nor from an empirical point of view. Consequently, research could, on the one hand, look for more promising candidates or, on the other hand, use and develop other approaches and designs that do not infer underlying processes from performance only (Converse & DeShon, 2009).

1.3 Model comparison

Table 1 provides an overview of important characteristics of the presented models explaining ego depletion in Chapter 1.1.

Table 1

Overview of important characteristics of the models explaining limited consecutive self-control

Characteristic	SMSC	MAPM	OCM	CMT
Reason for limited short-term consecutive self-control performance	Self-control resource that gets depleted by initial self-control acts	Shift from control to reward motives	Imbalance between effort and reward	Switch costs due to impaired adaptation of self-control demands in the first task
Number of supporting evidence	High	Low	Low	Low
Internal validity of evidence	Medium	Medium	Medium	Medium
External validity of evidence	Low	Medium	Medium	Low
Effect of increasing task duration on consecutive self-control performance	Performance declines	Performance declines	Performance declines	Performance improves
Effect of task similarity on consecutive self-control performance	Null effect	Null effect	Null effect	Performance declines only in dissimilar self-control tasks
Effect of motivation on consecutive self-control performance	Null effect	Performance declines vanish with increasing motivation	Performance declines vanish with increasing motivation	Null effect

Note. SMSC = strength model of self-control; MAPM = motivational-attentional process model; OCM = opportunity costs model; CMT = conflict-monitoring theory.

1.4 Dissertation Outline

This dissertation aims at investigating the boundaries and conditions of ego depletion by comparing competing explanations of the ego depletion effect and by testing the internal and external validity of this effect and associated processes. Thus, this work integrates research from the models outlined in Chapter 1.1 in order to lay out the problems of and chances for research on limited self-control.

The remainder of this dissertation consists of four chapters, which research objectives are graphically represented in Figure 2.

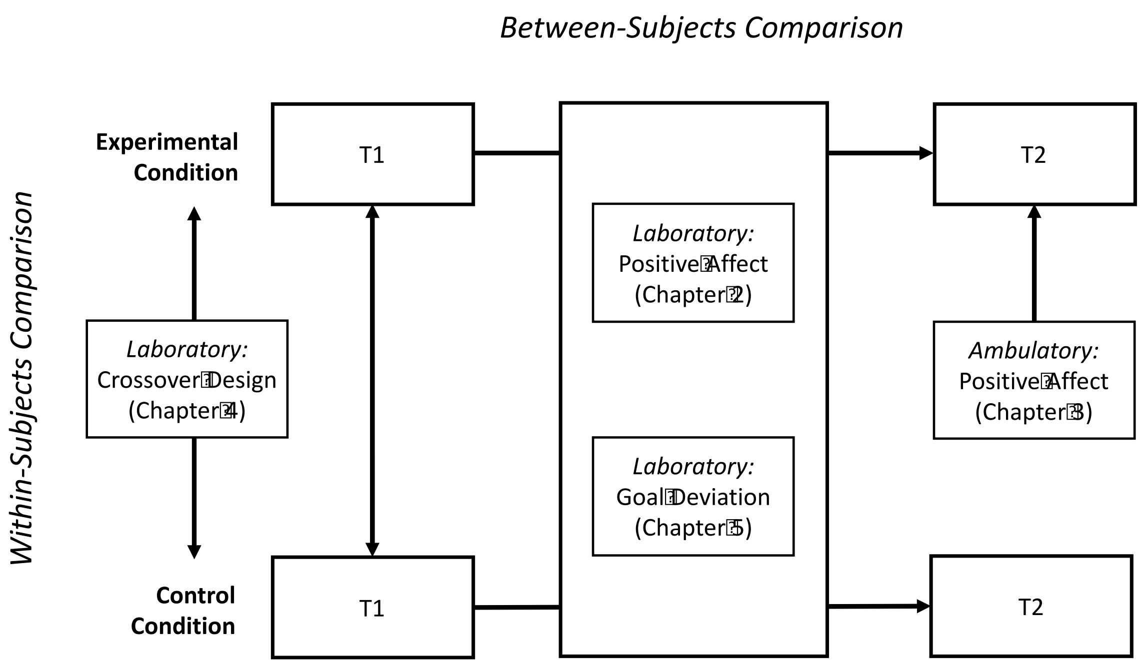

Figure 2. Overview of research objectives and the four studies (Chapter 2 to 5) of the present dissertation.

Chapter 2 focuses on the question if positive affect improves self-control performance after initial exertion of self-control and if this effect can be understood as replenishing a limited self-control resource, as predicted by the SMSC, or as an adaptation process, as predicted by the CMT. Specifically, I employed two dual-task experiments with a mild induction of positive affect and a manipulation of the similarity of the two tasks in the dual-task paradigm.

To tie in with the limited ecological validity discussed in Section 1.2.3, *Chapter 3* investigates the assumptions of the study tested in Chapter 2 in an ecologically more valid context. Specifically, I investigated the flexibility-enhancing effects of positive affect on the self-reported success of self-control strategies followed in daily life in which participants completed a 13-day daily diary study that included measures of positive affect, desire, and habit strength as well as three self-control strategies (i.e., monitoring, distraction, and stimulus control).

Given the ambiguous evidence regarding the ego depletion effect (see Section 1.2.2), I conducted a crossover study in *Chapter 4* that enables high statistical power with only a single study, pursuing three aims: (a) to investigate the robustness of the ego depletion effect between and within subjects; (b) to compare ego depletion across a range of established measures of a self-control task, thus allowing us to judge the influence of the researcher's choice of analysis method; and (c) to estimate the reactivity effects of an explicit manipulation check of ego depletion.

Since evidence regarding glucose as a proxy for self-control strength is rather discouraging, I followed Converse's and DeShon's (2009) recommendation to employ alternative designs that capture limited self-control. The aim of *Chapter 5* was, thus, to test the different assumptions in two dual-task experiments where I operationalized ego depletion as a performance deviation from a self-set goal. Instead of comparing performance in the experimental with the control condition, participants were asked before the second task how much cookies they wanted to eat during the product test of cookies that was introduced as the second task. This assessment, thus, provided a direct measure whether participants are able or not to follow through with their intentions after initial self-control.

Finally, in *Chapter 6* I summarize and integrate the main findings of the empirical studies presented in Chapter 2 to 5 and provide implications and future directions for studies on limited self-control in laboratory and ambulatory research.

Please not that several studies in this dissertation are already published or were under review at the time this dissertation was submitted, which is indicated in more detail in a footnote for each chapter. Consequently, each chapter consists of a specific theoretical background, sometimes leading to overlaps and repeating explanations of terms and approaches due to common theoretical foundations. Moreover, due to the large number of relatively diverse studies in this dissertation (four studies with six experiments), limitations and implications of the results of each study will be discussed separately in the respective chapter and the general discussion in Chapter 6 will include a short discussion of the overall implications for future research on ego depletion.

Chapter 2:

The Effect of Positive Affect on Consecutive Self-Control as a Test of Competing Explanations of Limited Self-Control[1]

2.1 Introduction

Self-control reflects one's ability to adaptively alter dominant responses (Baumeister, 2001). It enables people to avoid routinized, but possibly dysfunctional behaviors, and to flexibly act in daily life. Without the ability to exert self-control, people would merely stick to dominant behaviors and goals and pursue "old habits" (Baumeister, Vohs, & Tice, 2007). The ability to exert self-control is limited though: People cannot control themselves all the time. A common metaphor of the strength model of self-control (SMSC) is that self-control is a "muscle" that gets tired from exertion and has to recover before self-control can be exerted again (Baumeister et al., 2007). Findings consistent with the SMSC have been shown experimentally in the dual-task paradigm (e.g., Vohs & Heatherton, 2000), in which consecutive self-control performance usually deteriorates after initial self-control compared to performance following a first task that did not require self-control. Evidence for this effect on self-control performance has accumulated in many different domains including for the control of emotions (Baumeister, Bratslavsky, Muraven, & Tice, 1998), thoughts (Muraven, Tice, & Baumeister, 1998), food intake (Baumeister et al., 1998), and stereotypes (Vohs, Baumeister, & Ciarocco, 2005). Despite the intuitiveness of the muscle metaphor, it is not known whether the SMSC reflects the actual cognitive processes underlying the limits of self-control. The goal of this article is to examine the cognitive processes underlying self-control

[1] This study was published in 2013 in a peer-reviewed journal. Wenzel, M., Conner, T.S., & Kubiak, T. (2013). Understanding the limits of self-control: Positive affect moderates the impact of task switching on consecutive self-control performance. *European Journal of Social Psychology, 43*, 175-184.

and to compare two competing models by investigating the influence of positive affect on self-control performance.

2.1.1 The Strength Model of Self-Control

According to the SMSC (e.g., Baumeister et al., 2007), decreased performance in two consecutive self-control tasks occurs because every act of self-control, independent of its specific domain, is based on a limited resource, that is, *self-control strength*. When this resource is depleted, the ability to resist an active impulse is impaired. This effect is termed *ego depletion* to describe the detrimental effects of exertion on self-control performance. The SMSC offers explanations for many findings such as the deterioration of self-control performance in dieting (Vohs & Heatherton, 2000) and emotion suppression (Schmeichel, Harmon-Jones, & Harmon-Jones, 2010). Furthermore, associations between self-control decrements and glucose metabolism (e.g., Gailliot, Peruche, Plant, & Baumeister, 2009) and trait heart rate variability (e.g., Segerstrom & Nes, 2007; Geisler & Kubiak, 2009) might indicate a biological basis of limited self-control strength with trait-like features.

However, recent studies question the generalizability of ego-depletion effects. Job, Dweck, and Walton (2010) showed that self-control decrements might not be resource depletion effects per se but rather reflect people's implicit lay conceptions about the limited resource. Dahm and colleagues (2011) found that self-control performance in two consecutive tasks was only impaired for participants under 25 years old but not for older participants. Moreover, Converse and DeShon (2009) demonstrated that those decrements are limited to dual-task designs but do not appear in three-task designs with two initial self-control tasks instead of one. Furthermore, a recent study reported by Wan and Sternthal (2008) highlighted the important role of monitoring one's own behavior in overcoming decrements of self-control performance. Since cognitive processes may explain this evidence, an elaboration on cognitive processes underlying actual self-control performance would be helpful in order to understand the nature of ego depletion effects.

2.1.2 The Conflict Monitoring Theory

As suggested by others (Dewitte, Bruyneel, & Geyskens, 2009; Robinson, Schmeichel, & Inzlicht, 2010), a cognitive control perspective can help to elucidate the cognitive processes of decrements in self-control beyond the SMSC. In the *conflict monitoring theory* (CMT), self-control tasks are characterized by *response conflicts* that occur when alternative and, thus, incompatible responses are simultaneously activated (e.g., Botvinick, Braver, Barch, Carter, & Cohen, 2001). Therefore, the cognitive system must solve the response conflict, which requires cognitive control to adapt to the task demands. A common example for response conflicts is the Stroop task (see Miller & Cohen, 2001, for a detailed description) where participants have to name the color of a word, regardless of its meaning. When the display color is congruent to its meaning (e.g., the word “green” displayed in a green color), the response tendency to the meaning and to the color are the same (e.g., press the green key). Consequently, the task is unambiguous and does not exert self-control. However, when the word “green” is displayed incongruently in “red”, a response conflict arises because the response to the word meaning (green key) differs from the response to the displayed color (red key). Since word reading is a more automated process than color naming, the spontaneous response to name the word must be inhibited in favor of the slower response to the color, resulting in exerted self-control.

The CMT distinguishes between two control systems which are involved in solving response conflicts and adapting to task demands: the conflict monitoring system and the regulatory system (Inzlicht & Gutsell, 2007). The *conflict monitoring system* (Botvinick et al., 2001; Holroyd & Coles, 2002) monitors ongoing behavior in order to detect response conflicts. If a conflict is registered (e.g., the word "green" displayed in a red color), the conflict monitoring system passes the information to the *regulatory system*, which resolves the detected ambiguities by selecting the appropriate response (Kerns, Cohen, MacDonald, Cho, Stenger, & Carter, 2004). Hence, detecting and resolving the conflict are conceptualized as distinct (but interlinked) processes that can be linked to separate neurobiological correlates (conflict monitoring system to anterior cingulate activity, e.g., Botvinick, Nystrom, Fissell, Carter, & Cohen, 1999; regulatory system to prefrontal cortex activity, e.g., Miller & Cohen, 2001; Knight & Stuss, 2002).

2.1.3 Decrements in Self-control: SMSC vs. CMT

Both theories, the SMSC and the CMT, explain decrements in consecutive self-control performance within a dual-task paradigm differently. The SMSC predicts decrements in performance will occur whenever both tasks require the limited self-control strength. Exertion of self-control in one task should impair performance in the next task to the extent that both 'use up' the scarce self-regulatory resources. Thus, performing the same self-control task twice (e.g., Stroop followed by Stroop) or two different tasks (e.g., thought suppression followed by Stroop) will result in a decreased performance in the second task.

However, from the CMT, decrements in self-control performance are not due to a limited capacity of the cognitive system per se, but are instead due to impaired conflict monitoring of the second task. Specifically, when participants perform a second task with a novel conflict (e.g., thought suppression followed by a Stroop task), they have to shift back from using the regulatory system to engaging the conflict monitoring system first before engaging with the new conflict of the second task. Decrements in self-control performance may occur because of decreased activity in the conflict monitoring or in the regulatory system. If the regulatory system is impaired, self-control performance in the second task should be independent from task switching, which would be in line with the SMSC. However, we argue that the conflict monitoring system is impaired due to different response conflicts. So, decrements in consecutive self-control performance should only occur when the response conflict of the second task is different to the conflict of the first task. If participants simply perform the same self-control task twice (Stroop followed by Stroop), they do not have to switch back from regulation to conflict detection and, therefore, should not show impaired self-control performance in the second task as long as the participants can maintain the regulation of the response conflict. Consequently, a cognitive control view and a regulatory strength view hold different predictions about performance effects for performing the same self-control task twice.

Several lines of research support this cognitive control perspective on self-control. First, Inzlicht and Gutsell (2007) showed that participants who suppressed their emotions while watching an emotional film clip in a first task performed worse in a subsequent Stroop task compared to participants who did not have to suppress their emotions. Although this pattern can be interpreted from either the SMSC or the CMT, electroencephalography data suggested that decrements were mediated by weaker error-related negativity signals originating from the anterior cingulate cortex (van Veen & Carter, 2002). Given that the anterior cingulate cortex is modeled as the neural basis of conflict monitoring systems (Kerns et al., 2004), these findings suggest a link between self-control decrements and impaired conflict monitoring.

Second, Schmeichel (2007) demonstrated that decrements in self-control performance were not necessarily restricted to self-control tasks involving modification of dominant responses, but were linked to switching demands within a cognitive task. This evidence suggests that deteriorated performance occurs when cognitive tasks involve multiple operations that require switching from one mindset to another (e.g., tasks requiring division followed by multiplication). Although this evidence is not directly related to the conflict monitoring and regulatory systems per se, it illustrates the importance of switching costs.

Third, as shown by Dewitte and colleagues (2009), performance in a dual-task paradigm is impacted by *task switching*. Specifically, they showed that performance was only impaired in a second self-control task when participants had to solve a different conflict in the first self-control task (e.g., a thought suppression task followed by Stroop task). When the conflict was the same (e.g., Stroop task twice), performance improved in the second task compared to the control condition with no response conflict in the first task. Again, these findings are consistent with the CMT, which predicts that adaptation to the same regulatory task leads to a better and not impaired self-control performance in the second task. The SMSC does not predict this result because both tasks require self-control and, therefore, the use of the limited self-control resource.

Taken together, it is apparent that switch costs – within one task (Schmeichel, 2007) or between two tasks (Dewitte et al., 2009) – play an important role in decrements of self-control performance, suggesting there is merit in a cognitive control perspective. However, the evidence from Dewitte and colleagues (2009) is not necessarily inconsistent with the SMSC. It can be argued that performing the same task twice simply leads to practice effects and that these effects are somewhat stronger than the depletion of the limited resources. Thus, the findings do not rule out the predictions made by the SMSC; adaptation to a given self-control task (i.e., practice effects when performing the same task twice) could be a complementary part of the SMSC. Hence, it is not known to what extent both theories conflict or complement each other: Can the CMT explain decrements in self-control without recourse to an additional self-control resource? To answer this question, we turned to the role of positive affect. Interestingly, each view — the SMSC complemented by practice effects and the CMT with adaptation effects only — make different predictions about the effect of positive affect on consecutive self-control performance.

2.1.4 Positive Affect

Within the SMSC, positive affect is thought to play a role in restoration of the limited self-control resource. This effect was shown by Tice and colleagues (2007) who found that positive affect can overcome decrements in self-control task performance. For example, if participants watched a comedy film sequence after the first self-control task where they were to cancel out all letters "e" in page of text, self-control performance was higher in a difficult and frustrating ball game in the second task, compared to participants who watched a neutral film sequence. Moreover, participants in the positive affect group did not differ from participants who performed a non-self-control first task. If this beneficial effect of positive affect were based on the restoration of a depleted resource, it should be largely independent from task switching as long as both tasks within a dual-task paradigm consume the limited resource.

A cognitive control perspective of positive affect offers a different interpretation, leading to diverging predictions of the effect of positive affect on consecutive self-control performance. Dreisbach and Goschke (2004) view self-control as a balance between maintaining stability and having the flexibility to switch to different cognitive sets. This concept can be linked to the two control systems. Whereas activation of the regulatory system requires stability to solve conflicts with a given cognitive set, activation of the conflict monitoring system requires flexibility to disengage from the previous cognitive set and switch to a new one.

According to Dreisbach and Goschke (2004), positive affect impacts the regulation of this stability-flexibility balance profoundly by favoring flexibility over stability, as positive affect directs attention to novel information. This mechanism could be beneficial by enhancing flexibility when the information is useful for the regulation of the new task, but detrimental by increasing distractibility and lowering stability when the new information is not helpful. To discriminate between the demands of stability and flexibility, they developed a Stroop like cognitive set-switching paradigm where each trial consisted of two characters. Participants were trained to respond to a target stimulus in a pre-specified color (e.g., green characters like a green "N") and to ignore a different-colored distractor (e.g., red characters like a red "A"). Thereafter, they had to switch to one of two new conditions: In the flexibility condition, the new distractor appeared in the color of the former target (e.g., a green "5") while the target was presented in a new color (e.g., a yellow "9"). In this condition, the new information (color) was linked to the target and, consequently, switching costs under positive affect were reduced due to increased flexibility. In the stability condition, the new target shared the color of the former distractor (e.g., a red "O") while the distractor appeared in a new color (e.g., a yellow "X"). Here, the new information was linked to the distractor. Because positive affect should bias participants towards novel information, it should increase switching costs in the stability condition. Consequently, the new information determined whether positive affect was beneficial (increased cognitive flexibility) or harmful (increased distractibility),

Dreisbach and Goschke (2004) interpret this opposite effect as a modulation of the balance between flexibility and stability by positive affect, which can, in our view, be translated to the effects of affect in the dual-task paradigm. Consequently, we argue that the effects of positive affect on impaired performance in a second self-control task should depend on task switching: Participants in a flexibility-favoring condition with different conflicts in both tasks have to switch to solve a new conflict for the second task, which additionally requires flexibility. This mechanism could explain the restoration effects in the four studies reported by Tice et al. (2007) because all of the two consecutive tasks were of dissimilar conflicts (e.g., thought suppression and drinking of an unpleasant beverage). On the contrary, under a stability-favoring condition with the same response conflict in both tasks, participants do not have to switch to a different conflict in the second task. Therefore, flexibility is not required and so positive affect could have even detrimental effects to performance, as it reduces stability.

2.1.6 The Present Research

In the present research, we integrated the results from the SMSC and the CMT to elaborate on the influence of task characteristics and positive affect on consecutive self-control task performance. In Study 1a, we tested the diverging hypotheses that both models make with regard to the effects of positive affect on consecutive self-control performance. Self-control performance was tested using a dual-task paradigm with either different response conflicts (e.g., resisting sweets followed by a Stroop task, called the "flexibility condition") or the same response conflict twice (Stroop task followed by another Stroop task, or the "stability condition") together with a control condition (e.g., watching a movie followed by a Stroop task). For this design, both models predict a 3 (flexibility vs. stability vs. control) by 2 (positive affect vs. neutral affect) interaction because positive affect should not influence the control condition. However, the models differ in the 2 (flexibility vs. stability) by 2 (positive affect vs. neutral affect) single degree of freedom interaction: Whereas the SMSC predicts a main effect of affect due to the replenishment of the limited self-control resources, the CMT predicts an interaction, in that positive affect is beneficial in the stability condition but harmful in the flexibility condition. In Study 1b, we replicated this design and included a mild induction of negative affect to show that effects are unique to positive affect.

2.2 Study 1a

2.2.1 Method

We examined our hypotheses experimentally using a 3 (stability vs. flexibility vs. control) X 2 (positive affect vs. neutral) between-subjects ANOVA design. Participants were randomly assigned to the experimental conditions illustrated in Table 2. The dependent measures were the Stroop effect and the error rates in the second Stroop task.

Table 2

Randomized distribution of participants to experimental conditions in Study 1a / Study 1b.

Affect condition	Task switching condition		
	Stability	Flexibility	Control
Positive Affect	23 / 13	14 / 16, 14	18 / 11
Neutral Affect	22 / 17	13 / 17, 17	18 / 11
Negative Affect	- / 15	- / 16, 17	- / 12

Note. Values indicate the number of participants in each condition. Study 1a did not include a negative affect condition. The two flexibility conditions in Study 1b are thought suppression and attention control.

2.2.1.1 Participants

A total of 120 undergraduates from the University of Greifswald, Germany, participated in the Study 1a for partial fulfillment of course credit. Ten participants were excluded because of technical problems or dropout during the experimentation, one because of defective color vision, and one because of an extremely large Stroop effect in the Stroop test (> 3 *SD* of mean

Stroop effect). This left a total of 108 participants (78 female, age M = 23.53 years, SD = 3.92) for statistical analysis.

2.2.1.2 Procedure

We implemented a computer-based dual task paradigm with an affect induction in between both tasks. The experiment was realized in PHP 5.3 and was browsed with Firefox 3.6 (Mozilla Project, 2010) under Microsoft Windows XP. The screen resolution of the CRT display was 1024 x 768 pixels.

2.2.1.3 Materials

Participants completed two on-screen questionnaires before the experiment began. The first questionnaire contained socio-demographic characteristics; the second questionnaire tested their baseline mood through the short version of the validated German questionnaire Multidimensionaler Befindlichkeitsfragebogen (Multidimensional Mood State Questionnaire, MDBF; Steyer, Schwenkmezger, Notz, & Eid, 1997). The MDBF assesses momentary mood on three dimensions: *valence* (good, pleasant, bad, unpleasant), *energetic arousal* (awake, alert, sleepy, tired), and *tense arousal* (calm, relaxed, nervous, tense) with four unipolar items for each dimension, ranging from 1 (*not at all*) to 5 (*very much*). As shown by Schimmack and Reisenzein (2002), both arousal dimensions cannot be reduced to a combination of valence and activation, as suggested by others (e.g., Yik et al., 1999), resulting in two different arousal scales where only energetic arousal is influenced by a circadian rhythm (Thayer, 1989). Mean scores were computed for each dimension with higher scores representing higher valence (i.e. more positive affect), higher energetic, and higher tense arousal. Cronbach's alphas for the three scales of valence (.80), energetic arousal (.80), and tense arousal (.82) were in line with the previously reported consistency of .73 to .89 of the short version (Steyer et al., 1997).

2.2.1.3.1 First task

Participants in the *stability* condition performed a standard Stroop task on the computer as the first task described in the section "second task". The mean duration of the Stroop task was approximately 5 minutes (~303 seconds; *SD* = 36). Participants in the two other conditions (*flexibility* and *control condition*) were told that they would be watching a short film on the computer screen, which was the same film in both conditions. Close to the computer screen was a plate with cookies on it. In the *flexibility condition*, the experimenter informed the participants before the short film that he or she is leaving the room and that they must not eat the sweets because they have to be used in another experiment (Baumeister et al., 1998). In the *control condition,* there were no sweets in the room and the experimenter only informed the participants that he would leave the room before the film clip was shown. We opted for this procedure instead of permitting the participants to eat the sweets in this condition as such permission might be considered rewarding and, hence, influences affect (cf., Westermann, Spies, Stahl, & Hesse, 1996). This short film presented the daily routine of a common person and lasted five minutes. We pretested the affective content of this neutral film used in the first task and found no effects on the MDBF valence scale.

This first task was followed by a rating of perceived difficulty of the task using a 1 ("not at all difficult") to 5 ("very difficult") scale. Corresponding to the effortful and nature of self-control tasks, the perceived difficulty should be higher in both conflict conditions (Stroop test or resisting sweets while watching a film clip) than in the control condition (watching a film clip). There should be no differences between the stability and flexibility condition.

2.2.1.3.2 Induction of positive affect

Subsequently, the participants watched a film clip with the instruction to "put themselves as well as possible into the situation shown." Participants were randomized to watch a positive comedy film clip or a neutral film clip about a small political campaign of a local politician. We opted for this procedure, as according to the meta-analysis of Westermann and colleagues (1996), the combination of a film clip with an instruction is the most effective way to induce affect. After the film clips, participants completed the MDBF affect measure for a second time.

2.2.1.3.3 Second task

After completing the MDBF, participants in all experimental conditions performed the Stroop task which was programmed using Adobe Flash CS4 (Adobe Systems Inc., San Jose, CA, USA). In the Stroop task, participants were instructed to name the color of a word displayed on the screen as fast as possible by pressing a specific key on the keyboard. Words were either neutral words or color words (red, green, yellow, blue). Color words (e.g., red) could match the color it was displayed in (congruent trials; here red) or not (incongruent trials; e.g., green). On each of the one hundred trials, a fixation cross appeared for 1000 ms, after which the stimulus word appeared, followed by a feedback of the correct answer for 1000 ms. Besides error rates, our primary dependent variable was the Stroop effect ($RT_{incongruent}$ - $RT_{congruent}$ with RTs < 200 ms and > 5 *SD* of mean RT excluded) where higher scores indicate worse inhibition on incongruent trials, reflecting greater decrements in self-control performance.

2.2.2 Results

2.2.2.1 Manipulation Checks

2.2.2.1.1 Perceived difficulty

A one-way analysis of variance showed that the effect of task switching on perceived difficulty was significant, $F(2, 105) = 7.44$, $MSE = 0.609$, $p = .001$, $\eta^2 = .124$. Focused comparisons with the pooled error term of the ANOVA indicated that participants in the control condition ($M = 1.58$, $SD = 0.69$) found the task significantly easier than those in the flexibility ($M = 2.15$, $SD = 0.82$, $t(105) = -2.84$, $p = .005$, $d = -0.724$) and stability conditions ($M = 2.22$, $SD = 0.82$, $t(105) = -3.66$, $p < .001$, $d = -0.819$). Both self-control conditions were rated equally difficult, $t(105) = -0.39$, $p = .697$, $d = -0.095$. Hence, possible differences in self-control task performance in the second task were not mirrored by perceived difficulty of the first task in both conflict conditions. This finding indicates that our manipulation of the self-control tasks worked as intended.

2.2.2.1.2 Affect manipulation

A 3 (task switching) X 2 (affect) ANCOVA on valence after the affect manipulation including baseline valence as a covariate yielded a main effect for affect, $F(1, 101) = 9.56$, $MSE = 0.258$, $p = .003$, $\eta_p^2 = .086$. Participants in the positive affect condition ($M_{adjusted} = 3.87$, $SEM = .069$) were in a better mood after the mood induction than participants in the neutral affect condition ($M_{adjusted} = 3.57$, $SEM = .070$). We did not find a main effect for task switching, $F(2, 101) = 0.85$, $MSE = 0.258$, $p = .429$, $\eta_p^2 = .017$, nor an interaction effect, $F(2, 101) = 0.24$, $MSE = 0.258$, $p = .789$, $\eta_p^2 = .005$. For the arousal scales, neither the energetic nor the tense arousal scale of the MDBF was significantly influenced by the affect manipulation, $F(1, 101) = 1.04$, $MSE = 0.182$, $p = .309$, $\eta_p^2 = .010$ and $F(1, 101) = 1.14$, $MSE = 0.287$, $p = .287$, $\eta_p^2 = .011$, respectively. Taken together, the manipulation checks demonstrate that our induction of positive affect worked as intended, evidencing the medium effect size.

2.2.2.2 Stroop Performance

A 3 (task switching) X 2 (affect) ANOVA did not show any significant main effects of task switching, $F(2, 102) = 0.77$, $MSE = 13054$, $p = .464$, $\eta_p^2 = .015$, or affect, $F(2, 102) = 0.74$, $MSE = 13054$, $p = .392$, $\eta_p^2 = .007$, on the Stroop effect in the second task. However, as predicted by both models, the interaction of these two factors was significant, $F(2, 102) = 5.52$, $MSE = 13054$, $p = .005$, $\eta_p^2 = .098$. This is interaction is shown in Figure 3.

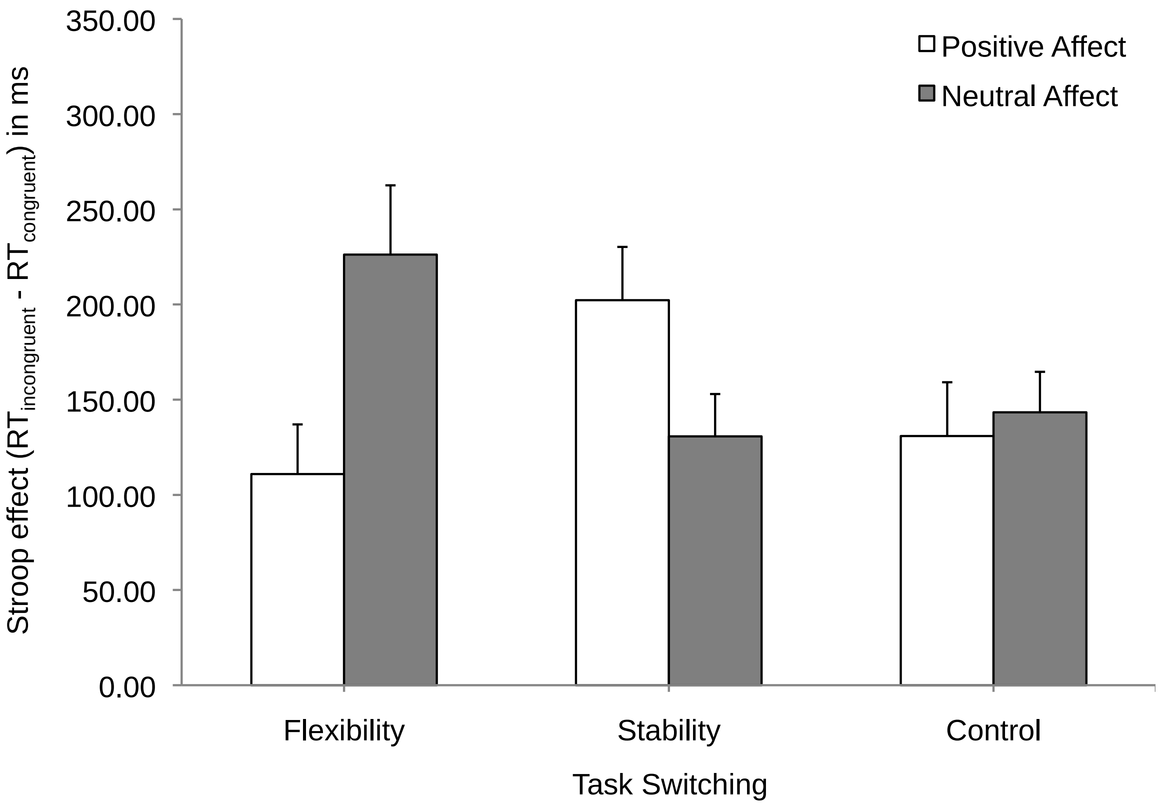

Figure 3. Mean Stroop effect in the second Stroop task varied as a function of task switching and affect (Study 1a). Higher scores indicated poorer performance. Task switching indicates

whether the second Stroop task was preceded by a different self-control task (Flexibility), by the same Stroop task (Stability) or by no self-control task (Control). Error bars reflect standard errors.

To evaluate this two degrees of freedom interaction, we first conducted a contrast analysis for the two competing *a priori* hypotheses (Abelson & Prentice, 1997). The (PA NE) weights for the SMSC were (-1 1) for the flexibility conditions, (-1 1) for the stability conditions, and (0 0) for the control conditions to test the predicted main effect of affect. That contrast test revealed no significant effect, $F(1, 102) = 0.67$, $p = .416$, $d = .279$. Conversely, the (PA NE) contrast weights for the CMT model were (-1 1) for the flexibility conditions, (1 -1) for the stability conditions, and (0 0) for the control conditions to test the predicted interaction, which yielded a significant contrast test, $F(1, 102) = 11.21$, $p = .001$, $d = 1.141$. This finding demonstrates that the data largely reflect predictions from the CMT.

Second, we further investigated the interaction of task switching and affect by evaluating the focused comparisons of affect separately for the three task switching conditions, as shown in Figure 3. In the flexibility condition, participants in the positive affect condition (M = 105.85 ms, SD = 96.10) showed smaller Stroop effects compared to participants in the neutral affect condition (M = 220.73 ms, SD = 136.41), $t(102) = -2.61$, $p = .010$, $d = -0.974$, demonstrating that positive affect benefitted performance when the response conflicts were different across the two tasks. Conversely, in the stability condition, participants in the positive affect condition (M = 202.22 ms, SD = 134.20) showed larger Stroop effects than participants in the neutral affect condition (M = 132.77 ms, SD = 100.00), $t(102) = 2.04$, $p = .044$, $d = 0.587$, demonstrating that positive affect was harmful and not helpful when the response conflicts were the same in both tasks. No differences were found for the control condition (positive affect: M = 130.87 ms, SD = 119.79; neutral affect: M = 143.38 ms, SD = 90.12), $t(102) = -0.33$, $p = .743$, $d = -0.118$.

The mean number of errors was $M = 4.27$ ($SD = 3.72$). We did not find any significant differences for task switching on the error rates in the second Stroop task, $F(2, 102) = 0.51$, $MSE = 13.74$, $p = .602$, $\eta_p^2 = .010$, affect, $F(1, 102) = 1.17$, $MSE = 13.74$, $p = .282$, $\eta_p^2 = .011$, nor the interaction of both, $F(2, 102) = 2.36$ $MSE = 13.75$, $p = .099$, $\eta_p^2 = .044$.

2.2.3 Discussion

Based on Dreisbach and Goschke (2004), we hypothesized that positive affect is linked to flexibility and that this connection may explain the role of affect on consecutive self-control task performance. Our findings were consistent with the study by Dreisbach and Goschke (2004) showing that positive affect impacts the stability-flexibility balance by favoring flexibility over stability. In the flexibility condition, where positive affect should be beneficial, performance in the second Stroop task was better (smaller Stroop effect) after induction of positive affect compared to the neutral condition. By contrast, when stability should be advantageous, as in the stability condition, positive affect hampered the performance compared to the neutral condition, resulting in a larger Stroop effect. This pattern suggests that positive affect influences adaptation to old and new tasks differently by increasing flexibility rather than by replenishing a limited resource.

2.3 Study 1b

The findings from Study 1a showed that positive affect improves adaptation to a situation following a different response conflict, but impairs adaptation following the same response conflict. These findings suggest that positive affect facilitates flexibility and conflict detection rather than replenishes self-control strength. The purpose of Study 1b was (a) to replicate these findings and (b) to examine the effects of negative affect on decrements in self-control task performance.

Due to the large number of conditions, we conducted an online experiment to achieve a sufficiently large sample size without running participants individually in the lab. Although online implementations of the dual-task paradigm have not been used until now, recent studies have demonstrated the feasibility and validity of this approach in self-control research (e.g., Burger, Charness, & Lynham, 2010). Moreover, Reimers and Stewart (2007) demonstrated that an online Stroop task implemented in Flash was equivalent to a conventional lab Stroop task implemented on a computer. To increase comparability with the results of Study 1a, we used the same material (film clips, Stroop task) in the online experiment where possible.

We extended our investigation to negative affect since evidence of the effect of induced negative affect on self-control performance is heterogeneous. Although Tice et al. (2007) did not find substantial effects of negative (compared to neutral) affect, Bruyneel and colleagues (2009) recently found that participants in a negative mood performed worse in a second task compared to those under a neutral condition, presumably because they cannot allocate resources due to active mood regulation. Consequently, negative affect should not interact with task switching because participants would need to allocate resources in both conflict conditions, however negative affect should decrease performance in the second task compared to the control condition.

2.3.1 Method

We used a 4 (stability vs. flexibility 1 vs. flexibility 2 vs. control) X 3 (positive vs. neutral vs. negative affect) between-subjects ANOVA to test our hypotheses, with two different flexibility conditions due to their novelty in online experiments. Participants were randomly assigned to the experimental conditions illustrated in Table 2.

2.3.1.1 Participants

A total of 268 participants started the online experiment, which lasted on average 22.5 minutes (*SD* = 6.5). We recruited the participants by advertising in e-mail newsletters from German universities. Of the starting sample, 176 participants completed the experiment (65.67%). We found no evidence that the dropout rate was systematically influenced by task switching or affect, with $\chi^2(11) = 8.70$, $p = .65$. A remaining total of 176 participants (128 female, 156 students, age *M* = 24.10 years, *SD* = 4.14) were included for statistical analysis.

2.3.1.2 Procedure

The procedure of Study 1b was similar to that of Study 1a, in that participants completed a first task followed by a Stroop task while watching a film clip in between which induced positive, negative, or neutral affect. We implemented the same stability condition as in Study 1a (Stroop in first task; followed by a Stroop). The mean duration of the first Stroop task in the stability condition was 292.41 seconds (*SD* = 31.94). However, due to the novelty of self-control tasks in online experiments, we implemented two different flexibility conditions by varying the first self-control task. In one flexibility condition, the first task was a thought suppression task of 'not thinking about a white bear' (Wegner, Schneider, Carter, & White, 1987) for five minutes. In the other flexibility condition, the first task was an attention control task where participants had to ignore peripheral stimuli while watching a video for five minutes (cf. Schmeichel, Vohs, & Baumeister, 2003). These stimuli were presented in white for 15 s each. Participants in the control condition watched the same video clip as participants in the attention control group but without the peripheral stimuli. As in Study 1a, participants watched a video clip between the two tasks to induce affect. For the negative affect condition, we used an unpleasant film (5 min) featuring a boy who tries to escape abuse and neglect from his parents. The film clips for the positive and neutral affect conditions were the same as in Study 1a.

2.3.2 Results

2.3.2.1 Manipulation Checks

2.3.2.1.1 Perceived difficulty

Participants were asked to rate the perceived difficulty of the first task, resulting in a significant main effect in a one-way ANOVA for task switching, $F(3, 172) = 17.05$, $MSE = 0.957$, $p < .001$, $\eta^2 = .229$. Focused comparisons revealed that, compared to the control condition ($M = 1.32$, $SD = 0.59$), the perceived difficulty of the three task switching conditions involving the Stroop ($M = 2.33$, $SD = 0.85$, $t(172) = 4.54$, $p < .000$, $d = 1.032$), thought suppression ($M = 2.59$, $SD = 1.19$, $t(172) = 5.81$, $p < .001$, $d = 1.297$) and attention control task ($M = 2.81$, $SD = 1.07$, $t(172) = 6.79$, $p < .001$, $d = 1.522$) differed significantly. The task switching conditions differed not among each other, $t(172) \leq 1.28$, $p \geq .202$, $d = 0.264$, except for the difference between the Stroop and the attention control task, $t(172) = -2.34$, $p = .019$, $d = -0.490$. Given that the regulation of a conflict requires effort, this result indicates that our self-control tasks worked as intended, although the attention control task was rated as more difficult than the Stroop task.

2.3.2.1.2 Affect manipulation

We measured affect twice to assess changes through our affect induction. For the valence scale of the MDBF, a 4 (task switching) X 3 (affect) ANCOVA with the baseline affect measure included as a covariate revealed a significant main effect of affect, $F(2, 163) =$

21.61, $MSE = 0.394$, $p < .001$, $\eta_p^2 = .210$. We used focused comparisons again to investigate the group differences: Whereas participants in the neutral affect condition ($M_{adjusted} = 3.49$, $SEM = .080$) were in a better mood than participants in the negative affect condition ($M_{adjusted} = 3.03$, $SEM = .081$), $t(163) = 4.06$, $p < .001$, $d = 0.736$, they felt significantly less positive than those in the positive affect condition ($M_{adjusted} = 3.82$, $SEM = .086$), $t(163) = -2.81$, $p = .006$, $d = -0.522$. There were neither a main effect for task switching, $F(2, 163) = 1.57$, $MSE = 0.394$, $p = .199$, $\eta_p^2 = .028$, nor an interaction between task switching and affect in mood scores, $F(2, 163) = 1.12$, $MSE = 0.394$, $p = .353$, $\eta_p^2 = .040$.

Whereas the affect induction did not influence the energetic arousal scale of the MDBF differently, $F(2, 163) = 1.22$, $MSE = 0.262$, $p = .298$, $\eta_p^2 = .015$, there was a main effect of affect on tense arousal after the affect induction with baseline tense arousal included as a covariate, $F(2, 163) = 10.96$, $MSE = 0.371$, $p < .001$, $\eta^2 = .12$. Simple comparison tests indicated that participants in the negative affect condition ($M_{adjusted} = 2.98$, $SEM = .079$) were more strained by the video clip than participants in the neutral ($M_{adjusted} = 2.48$, $SEM = .077$), $t(163) = 4.62$, $p < .001$, $d = 0.836$, or positive affect condition ($M_{adjusted} = 2.53$, $SEM = .083$), $t(163) = 3.99$, $p < .001$, $d = 0.749$, whereas neutral and positive affect did not differ from each other, $t(163) = -0.47$, $p = .640$, $d = -0.087$. To sum up, the manipulation checks showed medium to large effect sizes for affect demonstrating that our induction of affect worked as intended.

2.3.2.2 Stroop Performance

For the decrements in self-control task performance, a 4 (task switching) X 3 (affect) ANOVA on the Stroop effect in the second task did not show a significant main effect for affect, $F(2, 164) = 1.41$, $MSE = 12667$, $p = .247$, $\eta_p^2 = .017$, but a marginally significant main effect for task switching, $F(3, 164) = 2.59$, $MSE = 12667$, $p = .055$, $\eta_p^2 = .045$, and a significant interaction effect, $F(6, 164) = 3.13$, $MSE = 12667$, $p = .006$, $\eta_p^2 = .103$. This interaction is shown in Figure 4. Again, we first tested contrasts to examine the different predictions. The (PA NE) contrast weights for the SMSC were (-1 1) for the flexibility 1 conditions, (-1 1) for the flexibility 2 conditions, (-1 1) for the stability conditions, and (0 0) for the control conditions. The contrast test yielded no significant effect, $F(1, 164) = 1.17$, $p = .281$, $d = 0.408$. Conversely, the (PA NE) contrast weights for the CMT were (-1 1) for the flexibility 1 conditions, (-1 1) for the flexibility 2 conditions, (2 -2) for the stability conditions, and (0 0) for the control conditions, resulting in a significant contrast test, $F(1, 164) = 9.63$, $p = .002$, $d = 1.171$.

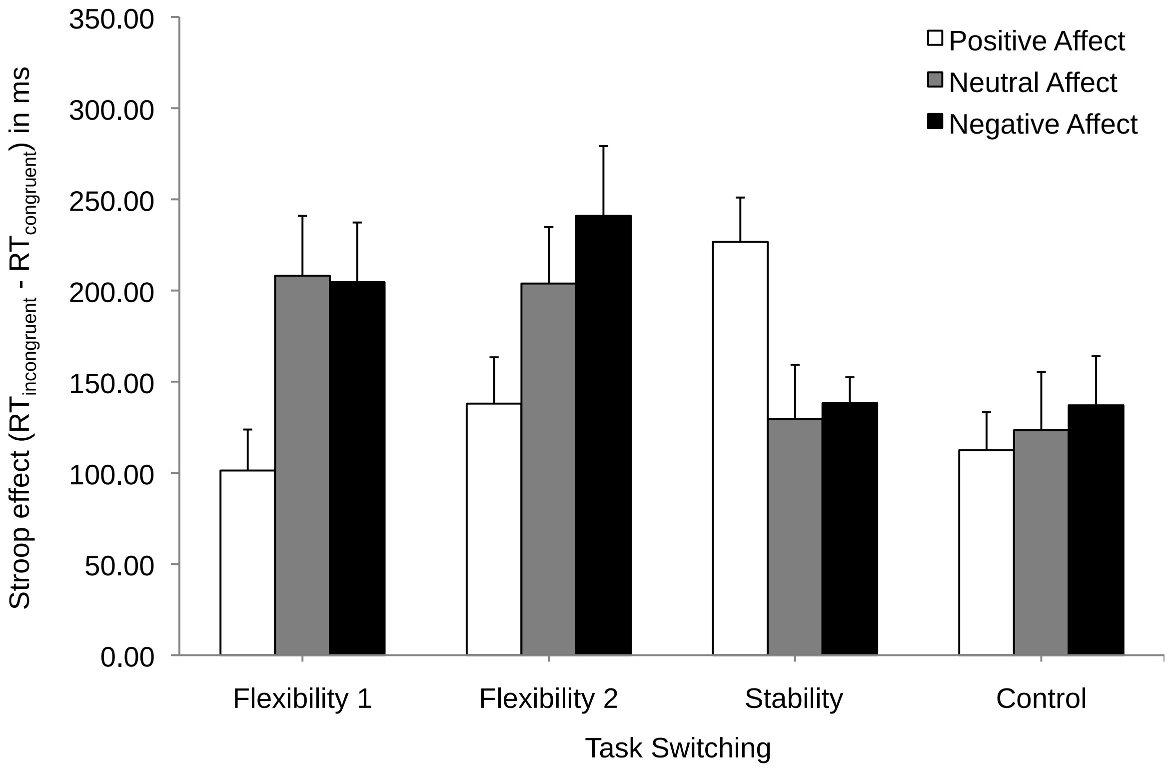

Figure 4. Mean Stroop effect in the second task as function of task switching and affect (Study 1b). Higher scores indicated poorer performance. Task switching indicates whether the second Stroop task was preceded by a different self-control task (Flexibility 1 = thought suppression,

Flexibility 2 = attention control), by the same Stroop task (Stability) or by no self-control task (Control). Error bars reflect standard errors.

To further evaluate the interaction shown in Figure 4, we computed simple main effects and focused comparisons for task switching and affect. The simple main effect of affect in the flexibility condition with thought suppression was significant, $F(2, 164) = 4.71$, $p = .009$, in that participants in the positive affect condition ($M = 101.29$ ms, $SD = 89.31$) showed smaller Stroop effects in the second task than participants in the neutral ($M = 208.21$ ms, $SD = 135.26$, $t(164) = -2.73$, $p = .007$, $d = -0.950$) or negative affect condition ($M = 204.60$ ms, $SD = 130.74$, $t(164) = -2.60$, $p = .010$, $d = -0.918$). Neutral and negative affect, however, did not differ from each other, $t(164) = -0.09$, $p = .927$, $d = -0.032$. This pattern reflects the beneficial effect of positive affect on self-control performance for different consecutive response conflicts, whereas negative affect was not influential. For the attention control flexibility condition, the simple main effect of affect was also significant, $F(2, 164) = 3.26$, $p = .039$, but only positive affect ($M = 137.95$ ms, $SD = 89.31$) differed significantly from negative affect ($M = 241.02$, $SD = 157.68$, $t(164) = -2.54$, $p = .012$, $d = -0.916$). Neither positive nor negative affect differed significantly from neutral affect, $t(164) = -1.62$, $p = .107$, $d = -0.585$ and $t(164) = 0.96$, $p = .337$, $d = -0.330$, respectively. As was the case with the other flexibility condition, positive affect improved self-control performance but failed to reach significance for the attention control flexibility condition. The simple main effect of affect in the stability condition was also significant, $F(2, 164) = 3.18$, $p = .041$. Participants in the positive affect condition ($M = 226.72$ ms, $SD = 87.73$) had significantly larger Stroop interferences than participants in the neutral ($M = 129.64$ ms, $SD = 122.36$, $t(164) = 2.34$, $p = .020$, $d = 0.863$) or negative affect condition ($M = 138.20$ ms, $SD = 55.16$, $t(164) = 2.08$, $p = .039$, $d = 0.786$), whereas the latter conditions did not differ significantly, $t(164) = -0.21$, $p = .830$, $d = -.076$. In contrast to both flexibility conditions, positive affect in the stability condition was not helpful but harmful for self-control performance whereas negative affect did not influence the Stroop performance. There were not any differences for affect in the control condition, $F(2, 164) = 0.14$, $p = .871$.

The ANOVA also yielded a marginally significant main effect of task switching. This was driven by differences between the flexibility and control conditions: Compared to the control condition ($M = 124.73$ ms, $SD = 88.61$), participants in the attention control flexibility condition showed a significantly larger Stroop interference ($M = 197.78$ ms, $SD = 135.25$, $t(164) = 2.90$, $p = .004$, $d = 0.649$) and participants in the thought suppression flexibility condition showed a marginally significant larger Stroop interference ($M = 172.12$ ms, $SD = 128.25$, $t(164) = 1.89$, $p = .061$, $d = 0.421$). The other differences were not significant, $t(164) \leq 1.59$, $p \geq .113$, $d \leq 0.330$.

As we found significant differences in tense arousal, we tested if they could account for the differences in the Stroop effect. Therefore, we computed an ANCOVA with task switching and affect and included the tense arousal measure after the mood induction and its interaction terms with the independent variables as covariates, testing the full ANCOVA model, according to Muller, Yzerbyt and Judd (2008). Using the adjusted means of that ANOVA, the contrast test for the prediction of the CMT revealed a smaller but still big effect size, $F(1, 152) = 6.74$, $MSE = 12507$, $p = .010$, $d = 0.979$, whereas the contrast for the SMSC was still not significant, $F(1, 152) = 0.07$, $MSE = 12507$, $p = .0.788$, $d = 0.102$. These data suggest that the included interaction of tense arousal and task switching did not account for the interaction of affect and task switching on the Stroop effect in the second task.

The mean number of errors was $M = 4.55$ ($SD = 3.39$), which did not differ from Study 1a, $t(283) = -0.66$, $p = .507$, $d = -0.080$. Like with Study 1a, we did not find any differences in the Stroop error rates, neither for task switching, $F(3, 164) = 1.74$, $MSE = 11.21$, $p = .161$,

η_p^2 = .031, affect, $F(2, 164) = 2.64$, $MSE = 11.21$, $p = .074$, η_p^2 = .031, nor their interaction, $F(6, 164) = 0.67$, $MSE = 11.21$, $p = .671$, η_p^2 = .024.

2.3.3 Discussion

We were able to replicate the findings from Study 1a showing that positive affect interacted with task switching on consecutive self-control performance. As with Study 1a, findings supported the predictions from CMT by demonstrating that positive affect was beneficial when the response conflicts were different in both tasks (flexibility condition with thought suppression) but harmful when the response conflicts were the same (stability condition), although the simple effect of affect for the flexibility condition with attention control was not significant. However, unlike the results of Bruyneel and colleagues (2009), we did not find any significant effects of negative affect on decrements in self-control performance, supporting the null findings by Tice and colleagues (2007). Consequently, the interactive effect of affect and task switching is unique to positive affect, which underscores the impact of positive affect on consecutive self-control performance.

2.4 General Discussion

We investigated the effects of task switching and positive (Study 1a & 2b) and negative affect (Study 1b) on performance in two consecutive self-control tasks by comparing the diverging predictions of the SMSC and the CMT. Across both studies, results favored the CMT by showing that positive affect only improved performance in consecutive self-control tasks when the tasks required switching a cognitive set. This interaction between positive affect and task switching was found in both experiments. When participants completed a resisting sweets (Study 1a) or thought suppression task (Study 1b) followed by a Stroop task (flexibility condition), participants who watched a funny film clip performed better than participants who watched a neutral film clip and they performed similarly to those who did not have to exercise self-control in the first task. However, when participants repeated the Stroop task with different trials twice (stability condition), participants watching a funny film clip performed worse compared to those who watched the neutral film clip.

These findings provide support for the CMT, which predicts an interaction between task switching and positive affect on performance without the recourse to a limited resource. Within a cognitive control view (Dreisbach & Goschke, 2004), positive affect improves flexibility at the cost of decreased stability and increased distractibility and, therefore, should benefit self-control performance when flexibility is required. For consecutive self-control tasks with different response conflicts, increased flexibility is functional because participants have to switch to a new conflict in the second task and, therefore, participants in a positive state perform better than participants in neutral state. For consecutive self-control tasks with the same response conflict, positive affect induces increased flexibility, which hampers performance in the second task compared to neutral affect. These differences cannot be explained within the SMSC and the restoration hypothesis of positive affect, which would predict that a mild positive mood induction should increase self-control performance regardless of task switching or, at least, should not decrease performance compared to no induction. Therefore, the interaction of task switching and affect does not support the restoration hypothesis of positive affect within the SMSC.

We did not derive specific hypotheses for negative affect since evidence for the effects of negative affect is not consistent. Whereas Tice and colleagues (2007) could not find any effects, Bruyneel and colleagues (2007) showed that the induction of negative affect resulted in impaired self-control performance in a second self-control task. However, our results are in line with the null findings by Tice and colleagues, in that self-control performance in the negative affect conditions did not differ from the neutral affect condition within any task switching condition. Moreover, this finding suggests that the stability-flexibility-balance is not impacted by negative affect but by positive affect only.

Several limitations have to be noted with regard to our experiments. First, since we only manipulated valence in our mood induction, we cannot certainly rule out that differences in Stroop task performance across conditions were driven by arousal instead of valence. However, we did not find any differences in arousal after the mood induction in Study 1a. We did find them for tense arousal in Study 1b, but they did not account for the interaction between task switching and affect on the Stroop effect. Even if our two arousal measures were not sensitive enough to detect differences in activation so that participants found the positive film clip also more arousing than the neutral one, research about the interplay of valence and arousal on self-control performance argues against arousal as an alternative explanation for our results. Fedorikhin and Patrick (2010) demonstrated that elevated arousal dampens the beneficial effects of positive mood when exerting self-control, leading to no differences compared to neutral mood. We found that positive affect is beneficial (flexibility) or harmful (stability) depending on task switching. This finding would only be in line with the dampening effect reported by Fedorikhin and Patrick (2010) if somehow positive affect reflected high arousal in the stability condition but low arousal in the flexibility condition. However, the tense arousal main effect of affect was driven by differences in the negative affect conditions and not by any differences in the positive affect conditions. Taken together, we could not find any evidence that the difference in Stroop performance can be attributed to differences in tense arousal.

Second, we measured the perceived difficulty of the first task before the second task and found no difference between the self-control tasks. However, it would be interesting to assess if the effect of adaptation (task switching) and flexibility (via positive affect) is mirrored in the perceived difficulty of the second task. This may shed light on whether participants who are adapted to the conflict perceive this task as less difficult than participants who encounter a new conflict, and, whether these possible differences in perceived difficulty of the second task could be influenced by positive affect.

Third, the balance of stability and flexibility was tested only by the Stroop effect in the second task. It would be beneficial in future research to use a task with a more direct measure of adapting, task switching, and switching costs like the cognitive set-switching paradigm by Dreisbach and Goschke (2004).

Fourth, the samples of both experiments were medium sized, with 108 participants in Study 1a and 176 in Study 1b. This led to low numbers of participants in some cells. However, the statistical power was large enough to detect the contrast effect of the predictions of the CMT in both experiments, whereas the contrast effect size of the SMSC was very small.

To summarize, the interaction between task switching and positive affect, which was derived from an integration of findings from ego depletion and cognitive control literatures, was supported. Positive affect and its connection to flexibility counteract decrements in self-control task performance only when flexibility is helpful. In our view, drawing on the CMT could add substantially to the self-control research and provide appropriate theoretical guidance about the mechanisms involved in self-control effort and decrements in self-control.

Chapter 3:

Ecological Validity of the Effect of Positive Affect on Limited Self-Control[2]

3.1 Introduction

People are faced with numerous challenges to self-control in everyday life. They may be trying to eat healthier, exercise more, drink less alcohol, work harder, or curb what they say (or do not say) – all in an effort to modify unwanted behavior. To date, most studies have examined self-control success and failure in controlled laboratory settings (e.g., Hagger, Wood, Stiff, & Chatzisarantis, 2010). Recent studies, however, expand on these approaches emphasizing the importance of everyday behavior in self-control (e.g., Berkman, Falk, & Lieberman, 2011; Hofmann, Baumeister, Förster, & Vohs, 2012). Studies about the effectiveness of everyday self-control strategies are still scarce (Quinn, Pascoe, Wood, & Neal, 2010). Specifically, more research is needed that examines the effectiveness specific strategies people employ in daily life to regulate temptations and habits and to identify possible moderating factors associated with self-control success. One promising candidate is positive affect: Laboratory research showed that positive affect can facilitate or hinder self-control depending on self-control demands (Aspinwall, 1998; Wenzel, Conner, & Kubiak, 2013). However, to the best of our knowledge, it is still unclear if and how this evidence translates to self-control as it occurs in the everyday life of individuals. Thus, the goal of the present daily diary study was to investigate how positive affect influences self-control success by facilitating (or thwarting) different self-control strategies.

[2] This study was published in 2016 in a peer-reviewed journal: Wenzel, M., Kubiak, T., & Conner, T. S. (2016). Self-control in daily life: How affect may boost or sabotage efforts at self-control. *Social Psychological and Personality Science, 7,* 195-203.

Self-control, often defined as the altering of one's own impulses to attain goals (Muraven, Tice, & Baumeister, 1998), requires both stable maintenance of current goals and flexible goal switching (Goschke, 2003). On the one hand, currently active goals need to be maintained and protected from distraction in order to achieve goal attainment. On the other hand, flexibility is needed when switching between goals or when disengaging from an unachievable goal. Thus, successful self-control requires a context-sensitive balance between stability (maintaining goal intentions) and flexibility (switching goal intentions). If this balance is off, individuals could suffer from perseveration or they could suffer from impulsivity and distractibility (Dreisbach & Goschke, 2004).

Research has shown that positive affect influences this stability-flexibility-balance. For example, positive affect is associated with increased cognitive flexibility in adopting abstract future goals, particularly self-control goals, whereas negative affect is linked to focused attention on immediate goals like mood management (Fishbach & Labroo, 2007; Labroo & Patrick, 2009). Positive affect is also known to broaden one's attentional focus (Fredrickson, 2004) and lead to a broader repertoire of action plans (Fredrickson & Branigan, 2005). Dreisbach and Goschke (2004) showed that positive affect enhances flexibility and distractibility at the cost of reduced stability and focus in a cognitive task. When new information was linked to a target response, participants in a positive affective state performed better than participants in a neutral affective state since directing attention to the novel (helpful) information promoted the processing of the target. Conversely, when novel information was linked to the distractor, participants in a positive state performed worse because the increased focus on the new (distracting) information interfered with the target response. In a similar vein, Wenzel, Kubiak, and Conner, (2014) demonstrated previously that the flexibility-enhancing effect of positive affect also holds for self-control and that greater cognitive flexibility explained the influence of positive affect on consecutive self-control performance: If novel information in the second task was linked to its target, positive affect improved performance compared to neutral affect. In turn, positive affect led to worse performance if the novel information was linked to the distractor in the second task. These results demonstrate that the effect of positive affect on cognitive and self-control depends on whether the novel information individuals switch to is beneficial or not for the given task goal.

However, so far, the evidence for the flexibility-enhancing effect of positive affect is based on rather short laboratory tasks. It is unclear whether such an effect occurs in daily life as well. Ambulatory assessment represents a necessary complementary strategy to investigate the ecological validity of results found in laboratory research (Kubiak & Stone, 2012). In order to assess demands for stability and flexibility in self-control situations in daily life, we adopted the self-control strategies from Quinn and colleagues (2010) since they affect the stability-flexibility-balance differently, as explained later. Quinn and colleagues (2010) identified three main strategies of self-control based on the delayed gratification research: *vigilant monitoring* of one's own behavior, *distraction* (e.g., thinking about something else) and *stimulus control* that is reducing or avoiding the salience of a cue (e.g., leaving the situation). Whereas in monitoring, one focuses attention on the behavior or the cognitions that should be controlled, distraction concentrates attention on something else. Quinn and colleagues (2010) further demonstrated that the effectiveness of these self-control strategies depends on whether the unwanted behaviors are temptations or habits. *Temptations* arise when an impulse like eating a tempting donut interferes with a long-term goal like reducing weight (Nordgren & Chen, 2011). *Habits* are automatically activated responses in a consistent context (Wood & Neal, 2007) like reading the newspaper every morning or taking a walk every afternoon. Quinn and colleagues (2010) found that monitoring was less successful than stimulus control in resisting temptations, whereas monitoring was more successful than stimulus control for regulating habits, suggesting that focusing on a tempting impulse is detrimental (Hofmann, Deutsch, Lancaster,

& Banaji, 2009) but focusing on a habit is not. However, contrary to Quinn and colleagues' expectations, distraction did not differ from stimulus control or monitoring.

In the present research, we propose that positive affect interacts with self-control strategies in explaining differences in self-control success in everyday life. Specifically, we predict that positive affect will enhance the success of distraction in daily life when confronted with strong temptations, whereas it impairs the effectiveness of monitoring. Distraction favors flexibility over stability, in that to distract oneself from a temptation, individuals have to modify the response disposition to the tempting situation by stepping back and disengaging oneself from the impulse in order to think about something else. If positive affect helps people to improve cognitive flexibility by facilitating processing of new information (Dreisbach & Goschke, 2004; Wenzel et al., 2014), it should increase self-control success when people are using distraction as a self-control strategy in daily life. If the effectiveness of distraction depends on positive affect, it could explain the null findings for distraction by Quinn and colleagues (2010). In contrast, monitoring favors stability over flexibility, in that stable attention is required in order to ensure that a response is not performed. Thus, positive affect should negatively interact with monitoring. This two-way interaction (positive affect by self-control strategy) should be also influenced by temptation and habit strength, in that distraction should be particularly effective in regulating strong temptations compared to the other self-control strategies when people feel positive.

3.2 Method

3.2.1 Participants

A total of 327 undergraduate students at the University of Otago, New Zealand, participated for an opportunity to earn research credits or payment (maximum of NZ$55).

Those 327 participants reported 3,670 observations. Data from nine of these participants (*n* = 37 observations) were excluded because they completed fewer than 50% of the daily surveys. On 922 days (25.4% of actual answers), participants did not report an unwanted behavior, which reduced the total observations for our analyses to a total of 2,711 and the total number of participants to 310. Since missings in other measures of interests were not imputed, another 13 participants (*n* = 214 observations) were excluded due to missings, leaving a total of 297 participants (65% female) and 2,497 observation for analysis. These participants were aged 17 to 30 years (*M* = 19.93 years, *SD* = 2.32). In terms of ethnicity, 73.6% self-identified as European, 10.4% as Asian, 4.0% as Maori/Pacific Islander, and 12.0% from other ethnic backgrounds.

3.2.2 Procedure

This study was part of a combined laboratory and 13-day daily diary study called the Daily Life Study. At the start of this study, participants were invited to the lab, where they provided informed consent and completed self-report measures of demographics over the computer. Starting the next day, for 13 days total, participants logged onto a secure website between 3pm and 8pm to complete a brief 5-minute questionnaire that included the following measures embedded within other measures of daily experience. Thus, the daily diary comprised one query per day that retrospectively assessed the domain and nature of the unwanted behavior and how it was regulated. After two weeks, participants returned to the lab and were debriefed and reimbursed for their participation.

3.2.3 Measures

3.2.3.1 Domain

To control for the influence of the type of event that the individuals encountering at a given moment, participants were given a list of examples of unwanted behaviors derived from the list of frequently reported unwanted activities reported by Quinn and colleagues (2010), e.g. procrastinating or sleeping. In each daily diary, participants picked one that they tried to inhibit or change that respective day. If they experienced more than one unwanted behavior on that day, they were instructed to indicate the strongest one, which was subject to the further measures.

3.2.3.2 Affect

Participants rated how they felt "before the situation with the unwanted behavior" using two items of the valence subscale of the short form of the German questionnaire Multidimensional Mood Questionnaire (MDMQ; Steyer, Schwenkmezger, Notz, & Eid, 1997). The MDMQ assesses affect on three dimensions: valence (items: good, bad), energetic arousal, and tense arousal on a scale ranging from 1 (*not at all*) to 5 (*extremely*). Mean scores were computed for the valence dimension (with one of the items reversed-scored) with higher scores representing higher valence (i.e. more positive affect). Summary statistics for affect and the other measures are presented in

Table 3.

3.2.3.3 Self-control strategies

Participants then checked the strategy they used to inhibit or change the unwanted behavior, with: (1) monitoring (36.2% of the total number of diary entries across individuals), (2) distraction, (26.8%) stimulus control (16.0%), (4) nothing (17.2%) or (3.8%) other. Definitions and examples of each strategy were listed in the diary instructions. The examples were monitoring (using a watch to be ontime; monitoring alcohol use), distraction (thinking about something else; doing another activity), stimulus control (removing myself from the situation or removing the opportunity to do it such as avoiding certain triggers; leaving early from a pub), nothing (I didn't try any strategies to inhibit or change unwanted behaviour). We excluded the "other" category since we were interested in the interactions of positive affect and specific self-control strategies.

3.2.3.4 Desire strength

Since temptations are based on desires, the enactment of these should lead to positive feelings in that moment. Therefore, temptations were measured by asking participants "how much would performing the unwanted behavior make you feel good" in that situation from 1 (*not at all*) to 5 (*extremely*), with higher scores indicating greater desire strength.

3.2.3.5 Habit strength

Habit strength (M = 2.60, SD = 0.78) was assessed with the single item "How often have you performed the unwanted behavior in the past?" on a scale ranging from 1 (*monthly or less often*) to 4 (*several times per day*).

3.2.3.6 Self-control success

The dependent variable of self-control success was measured by "how successful was the attempt to change or inhibit the unwanted behavior?" from 1 (*not at all*) to 5 (*completely*).

Table 3

Range, means, standard deviations, reliability, and intercorrelations on Self-Control Success, Desire Strength, Habit Strength, Valence (both before and after the act of self-control)

Measure	Range	*M*	*SD*	R_{KRN}	R_{CN}	1	2	3	4
1. Self-Control Success	1 - 5	2.64	1.11	.82	-				
2. Desire Strength	1 - 5	2.22	1.23	.85	-	.01			
3. Habit Strength	1 - 4	2.55	0.80	.88	-	-.10	.05		
4. Valence (before act of self-control)	1 - 5	3.55	0.80	.81	.56	.12	.04	-.04	
5. Valence (after act of self-control)	1 - 5	3.52	0.88	.77	.66	.42	.08	-.10	.39

Note. R_{KRN} = between-person reliability and R_{CN} = within-person reliability (day-to-day changes, see Shrout & Lane, 2012). The presented variables are not person-mean centered.

3.2.3.7 Control variables

Control variables in all analyses were age, gender, domain, day of study, and day of the week. Since daily diaries involve retrospective assessment of recent thoughts and behavior, individuals may not be able to distinguish their affect before the unwanted behavior from their affect afterwards. For instance, the affect measures may be biased by individuals' feeling about how they managed these events, in that if they managed them well, positive affect after the situation increased, which may lead to higher positive affect measure before the situation. Therefore, participants, also rated how they felt "after the situation with the unwanted behavior" using the same MDMQ valence scale. There was no evidence for multicollinearity, with both measures correlating only moderately, $r = .39$. Thus, controlling for affect after the self-control, an independent affect coefficient for positive affect before the situation could be obtained.

3.2.4 Analytic Approach

We computed multilevel models with random intercepts and coefficients in Stata 13 (Stata Corporation, College Station, TX, USA) with daily diary observations (level 1) nested within participants (level 2). All continuous level-1 independent variables were person-mean centered and all continuous level-2 independent variables were grand-mean centered (Enders & Tofighi, 2007). Since we were only interested in the within-subject processes, the averaged continuous level-1 independent variables were not entered into the model (Bolger & Laurenceau, 2013). Categorical variables were dummy-coded (self-control strategies reference: "doing nothing"; gender reference: "male"; domain: "other") since we were primarily interested in continuous x categorical interactions (Cohen, Cohen, West, & Aiken, 2003).

3.3 Results

Adherence with the daily diary protocol was good, with participants completing more than 11 of the 13 possible daily diaries on average ($M = 11.11$, $SD = 2.14$; 86% adherence). To test whether participants completed fewer diaries across the course of the study, we computed a chi-square goodness-of-fit test which did not differ from an equal distribution, $\chi^2(12) = 17.19$, $p = ns$. Consequently, compliance did not deteriorate over the two weeks of the study.

We first fitted the unconditional model without predictors on self-control success, which revealed that 74.7 % of the total variance in self-control success was attributable to within-person variability. We, then, included all level-1 and level-2 variables to compute the main effects and added the two- and three-way interactions in the next step. We tested whether treating the level-1 predictors as random improved model fit, which was the case, $\chi^2(4) = 103.60$, $p < .001$. Since including random covariances did not further improve the model fit significantly, $\chi^2(6) = 4.33$, $p = .614$, we only included the random effects without their covariances. The final model (Table 4) accounted for 37.3% of the variance in self-control success. To facilitate the presentation of the results in Table 4, we do not present the coefficient for each strategy-by-moderator interaction but an omnibus test, which indicates a significant interaction. This omnibus can be interpreted like the F-value in an ANOVA and F-values can be obtained by dividing the chi-square values by its degrees of freedom. To interpret significant differences, we then show the simple slopes (Cohen et al., 2003) for each self-control strategy (Table 5).

Table 4

Fixed Effects Estimates (Top) and Variance Estimates (Bottom) of Self-Control Success as Function of Self-Control Strategies (SCS), Desire and Habit Strength, and Valence

Within-subjects fixed effects	Estimate	*SE*	*z*	p^a	CI_{95} Lower	CI_{95} Upper
Step 1						
SCS	$\chi^2(3, N = 2497) = 136.58$			<.001		
Desire Strength	0.01	0.02	0.22	.828	-0.04	0.05
Habit Strength	-0.05	0.04	-1.44	.149	-0.12	0.02
Valence (before act of self-control)	0.04	0.03	1.23	.218	-0.02	0.10
Valence (after act of self-control)	0.43	0.03	17.29	<.001	0.38	0.48
Step 2						
Desire Strength * Valence	0.04	0.03	1.41	.158	-0.02	0.10
Habit Strength * Valence	0.05	0.05	1.08	.280	-0.04	0.15
SCS * Desire Strength	$\chi^2(3, N = 2497) = 6.85$			.077		
SCS * Habit Strength	$\chi^2(3, N = 2497) = 6.27$			.099		
SCS * Valence	$\chi^2(3, N = 2497) = 8.98$			.030		
Step 3						
SCS * Desire Strength * Valence	$\chi^2(3, N = 2497) = 15.06$			.002		
SCS * Habit Strength * Valence	$\chi^2(3, N = 2497) = 3.13$			.366		
Intercept	2.30	0.39	5.84	< .000	1.53	3.07
			Random parameters			
SCS	0.18	0.03	5.56	< .001	0.14	0.26
Desire Strength	0.04	0.01	3.39	.001	0.03	0.08
Habit Strength	0.07	0.03	2.56	.010	0.03	0.15
Valence	0.02	0.02	1.26	.207	0.01	0.11
Intercept	0.21	0.04	5.96	< .001	0.14	0.28

Notes. SCS = Self-control strategies. *N* = 297 persons, 13 days, 2,497 observations. Self-control success was measured on a scale from 1 to 5. Chi-square tests reflect omnibus tests for interactions of the self-control strategies dummy variables.

[a]All *p*-values are two-tailed except in the case of the random parameters, where one-tailed *p*-values are used (because variances are constrained to be non-negative).

Table 5

Simple slopes for each Self-Control Strategy (SCS) and Valence, Desire Strength, Habit Strenght, Valence x Desire Strength

					CI_{95}	
Measure	Estimate	*SE*	*z*	p^a	Lower	Upper
Valence						
Monitoring	-0.05	0.05	-0.93	.353	-0.14	0.05
Distraction	0.17	0.06	3.03	.002	0.06	0.28
Stimulus Control	0.04	0.08	0.42	.675	-0.13	0.20
Doing Nothing	0.11	0.08	1.38	.169	-0.05	0.26
Desire Strength						
Monitoring	-0.02	0.04	-0.49	.623	-0.09	0.05
Distraction	0.10	0.04	2.48	.013	0.02	0.18
Stimulus Control	-0.03	0.05	-0.51	.611	-0.12	0.07
Doing Nothing	-0.01	0.05	-0.18	.859	-0.11	0.09
Habit Strength						
Monitoring	0.01	0.05	0.26	.791	-0.09	0.12
Distraction	-0.01	0.06	-0.07	.948	-0.13	0.12
Stimulus Control	-0.20	0.08	-2.42	.016	-0.36	-0.04
Doing Nothing	-0.11	0.08	-1.47	.143	-0.27	0.04
Valence x Weak Desires						
Monitoring	-0.11	0.07	-1.70	.089	-0.25	0.02
Distraction	-0.01	0.07	-0.22	.829	-0.15	0.12
Stimulus Control	0.07	0.11	0.66	.506	-0.14	0.28
Doing Nothing	0.25	0.10	2.52	.012	0.06	0.44
Valence x Strong Desires						
Monitoring	0.02	0.07	0.33	.739	-0.11	0.16
Distraction	0.35	0.09	3.97	<.001	0.18	0.52
Stimulus Control	-0.01	0.11	-0.01	.992	-0.23	0.22
Doing Nothing	-0.04	0.11	-0.37	.714	-0.25	0.17

3.3.1 Effectiveness of self-control strategies

In the first step, we entered the main effects into the model (Table 4), which indicated significant differences between the three self-control strategies. Simple main effects analyses with pair-wise comparisons of the marginal means revealed that individuals reported significantly more self-control success if they had used any of the three self-control strategies monitoring (M = 2.79, SE_M = 0.05, p <.001), distraction (M = 2.80, SE_M = 0.05, p <.001), and stimulus control (M = 2.98, SE_M = 0.07, p <.001) compared to doing nothing (M = 2.05, SE_M = 0.07). Moreover, among the self-control strategies, stimulus control was more effective than to monitoring, p = .007, and distraction, p = .013, with no difference between the latter two strategies, p = .861. Neither desire or habit strength nor valence exhibited significant main effects on self-control success.

In the second step of the analysis, we included the two-way interactions. Self-control strategies significantly interacted with valence, p = .030, and marginally significantly with desire strength, p = .077, and habit strength, p = .099. As illustrated in Table 5, simple slop analysis revealed that positive affect only influenced distraction significantly, in that distraction was more effective with increasing positive affect. In turn, the success of monitoring and stimulus control as well as doing nothing was not significantly influenced by positive affect. The same pattern was revealed for the simple slope analysis with desire strength, in that only distraction was more effective with increasing desire strength. For habit strength, stimulus control was less effective when dealing with stronger habits, whereas monitoring was not impacted by habit strength.

Both two-way interactions between self-control strategies and valence or desire strength, respectively, were qualified by a three-way interaction between self-control strategies, desire strength, and valence in the third step of the hierarchical analysis, p = .002. Simple slope analyses revealed that the positive effect of positive affect on the success of distraction increased with increasing desire strength, with a large effect for strong desires (+1 *SD* below the mean; see *Figure 5*, right panel) and no significant relationship for weak desires (-1 *SD* below the mean). Moreover, the simple slope of monitoring for weak desires was marginally significant, that is positive affect yielded a negative effect when participants reported using monitoring for weak desires (see *Figure 5*, left panel). These results demonstrate the flexibility-enhancing effect of positive affect, in that the success of distraction in regulating strong tempting desires highly depends on valence.

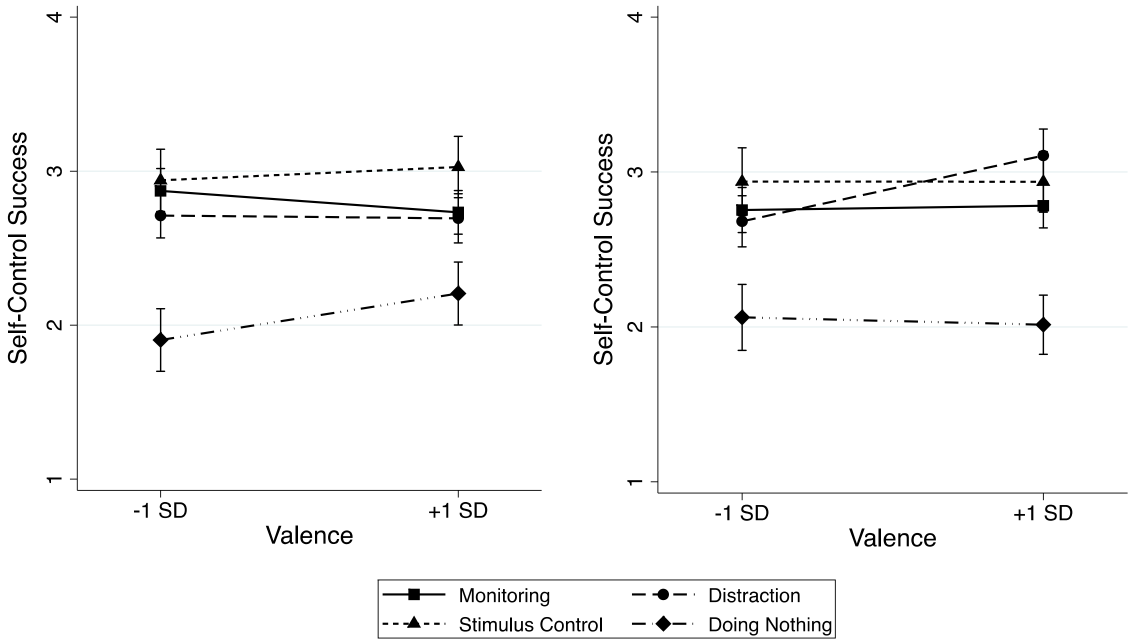

Figure 5. Marginal means of self-reported self-control success as a function of valence for weak (-1 SD of mean of desire strength, left panel) and strong temptations (+1 SD of mean of desire strength, right panel). The bars at each marker reflect 95% CI.

3.3.2 Choice of self-control strategy

Next, we investigated how the primary constructs influenced the choice of self-control strategies. To this end, we computed a multilevel multinomial logistic regression with self-control strategies as the outcome. We examined self-control success, desire strength, habit strength, valence, and its interactions as the predictors, using the gsem command in Stata 13. As illustrated in Table 6, participants facing strong habits chose more likely monitoring and less likely distraction or stimulus control. Although desire strength did not directly influence the likelihood for any self-control strategy, it interacted with self-control success: Less effective regulators (1 *SD* < *M* of self-control success) chose monitoring marginally significantly more often, $b = 0.03$, $SE = 0.02$, $p = .076$, 95% CI [-0.00, 0.06], and distraction significantly less often, $b = -0.04$, $SE = 0.02$, $p = .020$, 95% CI [-0.07, -0.01], when confronted with strong compared to weak desires, whereas effective regulators (1 *SD* > *M*) did not, $b = -0.02$, $SE = 0.02$, $p = .348$, 95% CI [-0.05, 0.02] and $b = 0.01$, $SE = 0.02$, $p = .399$, 95% CI [-0.02, 0.04], respectively. However, as noted before, the interaction on the likelihood of monitoring failed to reach significance.

With increasing valence, participants were less likely to choose distraction and more likely to choose stimulus control, even if they reported to be more successful when using distraction under positive affect. Moreover, valence interacted with self-control success if distraction was chosen: Less effective but not more effective regulators reported to choose distraction less often when regulating strong desires, $b = -0.09$, $SE = 0.02$, $p < .000$, 95% CI [-0.14, -0.05] and $b = -0.02$, $SE = 0.02$, $p = .318$, 95% CI [-0.07, 0.02], respectively. None of the interactions for stimulus control and doing nothing reached significance.

Table 6

Simple Effects of the Multilevel Multinomial Logistic Regression with Self-Control Strategies as the outcome and Self-Control Success, Desire Strength, Habit Strength, Valence and their two-way interactions as the predictors

Measure	Estimate for Monitoring [95% CI]	Estimate for Distraction [95% CI]	Estimate for Stimulus Control [95% CI]	Estimate for Doing nothing [95% CI]
Self-Control Success	0.01 [-0.02, 0.03]	0.01 [-0.02, 0.03]	0.04*** [0.02, 0.06]	-0.05*** [-0.08, -0.03]
Desire Strength	0.01 [-0.01, 0.03]	-0.01 [-0.04, 0.01]	-0.01 [-0.02, 0.01]	0.01+ [-0.00, 0.02]
Habit Strength	0.08*** [0.04, 0.12]	-0.04* [-0.07, -.00]	-0.04* [-0.07, -0.01]	-0.01 [-0.03, 0.00]
Valence	0.01 [-0.03, 0.05]	-0.05** [-0.09, -0.02]	0.03* [0.01, 0.06]	0.01 [-0.00, 0.03]
Self-Control Success x Desire Strength	-0.02 [-0.05, 0.00]	0.02* [0.00, 0.05]	-0.00 [-0.02, 0.02]	0.00 [-0.01, 0.1]
Self-Control Success x Habit Strength	0.01 [-0.03, 0.05]	0.01 [-0.03, 0.04]	-0.01 [-0.04, 0.03]	-0.01 [-0.3, 0.01]
Self-Control Success x Valence	-0.02 [-0.06, 0.02]	0.05* [0.01, 0.08]	-0.02 [-0.06, 0.02]	0.01 [-0.01, 0.02]
Desire Strength x Valence	-0.00 [-0.04, 0.03]	-0.00 [-0.03, 0.03]	0.01 [-0.02, 0.03]	-0.00 [-0.02, 0.01]
Habit Strength x Valence	-0.01 [-0.07, 0.05]	-0.01 [-0.06, 0.05]	-0.00 [-0.05, 0.05]	0.02 [-0.01, 0.04]

Notes. In order to improve clarity of the presentation of our findings, we only present the averaged marginal effects.
* = $p < .05$; ** = $p < .01$; *** = $p < .001$.

3.4 Discussion

We investigated the role of affect on the success of different self-control strategies in daily life. In general, stimulus control was the most effective self-control strategy. In turn, distraction was less effective overall, which was offset under certain circumstances: When participants were experiencing greater positive affect, distraction was reported to be a successful self-control strategy, especially when confronted with strong desires. The results for distraction are in line with the flexibility-enhancing effect of positive affect (Dreisbach & Goschke, 2004) as well as the broaden and build model of positive emotions (Fredrickson, 2004) and provide ecologically valid evidence for the flexibility-enhancing effect. On the grounds of laboratory research, positive affect is linked with higher flexibility and higher distractibility. Therefore, positive affect should help distract oneself from tempting desires, which was found in this study. Thus, the lower self-control success of distraction in general may be due to the increased demand of additional cognitive flexibility, which is facilitated by positive affect. Our findings illustrate why the effects of affect on self-control may be complex at times: When faced with tempting desires, it is better to distract oneself when feeling positive than negative. In short, it is not sufficient to learn which strategy works for which situation – it is also important to consider one's affective state.

Participants with higher self-control success surprisingly did not report weaker habits or stronger desires in general. Basically, stronger temptations or habits should be mirrored in lower self-control success on average: Quinn and colleagues (2010) found a significant main effect for habit strength, in that regulating strong habits was less successful than regulating weak habits, whereas the main effect of habit strength in our study failed to reach significance. It has to be noted, though, that we used person-mean centering of habit strength for our analyses, which differs from the analytical approach in the study by Quinn and colleagues

(2010) where grand-mean centering was used. We deliberately opted for person-mean centering, as grand-mean centering does not allow for separating between within- and between-person processes (Hamaker, 2012), which was crucial for our research questions. Interestingly, when using the grand-mean centered approach by Quinn and colleagues (2010) on our dataset, we find significant negative effect of habit strength on self-control success, $b = -0.09$, $SE = .04$, $p = .012$, as well as a marginally significant positive effect of desire strength, $b = 0.04$, $SE = .02$, $p = .101$, which is in line with the Quinn et al. (2010) findings. Since the person-mean centered measure of habit strength was not significant in our study, the effects found when using the grand-mean centered measures has to be due to between-person differences. Thus, our results demonstrate that participants, who face stronger habits than the ones they usually have to deal with, are not less successful in regulating them (within-person process). Instead, participants who generally face stronger habits than other participants reported less success at doing so (between-person process). However, the random effect of the within-subjects habit strength measure was significant in the final model (Table 4), in that approximately 95% of the participants were within -0.56 and 0.46 ($-0.05 \pm 0.26 \times 1.96$) of the typical value for habit strength. Some individuals reported less success and some more success when confronted with stronger habits. Future research should examine individual differences such as trait self-control in order explain the variability between the individuals in more detail.

In the case of desire strength, neither the within-subject nor the grand-mean centered main effect was significant, although the latter approached significance ($p = .101$), in that stronger desires were more difficult to control on average. This may seem odd since stronger temptations should pose a greater self-control dilemma and be more likely to result in indulgence. Counteractive self-control theory (Fishbach, Zhang, & Trope, 2010) may explain these null findings, as it suggests that temptations could, in fact, also strengthen a long-term goal by indicating a need for self-control. Thus, some individuals indulge when strongly tempted, whereas others could increase their self-control efforts in order to avoid putting their long-term goal at risk. This individual difference in the influence of desire strength on self-control is supported by the significant random effect of desire strength (Table 4). Future research could tap into this variability by directly assessing short- vs. long-term goals (e.g., Hofmann et al., 2012), which enables investigating whether trait self-control is associated with strengthening long-term goals when faced with tempting desires.

Our results also shed some light on some of the null findings by Quinn and colleagues (2010): Contrary to their expectations, distraction did not differ from stimulus control or monitoring in the success of regulating strong desires. We found that the success of distracting from strong desires depended on positive affect, in that distraction was only more effective than stimulus control or monitoring when individuals reported positive affect. In line with the results by Quinn and colleagues (2010), we found self-control control to be less successful in regulating strong habits. We could not replicate, however, that monitoring was the least successful strategy for desires and the most successful one for habits. One possible explanation might be that Quinn and colleagues' (2010) measure of habit strength not only assessed how often the behavior was performed but additionally included the location where the unwanted behavior has been performed in the past. Thus, when an unwanted behavior is often performed but in different locations, it is a weak habit in Quinn et al.'s (2010) study but a strong one in ours, which may explain the differences.

We also examined the choice of specific self-control strategies when facing habits and tempting desires in order to clarify whether participants choose the more effective strategy. Two of our findings of these analyses are noteworthy: First, participants chose stimulus control more likely with increasing habit strength, which was in line with the decreased effectiveness of stimulus control for strong habits. Moreover, although the success of monitoring and distraction was not influenced by habit strength, participants chose them more (monitoring) or less likely (distraction), which is in line with the theoretically predicted effectiveness. Thus, since continually monitoring of persistent habits is wearisome (Baumeister, Vohs, & Tice, 2007) and may increase the likelihood for indulgence (Hofmann, Vohs, & Baumeister, 2012), individuals might switch to a less effortful strategy at some point in order to avoid the strenuous monitoring.

Second, desire strength did not influence the choice of a particular self-control strategy alone. Only the interaction with self-control success was significant: More effective regulators (high self-control success) were not more likely to pick the (theoretically) more effective strategy. Less effective regulators chose monitoring and distraction more likely when a different would be more effective. Third, although distraction was more successful under positive affect, participants chose it less likely with increasing valence. To sum up, based on our results it seems that individuals may not necessarily recognize the optimal strategy since even individuals who reported higher self-control success did not systematically choose the self-control strategies that were most effective given our data regarding affect, desire and habit strength. Thus, future research clearly should include an in-depth examination of predictors and concomitants of the choice of particular self-control strategies in daily life.

There are a number of limitations of the present research. First, people were asked to remember how they felt *before* enacting the control strategy. It is possible that people might misremember how they felt to be in line with their current feelings (Ross & Conway, 1986). However, the daily diary was completed relatively early after the events' occurrence which should reduce a potential retrospective memory bias. We also found interaction effects and differences in marginal means for negative and positive affect and self-control strategies, which cannot be simply explained by a retrospective memory bias. Finally, we also assessed affect *after* enacting the control strategy and controlled for its influences on daily self-control success.

Second, we assessed requirements of stability and flexibility only indirectly by assessing self-control strategies that differently affect the stability-flexibility-balance. Future research could study the influence on this balance more directly, for example by using an Ambulatory Assessment design with several prompts per day that assess current demands for stability or flexibility.

Third, we adopted the measures of habit and temptation strength from Quinn and colleagues (2010). There, temptation strength is a measure of the strength of desires (in conflict), in that it assesses the positive affective consequences only of enacting a strong desire (feeling better). However, the measure fails to capture negative affective consequences such as guilt (Hofmann & Fisher, 2012) that may arise since giving in to desires may also jeopardize a long-term goal. Thus, future research should improve measurement of temptation strength by either separate desire strength from conflict strength (Hofmann, Baumeister, Förster, & Vohs, 2012) or assess temptation strength more directly.

Fourth, we did not distinguish between different facets of positive affect, such as motivational intensity, that might differently influence cognitive flexibility and, thus, self-control. According to Harmon-Jones, Gable, and Price (2012), high motivationally intensive positive affect before goal attainment (e.g., enthusiasm) is associated with narrowed attention in order to foster the completion of that goal. Conversely, low motivationally intensive positive affect after goal attainment (e.g., satisfaction) is associated with broadened attention in order to facilitate one's search for new goal opportunities. However, temptations are high in approach motivation and our results suggest that positive affect rather helps than hinders the success of distraction in regulating strong temptations. Nevertheless, this distinction offers a promising future research avenue to investigate the influence of positive affect on self-control in more detail.

Fifth, the correlational design cannot provide causal evidence for the effect of positive affect and, thus, cannot rule out alternative explanations like affect covarying with other factors possibly affecting self-control. While our data cannot replace experimental studies about the flexibility-enhancing effect of positive affect (e.g. Dreisbach & Goschke, 2004; Wenzel, Conner & Kubiak, 2013), it demonstrates its potential in explaining the role of affect on self-control efforts in naturalistic situations and corroborates the experimental evidence mentioned above with ecologically valid data.

To conclude, the impact of affective experience on self-control depends on the chosen self-control strategy. We found that self-control success is higher under positive affect when flexibility is required. Feeling good helps people to distract themselves from strong tempting desires and focus on something new by fostering the required flexibility. Therefore, when trying to distract yourself from that tempting donut, it may be better to do it when feeling good. Otherwise it may be better to avoid the presence of this tempting treat.

Chapter 4:

Internal Validity of Limited Self-Control in a Laboratory Research Setting[3]

4.1 Introduction

Self-control is usually framed as an individual's ability to control his or her desires and behaviors in the pursuit of longer-term goals (Muraven & Baumeister, 2000). The strength model of self-control (SMSC), which sets out to explain self-control and its failures, has been a popular theory in the field for the past two decades (e.g., Baumeister, 2001). In this model, an individual's capacity to exert self-control is considered a limited resource that acts of self-control can draw from. When this resource is depleted, people are thought of to be in a state of *ego depletion*, which results in a decline in performance in subsequent acts of self-control.

Over the past two decades, researchers have collected a large body of evidence for the ego depletion effect: Hagger, Wood, Stiff, and Chatzisarantis (2010) conducted a meta-analysis of 83 studies and reported a medium-to-large pooled effect size of $d = 0.62$ for the ego depletion effect. Although an impressive amount of evidence has been amassed so far, the notion of ego depletion and its supporting evidence have recently been challenged. Carter and McCollough (2013; 2014) reanalyzed the studies used in the meta-analysis by Hagger and colleagues (2010) and found evidence for small-study effects such as publication bias. By applying two bias-correction methods, they found average effect sizes of $d = 0.25$ and $d = -0.10$, respectively. In addition, Carter, Kofler, Forster, and McCollough (2015) conducted a meta-analysis in which they included non-published effect sizes and excluded studies with tasks that did not directly relate to self-control (e.g., imagining cheating on a hypothetical

[3] This study is currently submitted to a peer-reviewed journal: Wenzel, M., Lind, M., Rowland, Z., Zahn, D., & Kubiak, T. (2016). *The limits of ego depletion: A crossover study on the robustness of performance deterioration in consecutive tasks.* Manuscript submitted for publication.

romantic partner). They reported a pooled effect size of Hedge's $g = 0.43$. This effect, again, was reduced to zero when small-study effects were accounted for. Finally, in a multi-lab pre-registered replication study of the ego-depletion effect, Hagger and colleagues (2016) found a non-significant effect size of $d = 0.04$: Out of 23 participating labs with a total of 2141 participants, three labs found significant effect sizes, with one effect size in the opposite direction (i.e., better performance in the experimental condition).

Single experimental studies using the traditional dual-task paradigm to investigate ego depletion are often underpowered, with statistical powers ranging from 0.31 to 0.69 for the studies in the meta-analysis by Hagger and colleagues (2010; Carter & McCollough, 2014). One efficient way to increase statistical power in a single study is to build on within-subject designs: In a dual-task setup, a baseline measure equivalent to the second task can be employed to, then, compare performance in the second task with the baseline performance of the respective participant (e.g., Xu et al., 2014). This reduces error variance due to controlled baseline differences, and thus increases statistical power. Controlling baseline differences also has an additional advantage, as Lange and Eggert (2014) demonstrated that randomization does not guarantee the elimination of potential confounding variables (Krause & Howard, 2003). Lange and Eggert (2014) found significant performance differences in their self-control tasks; while these differences could have been misattributed to the experimental manipulation, but were instead the result of baseline performance differences. This could also partially explain the substantial between-laboratory variance found in the multi-lab replication project (Hagger et al., 2016) for which all study centers used the same instructions and materials.

While these studies demonstrate that assessing baseline performance is important, there are three challenges to assessing baseline performance: First, the wash-out period between the baseline and the two experimental tasks is usually less than an hour; this can lead to practice effects, which may, in turn, obscure possible ego depletion effects. Second, it is unclear how performing a baseline task that requires self-control affects performance in the dual-task paradigm (i.e., beyond practice) because studies varying the number of initial tasks are scarce. Converse and DeShon (2009) observed an ego depletion effect when participants in the experimental condition performed one depleting task, but not if they performed two depleting tasks. Third, assessing baseline performance only allows for within-subject analysis within a specific condition (e.g., experimental condition). Between-subject variance is therefore still a factor, since performance in the experimental condition is aggregated across participants and then compared with aggregated performance in the control condition.

We contend that the strongest within-subjects test of ego depletion would be to compare whether the same person performs worse after initial depletion in the experimental condition compared to no depletion in the control condition. We therefore propose a crossover design to reduce the problems associated with the assessment of baseline performance. In a crossover study, all participants go through both study conditions, separated by a wash-out period to reduce carry-over effects. Half of the participants start in the experimental condition and then cross over to the control condition, while the other half start with the control condition and switch to the experimental condition. This design not only allows for a longer wash-out period, reducing possible effects due to practice or the experimental setup, but also allows researchers to quantify performance differences between the experimental and control conditions for each individual. This measure can therefore be used to (a) assess the variability of the ego depletion effect, answering the question of "How many individuals show ego depletion effects?" and (b) investigate possible moderators, answering the question of "Why do some individuals not show any ego depletion effects?"

In addition to the issues of statistical power and baseline performance impact, there are several other problems in existing ego depletion research. The first problem is that the researcher generally makes largely arbitrary choices in selecting outcome measures and analytic approaches, which may impact the results (Simmons, Nelson, & Simonsohn, 2011). A good example of this is in the analysis of the Stroop task, a popular second task in the ego depletion literature (Hagger et al., 2010). At least five dependent variables are commonly used or reported in the Stroop task and in related tasks: (1) Mean reaction time on incongruent trials (e.g., Muraven, Shmueli, & Burkley, 2006), (2) difference between mean reaction time on incongruent and congruent trials (e.g., Wenzel, Conner, & Kubiak, 2013), (3) a ratio score $(M_{RT\_incongruent} - M_{RT\_incongruent}) / (M_{RT\_incongruent} + M_{RT\_incongruent})$ (e.g., Inzlicht & Gutsell, 2007), (4) reaction time variability on incongruent trials (e.g., Sripada, Kessler, & Jonides, 2014; Hagger et al., 2016), and (5) number of errors participants made (e.g., Job, Dweck, & Walton, 2010). Since it is unclear which measure is best suited to assessing self-control performance, as well as how these measures correlate with each other, research is needed to clarify whether the wide range of researcher methodological choices may explain the heterogeneous results reported in the ego depletion literature.

Another problem is that the manipulation checks that are conducted to check whether individuals are initially depleted are rarely validated. Manipulation checks often build on scales that comprise a seemingly ad-hoc selection of items assessing effort, perceived difficulty, or fatigue (Hagger et al., 2010), making it unclear to what degree each of these concepts measure self-control depletion. To the best of our knowledge, there is only one scale that was directly constructed and validated to assess self-control depletion: the State Self-Control Capacity Scale (SSCCS; Ciarocco, Twenge, Muraven, & Tice, 2004). The SSCCS was translated into German by Bertrams, Unger, and Dickhäuser (2011) and showed good reliability and validity. However, assessing self-control depletion may trigger or induce implicit theories about willpower that have been shown to moderate the ego depletion effect. For example, Job et al. (2010) found that the ego depletion effect only manifested when participants completed six items that indicated limited self-control capacity (e.g., “After a strenuous mental activity your energy is depleted and you must rest to get it refueled again"). In contrast, this effect was not observable when the items indicated non-limited self-control capacity ("Your mental stamina fuels itself; even after strenuous mental exertion you can continue doing more of it"). More research is clearly needed to test whether assessing self-control depletion directly as a manipulation check itself induces ego depletion.

4.1.1 The Present Research

In the present paper, we discuss a crossover study in which all participants completed both the experimental and the control condition in a dual-task paradigm with a wash-out period of two weeks. This design enables to test for ego depletion with high statistical power and to determine the existence of any within-subjects effects, thereby allowing us to control for sequence effects and to quantify inter- and intraindividual variability of the ego depletion effect. However, since recent evidence regarding the ego depletion effect is heterogeneous, we wanted to examine to what extent this heterogeneity may be caused by the range of ways researchers analyze data. To that end, we compare the results of the various indicators of self-control in the second task of the dual-task paradigm. Finally, we wanted to test whether the SSCCS can serve as a manipulation check of ego depletion without inducing this state.

4.2 Method

4.2.1 Participants

To achieve a 95% power to detect an effect size equivalent to that reported by Carter et al. (2014) (i.e., $d = 0.25$, $f = 0.125$) with $\alpha = .05$, a conservative correlation among the repeated measures of .60, and an attrition rate of no more than 10%, we recruited a total of 192 participants through bulletins, social networks, flyers, mailing lists, and direct approaches.

Each participant received partial course credit, which was not contingent on performance, as compensation for participating. Additionally, participants who completed the whole experiment had the chance to win one of 13 5€ vouchers (approx. 6 US$) for a local ice cream parlor. None of the participants had to be excluded due to performance issues ($z > 3$) or due to having an error rate of over 33% in the first or second task. However, one participant was excluded due to technical problems. The final sample therefore consisted of 191 students (170 females, age $M = 23.0$ years, $SD = 5.3$).

4.2.2 Design

This study was a randomized, 2 x 2 x 2 crossover study, with two conditions for the within-subject factor of ego depletion (depletion vs. control condition), the between-subject factor of manipulation check (SSCCS assessment vs. no SSCCS assessment condition), and the within-subject factor of period (first vs. second dual-task trial). Participants were randomly assigned to the experimental conditions with $n_{\text{depletion condition first}} = 47$ and $n_{\text{depletion condition second}} = 46$ for the SSCCS condition and $n_{\text{depletion condition first}} = 50$ and $n_{\text{depletion condition second}} = 48$ for the no SSCCS condition.

4.2.3 Procedure

Participants signed up to take part in a study testing various cognitive tasks. In the first lab session, participants received information about the study, signed an informed consent form, and indicated their baseline affect. Participants then completed a computer version (Sripada et al., 2014) of the crossing-out letters task (Baumeister, Bratslavsky, Muraven, & Tice, 1998) as their first task. This task either required self-control (complex rules condition) or not (simple rules condition). After receiving the instructions but before completing the test, participants answered a questionnaire on task motivation. After completing the task, participants rated their affect again, as well as the perceived difficulty of the task. Participants in the SSCCS condition also rated their capacity for self-control while participants in the no SSCCS condition rated their affect, using a parallel form of the affect questionnaire that they had already completed. This parallel form had a similar number of items compared to the SSCCS, to ensure equal duration between the two groups. Participants then completed the second task, which was the same for all conditions, and answered the task motivation questionnaire before completing the task again. They ended by rating their affect for a third time and rating the difficulty of the second task. After finishing the first lab session, participants were instructed to complete an online questionnaire before participating in the second lab session two weeks later. The procedure for the second lab session mirrored that of the first lab session, except that participants who had previously been in the depletion condition were assigned to the control condition and vice versa.

4.2.4 Materials

4.2.4.1 First task: Letter e-task

We opted for a computer version (Sripada et al., 2014) of the letter-e task (Baumeister et al., 1998) since Hagger and colleagues (2010) reported in their meta-analysis that crossing out letters as the first task yielded the highest ego depletion effect sizes. In the letter-e task, single words were consecutively displayed and participants were instructed to press the letter g on the keyboard as quickly and accurately as possible when the letter e was displayed. Participants in the depletion condition received an additional instruction to refrain from pressing the letter g when the letter e was next to or one letter away from a vowel. This required more self-control, since participants in the experimental (i.e., depletion) condition could not react to every e, but instead needed to inhibit their impulse when the restrictive rule condition was met. The main session consisted of 50 practice trials and 150 recorded trials for seven and a half minutes. In the practice trials, all participants from both groups were instructed to react to each letter e by pressing the letter g. This established a behavioral pattern in the experimental group that had to be suppressed in those individuals who received additional instructions. Practice trials were excluded from the computation of mean reaction time and number of errors made.

4.2.4.2 Second task: Multi-source interference task (MSIT)

For the second task, we used the MSIT (Bush, Shin, Holmes, Rosen, & Vogt, 2003), which has previously been employed in ego depletion research (Shamosh & Gray, 2007; Sripada et al., 2014; Wenzel, Kubiak, & Conner, 2014) and which has shown large interference effects. In the MSIT, in the present study, three characters are displayed as a combination of the characters 1, 2, 3, and x. Participants are told to indicate as quickly and as accurately as possible the character that differed from the other identical ones, regardless of its position, by pressing the corresponding key on the keyboard. In control trials, the target is (a) always a number, (b) always on the corresponding place (target '1' on the first place in '1xx'), and (c) always larger than the distractor; finally, (d) both distractors are always the letter "x." Thus, the congruent trials are 1xx, x2x, and xx3. In interference trials, the target (a) is sometimes larger or smaller in size on the screen, (b) does not always match its position, and (c) distractors are always other numbers. This study used 10 sets of incongruent trials (112, 131, 313, 221, 232, 233, 311, 322, 331, and 332). Participants completed 150 trials (50 congruent and 100 incongruent trials) in a fixed, interspersed order to hold trial-to-trial-adaptations constant across participants (Lorist & Jolij, 2012).

Stimuli were presented in the center of a computer screen positioned 90 cm in front of participants' eyes. Characters were printed in white (Font: Arial; Size: 18 and 24pt; Style: bold) on a black background. Before each trial, a white plus sign was displayed for 500 ms (to ensure that participants were focusing on the correct point); participants then had two seconds to complete each trial before hitting the response deadline. The first trial (0.67% of all trials), error trials (4.64%), trials faster than 200 ms (0.01%), and any trials with a reaction time (RT) above or below 3 *SD*s from each individual's mean RT (1.34%) were excluded from all subsequent analyses. Applying these rules excluded 6.65% of the total data.

Since the MSIT is a Stroop-like cognitive interference task, the following five measures (as outline in the introduction) can be derived from the MSIT, and served as indicators of self-control performance in the present study: (1) mean reaction time on incongruent trials, (2) difference between mean reaction time on incongruent and congruent trials, (3) ratio score ($M_{RT\_incongruent}$ - $M_{RT\_incongruent}$)/ ($M_{RT\_incongruent}$ + $M_{RT\_incongruent}$), (4) reaction time variability (RTV) on incongruent trials (i.e., the sum of the sigma and tau variability parameters using ex-Gaussian modeling; Sripada et al., 2014), and (5) the number of errors. Since all measures were positively skewed, we took the logarithm of them to achieve a normal distribution; the only exception was number of errors, for which we took the square root. Lower values indicate better self-control performance for all measures.

4.2.4.3 Additional dependent variables

4.2.4.3.1 Manipulation checks

Using the categories established by Hagger and colleagues (2010), after each task in the present study participants in all conditions completed self-report ratings regarding the difficulty (3 items: "difficult," "hard," "easy" (reversed)), required effort (3 items: "effortful," "laborious," "undemanding" (reversed)), and fatiguing nature (3 items: "depleting, "exhausting," "refreshing" (reversed)) of the task they had just completed. Each item was scored on a Likert scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). Factor analysis revealed only two factors with Eigenvalues larger than 1: The items "difficult," "hard," "easy," and "undemanding" all loaded onto one factor (Factor 1: difficulty), while the items "depleting," "exhausting," "refreshing," and "laborious" loaded onto the second factor (Factor 2: fatigue); "effortful" loaded onto both factors. After identifying these two factors we computed mean scores for all of the single-loading items, with higher scores indicating higher task difficulty or higher fatigue, respectively. Both scales correlated moderately with each other, with $r = .45$ for the first task and .26 for the second one. Although substantial, these correlations indicate a shared variance of only 7% to 20%. In classical test theory, reliability coefficients such as Cronbach's alpha are defined as the proportion of the variance of a measure that can be attributed to individual differences in the true score. However, in repeated measures, the random measurement error is affected by both variations between and within individuals; measures can therefore vary in how well they capture those variations. To separate the reliabilities between and within individuals, we did not compute Cronbach's alpha either for each measurement point or across them; rather, we used the formula provided by Shrout and Lane (2012). The between-person reliabilities for task difficulty ($R_{KRN} = .65$) and fatigue ($R_{KRN} = .77$), as well as the within-person reliabilities for changes from time point to time point ($R_{CN} = .82$ and $R_{CN} = .69$, respectively), were acceptable to good.

Participants in the SSCCS condition also completed the validated German short-version (Bertrams et al., 2011) of the State Self-Control Capacity Scale (SSCCS; Ciarocco et al., 2004). The scale comprises 10 items measuring one's capacity to regulate oneself in that particular moment (e.g., "I feel like my willpower is gone") on a Likert-type scale ranging from 1 (*not true*) to 7 (*very true*). We inverted this scale, so that higher scores reflected higher self-control depletion (i.e., less self-control). The scale showed good between- and within-person reliabilities ($R_{KRN} = .83$ and $R_{CN} = .84$, respectively). Participants in the no-SSCCS condition completed a 12-item parallel form of the affect questionnaire, which is described in the following paragraph.

4.2.4.3.2 Affect

Affect was measured before and after the first task and a third time after the second task, using the short version of the German Multidimensional Mood Questionnaire (MDMQ; Steyer, Schwenkmezger, Notz, & Eid, 1997). This instrument assesses momentary affect on three dimensions: valence (items: "good," "bad"), energetic arousal (items: "awake," "tired"), and calmness (items: "nervous," "calm"). Mean scores for agreement with these adjectives on a Likert-type scale ranging from 1 (*not at all*) to 5 (*extremely*) were computed for each dimension, with higher scores representing a stronger valence (R_{KRN} = .83, R_{CN} = .78), more energy (R_{KRN} = .85, R_{CN} = .75), or more calmness (R_{KRN} = .88, R_{CN} = .73).

4.2.4.3.3 Motivation

Since motivation is currently a popular explanation for the reduced performance in the second task of a dual-task setup (e.g., Inzlicht & Schmeichel, 2012), participants were asked to complete the short form of the Questionnaire for Assessing Momentary Motivation to assess task-related motivation (QAM; Rheinberg, Vollmeyer, & Burns, 2001; Freund, Kuhn, & Holling, 2011). This questionnaire uses 12 items on a Likert-type scale ranging from 1 (*not correct at all*) to 7 (*fully correct*) to measure four motivational factors in learning or achievement situations: probability of success, anxiety, challenge, and interest. Anxiety (R_{KRN} = .92), challenge (R_{KRN} = .87), and interest (R_{KRN} = .92) showed very good between-person reliabilities, whereas probability of success was only acceptable (R_{KRN} = .61). The within-person reliabilities were also only acceptable, ranging from R_{CN} = .56 to R_{CN} = .58; the only exception was challenge (R_{CN} = .38).[4]

4.2.5 Statistical Analyses

To detect the typical ego depletion effect—i.e., comparing the depletion and control conditions at one point in time—and to investigate the influence of the SSCCS assessment on this effect, we computed a MANOVA on the five MSIT measures of the first lab session. We included the two factors of ego depletion and SSCCS assessment and controlled for gender and age. Each dependent variable was rescaled as a percentage of the empirical maximum score for that particular measure. This procedure maintains the distribution of the variables while making the results easier to interpret and compare with each other. We then computed separate ANOVAs to disentangle the effects of the experimental manipulation and the control variables on the dependent variables.

We then analyzed the full data set of the crossover design. Since the full data set consists of multiple dependent variables and points in time, we computed hierarchical multivariate related outcomes multilevel models to estimate and compare the effects of our experimental manipulations on the dependent variables in a common model (Baldwin, Imel, Braithwaite, & Atkins, 2014). This allowed us to test the differential effects of an experimental manipulation with increased power. To build these models we first z-standardized the dependent measures to obtain a comparable scale. We then created four indicator variables for reaction time on incongruent trials, ratio score, RTV, and number of errors; we excluded the difference score due to its high correlation with the ratio score, with a goal of limiting multicollinearity. This resulted in a 3-level model in which dependent variables (Level 1) are nested within time

[4] Please note that the within-person reliability is more sensitive to the number of items, which may explain the lower values.

points (Level 2), which are nested within participants (Level 3). We then computed hierarchical multivariate multilevel models. First we entered the four indicator variables, which served as the random intercepts in the model, and included their two-way interactions with ego depletion, SSCCS assessment, and period. We also included the control variables of gender and age to obtain the main effect of each predictor for each dependent variable of the MSIT (Equation 1):

$$\begin{aligned} y_{hij} = {} & \beta_{10}l_j + \beta_{20}r_j + \beta_{30}v_j + \beta_{40}e_j + \beta_{11}Period_{ij}l_j + \beta_{21}Period_{ij}r_j + \beta_{31}Period_{ij}v_j \\ & + \beta_{41}Period_{ij}e_j + \beta_{12}ED_jl_j + \beta_{22}ED_jr_j + \beta_{32}ED_jv_j + \beta_{42}ED_je_j \\ & + \beta_{13}SSCCS_jl_j + \beta_{23}SSCCS_jr_j + \beta_{33}SSCCSv_j + \beta_{43}SSCCS_je_j + \beta_{14}Gender_jl_j \\ & + \beta_{24}Gender_jr_j + \beta_{34}Gender_jv_j + \beta_{44}Gender_je_j + \beta_{15}Age_jl_j + \beta_{25}Age_jr_j \\ & + \beta_{35}Age_jv_j + \beta_{45}Age_je_j + \mu_{1j}l_j + \mu_{2j}r_j + \mu_{3j}v_j + \mu_{4j}e_j + \vartheta_{1j}Period_{ij}l_j \\ & + \vartheta_{2j}Period_{ij}r_j + \vartheta_{3j}Period_{ij}v_j + \vartheta_{4j}Period_{ij}e_j + \varepsilon_{1ij}l_j + \varepsilon_{2ij}r_j + \varepsilon_{3ij}v_j \\ & + \varepsilon_{4ij}e_j \end{aligned} \tag{1}$$

where l, r, v, and e are the four indicator variables of the MSIT, with l = latency on incongruent trials, r = ratio score, v = RTV, and e = number of errors. β indicates the fixed effects, ν the random effects, and ε the error component. ED indicates the ego depletion manipulation, SSCCS the SSCCS assessment manipulation, and Period the lab session.

In the second step, we entered the three-way interactions (bolded in Equation 2), which reflect the two-way interactions of time, ego depletion manipulation, and SSCCS assessment manipulation on each dependent variable (Equation 2):

$$\begin{aligned} y_{hij} = {} & \beta_{10}l_j + \beta_{20}r_j + \beta_{30}v_j + \beta_{40}e_j + \beta_{11}Period_{ij}l_j + \beta_{21}Period_{ij}r_j + \beta_{31}Period_{ij}v_j \\ & + \beta_{41}Period_{ij}e_j + \beta_{12}ED_jl_j + \beta_{22}ED_jr_j + \beta_{32}ED_jv_j + \beta_{42}ED_je_j \\ & + \beta_{13}SSCCS_jl_j + \beta_{23}SSCCS_jr_j + \beta_{33}SSCCSv_j + \beta_{43}SSCCS_je_j + \beta_{14}Gender_jl_j \\ & + \beta_{24}Gender_jr_j + \beta_{34}Gender_jv_j + \beta_{44}Gender_je_j + \beta_{15}Age_jl_j + \beta_{25}Age_jr_j \\ & + \beta_{35}Age_jv_j + \beta_{45}Age_je_j + \boldsymbol{\beta_{16}Period_{ij}ED_jl_j} + \boldsymbol{\beta_{26}Period_{ij}ED_jr_j} \\ & + \boldsymbol{\beta_{36}Period_{ij}ED_jv_j} + \boldsymbol{\beta_{46}Period_{ij}ED_je_j} + \boldsymbol{\beta_{17}Period_{ij}SSCCS_jl_j} \\ & + \boldsymbol{\beta_{27}Period_{ij}SSCCS_jr_j} + \boldsymbol{\beta_{37}Period_{ij}SSCCS_jv_j} \\ & + \boldsymbol{\beta_{47}Period_{ij}SSCCS_je_j} + \boldsymbol{\beta_{18}ED_jSSCCS_jl_j} + \boldsymbol{\beta_{28}ED_jSSCCS_jr_j} \\ & + \boldsymbol{\beta_{38}ED_jSSCCS_jv_j} + \boldsymbol{\beta_{48}ED_jSSCCS_je_j} + \mu_{1j}l_j + \mu_{2j}r_j + \mu_{3j}v_j + \mu_{4j}e_j \\ & + \vartheta_{1j}Period_{ij}l_j + \vartheta_{2j}Period_{ij}r_j + \vartheta_{3j}Period_{ij}v_j + \vartheta_{4j}Period_{ij}e_j + \varepsilon_{1ij}l_j \\ & + \varepsilon_{2ij}r_j + \varepsilon_{3ij}v_j + \varepsilon_{4ij}e_j \end{aligned} \tag{2}$$

In the third and final step, we added the four four-way interactions (bolded in Equation 3):

$$
\begin{aligned}
y_{hij} = {} & \beta_{10}l_j + \beta_{20}r_j + \beta_{30}v_j + \beta_{40}e_j + \beta_{11}Period_{ij}l_j + \beta_{21}Period_{ij}r_j + \beta_{31}Period_{ij}v_j \\
& + \beta_{41}Period_{ij}e_j + \beta_{12}ED_jl_j + \beta_{22}ED_jr_j + \beta_{32}ED_jv_j + \beta_{42}ED_je_j \\
& + \beta_{13}SSCCS_jl_j + \beta_{23}SSCCS_jr_j + \beta_{33}SSCCSv_j + \beta_{43}SSCCS_je_j + \beta_{14}Gender_jl_j \\
& + \beta_{24}Gender_jr_j + \beta_{34}Gender_jv_j + \beta_{44}Gender_je_j + \beta_{15}Age_jl_j + \beta_{25}Age_jr_j \\
& + \beta_{35}Age_jv_j + \beta_{45}Age_je_j + \beta_{16}Period_{ij}ED_jl_j + \beta_{26}Period_{ij}ED_jr_j \\
& + \beta_{36}Period_{ij}ED_jv_j + \beta_{46}Period_{ij}ED_je_j + \beta_{17}Period_{ij}SSCCS_jl_j \\
& + \beta_{27}Period_{ij}SSCCS_jr_j + \beta_{37}Period_{ij}SSCCS_jv_j + \beta_{47}Period_{ij}SSCCS_je_j \\
& + \beta_{18}ED_jSSCCS_jl_j + \beta_{28}ED_jSSCCS_jr_j + \beta_{38}ED_jSSCCS_jv_j + \beta_{48}ED_jSSCCS_je_j \\
& + \boldsymbol{\beta_{19}Period_{ij}ED_jSSCCS_jl_j} + \boldsymbol{\beta_{29}Period_{ij}ED_jSSCCS_jr_j} \\
& + \boldsymbol{\beta_{39}Period_{ij}ED_jSSCCS_jv_j} + \boldsymbol{\beta_{49}Period_{ij}ED_jSSCCS_je_j} + \mu_{1j}l_j + \mu_{2j}r_j \\
& + \mu_{3j}v_j + \mu_{4j}e_j + \vartheta_{1j}Period_{ij}l_j + \vartheta_{2j}Period_{ij}r_j + \vartheta_{3j}Period_{ij}v_j \\
& + \vartheta_{4j}Period_{ij}e_j + \varepsilon_{1ij}l_j + \varepsilon_{2ij}r_j + \varepsilon_{3ij}v_j + \varepsilon_{4ij}e_j
\end{aligned}
\tag{3}
$$

4.3 Results

We first investigated how our ego depletion manipulation influenced the additional dependent variables, including the manipulation checks of ego depletion. To this end, we computed paired *t*-tests with the ego depletion factor as the independent variable, as well as dependency-corrected effect sizes (Morris & DeShon, 2002). As indicated in Table 1, participants reported that the first task in the depletion condition was more difficult and fatiguing than the first task in the control condition. Both effect sizes were large (Cohen, 1988) and the effect size of difficulty was nearly twice as large as that of fatigue, indicating that participants experienced the first task in the depletion condition as more difficult than fatiguing. However, when using SSCCS to assess self-control depletion, the effect was small and only marginally significant after applying the Šidák correction (Šidák, 1967). For the second task, the depletion factor did not reach significance for either fatigue or self-control depletion, but did reach significance for difficulty, with participants reporting that the second task was slightly more difficult in the control than in the depletion condition. However, this effect fell short of significance and was only small in size.

None of the affect scales reached significance for the first or second task when the results were Šidák corrected. The valence scale was almost marginally significant in the first task: Participants indicated that their affect was worse when performing the more difficult version of the letter-e task as their first task compared to the easier version.

For task motivation, participants in the depletion condition reported being less likely to perform well in the first task and being more afraid to fail than participants in the control condition. However, participants in both conditions were similar with respect to the challenge or interest scale of the QCM, appearing equally internally motivated to perform the first task. Our ego depletion manipulation did not influence any scale of the QCM in the second task, indicating that performing the more difficult and fatiguing first task did not influence motivation on the second task.

Table 7

Means, standard deviations, and inferential statistics of the additional dependent variables.

		First Task								*Second Task*							
		DC		*CC*						*DC*		*CC*					
	n	*M*	*SD*	*M*	*SD*	*r*	*t*	*p*	*d'*	*M*	*SD*	*M*	*SD*	*r*	*t*	*p*	*d'*
1. Difficulty	188	3.29	0.84	1.89	0.64	.31	21.57	**<.001**	1.60	2.65	0.88	2.89	0.91	.42	-3.52	**<.001**	-0.25
2. Fatigue	188	3.36	0.74	2.68	0.80	.44	11.51	**<.001**	0.84	3.02	0.91	3.08	0.86	.52	-0.94	.349	-0.07
3. Self-control depletion	93	3.35	0.94	3.04	0.95	.33	2.75	**.007**	0.28	3.28	1.02	3.26	1.09	.55	0.22	.830	0.02
4. MDMQ Valence	188	3.50	0.66	3.66	0.69	.30	-2.65	.009	-0.20	3.93	0.69	3.91	0.76	.38	0.46	.645	0.03
5. MDMQ Energetic Arousal	188	3.00	0.84	3.12	0.88	.39	-1.68	.095	-0.13	2.91	0.85	2.94	0.88	.40	-0.41	.680	-0.03
6. MDMQ Calmness	188	3.76	0.77	3.83	0.80	.42	-1.19	.237	-0.08	3.79	0.77	3.76	0.82	.53	0.53	.596	0.04
7. QCM Probability of success	188	4.95	1.08	6.05	0.80	.27	-13.16	**<.001**	-0.99	2.36	0.80	2.24	0.81	.50	2.07	.040	0.15
8. QCM Anxiety	188	3.11	1.45	2.72	1.42	.70	4.70	**<.001**	0.35	2.68	1.44	2.63	1.53	.74	0.62	.538	0.05
9. QCM Challenge	188	4.42	1.08	4.24	1.16	.65	2.51	.013	0.19	4.03	1.09	4.17	1.18	.75	-2.28	.024	-0.17
10. QCM Interest	188	3.07	1.35	3.13	1.28	.72	-0.80	.426	-0.06	3.01	1.43	3.00	1.33	.76	0.25	.802	0.01

Notes. First/Second Task = first / second task in the dual-task paradigm. DC = depletion condition; CC = control condition; MDMQ = Multidimensional Mood Questionnaire; QCM = Questionnaire on Current Motivation. P-values in bold are significant at the Šidák-corrected significance level of $\alpha = .0081$. 'The effect size is dependency-corrected (Morris & DeShon, 2002).

4.3.1 Robustness of the Ego Depletion Effect across Outcomes

4.3.1.1 Between-subject analyses

We assessed the robustness of the ego depletion effect using the various dependent measures of the MSIT, first focusing on the typical ego depletion effect by comparing performance in the depletion and control conditions at Time 1. To do this, we computed a MANOVA on the five dependent variables. As indicated in Table 2 (second column), this multivariate model indicated a substantial difference in the MSIT measures between the experimental conditions. However, the effects for the ego depletion manipulation and its interaction with SSCCS assessment did not achieve statistical significance. Moreover, the control variable of gender was highly significant. We therefore computed separate ANOVAs (see Table 2). To make the results more comparable, each dependent variable was rescaled as a percentage of the empirical maximum score for that particular measure, which did not alter the distribution of the variables. Out of the five measures, ego depletion exhibited a significant effect only on the MSIT difference and ratio score (without the Šidák correction), but in the opposite direction as was hypothesized by the SMSC: Participants in the depletion condition showed smaller, not larger, difference (M_{diff} = .42, SE_{diff} = .02) and ratio scores (M_{quot} = .44, SE_{quot} = .02) than controls (M_{diff} = .47, SE_{diff} = .02; M_{quot} = .49, SE_{quot} = .02). Moreover, gender also exhibited a significant effect on those two measures, in that females (M_{diff} = .43, SE_{diff} = .01; M_{quot} = .45, SE_{quot} = .01) performed better than males (M_{diff} = .57, SE_{diff} = .03; M_{quot} = .58, SE_{quot} = .03). Females also performed better in the other three MSIT measures, but the effects were not statistically significant. These results therefore demonstrate null effects, or, in the case of the difference and ratio scores, an effect in the opposite direction of the ego depletion manipulation, with effect sizes converted to Cohen's d ranging from d = -.34 for the ratio score to d = .15 for the number of errors.

Table 8

Analysis of Variance (ANOVA) between ego depletion manipulation and SSCCS assessment.

	MANOVA			ANOVA														
				Mean RT (log)			Difference (log)			Ratio (log)			RTV (log)			No. of errors (squared)		
	Wilks λ	*F*	*p*	*F*	*p*	η_p^2	*F*	*p*	η_p^2	*F*	*p*	η_p^2	*F*	*p*	η_p^2	*F*	*p*	η_p^2
ED	0.93	2.23	.054	0.17	.677	.00	5.10	.025	.03	5.24	.023	.03	0.37	.544	.00	0.97	.327	.01
SSCCS	0.98	0.79	.556	2.17	.143	.01	0.09	.765	.00	0.05	.830	.00	2.56	.112	.01	0.78	.380	.00
ED x SSCCS	0.94	1.99	.083	3.35	.069	.02	8.53	**.004**	.05	9.21	**.003**	.05	1.30	.256	.01	3.14	.078	.02
Gender	0.87	4.52	<.001	2.80	.096	.02	15.54	**<.001**	.08	14.36	**<.001**	.07	0.27	.604	.00	3.73	.055	.02
Age	0.50	1.07	.307	8.02	**.005**	.04	1.27	.261	.01	1.78	.184	.01	2.90	.091	.02	1.81	.180	.01

Notes. ED = Ego depletion manipulation; SSCCS = State self-control capacity scale assessment. p-values in bold are significant at a Šidák-corrected level at α = .0102 for the ANOVA results.

4.3.1.2 Within-subject analyses

4.3.1.2.1 Results of the hierarchical multivariate related outcomes multilevel model

The results of the hierarchical multivariate related-outcomes multilevel models—which compare performance differences across the depletion and control conditions within subjects in a common model with the four dependent variables of the MSIT—are presented in Table 3. For ease of interpretation, we will report the non-transformed, non-standardized means of the five dependent variables in Table 4. As indicated in the second and third column of Table 3, the effect of the ego depletion manipulation on MSIT performance differed significantly across the MSIT measures: When the first task required the use of self-control, participants responded significantly slower on incongruent trials (unadjusted) compared to when the first task consisted of a task without any self-control demands; this indicates a drop in performance due to lower initial levels of self-control for the second task. However, participants also showed a significantly lower mean ratio score for depletion compared to the control condition, which indicates better performance due to exerting initial self-control. Since neither the mean RTV nor the number of errors differed in response to different ego depletion manipulations, these results demonstrate that not only are these differences only small in size, but that they point into different directions, demonstrating substantial heterogeneity in the results.

The outcomes were influenced by gender and age in distinct ways: While the ratio score was sensitive to gender, with females having a lower ratio on average (i.e., better performance), mean reaction time on incongruent trials was sensitive to the age of the participants, with reaction time increasing with increasing age. The number of errors and the RTV were not influenced by the two control variables.

Period had a significant negative effect on all four dependent variables, indicating a practice effect, with participants performing better on all outcomes in the second lab session (which took place two weeks after the first lab session). However, when two-way interactions were included in the second step, the interaction between period and ego depletion manipulation was not significant for any of the four dependent variables, indicating no evidence for a sequence effect in our crossover design. Practice effects were therefore, to sum up, independent of the order of the ego depletion manipulation conditions and were much larger in size than any of the main effects of the ego depletion manipulation.

Table 9

Hierarchical multivariate related outcomes multilevel model of period, ego depletion, and SSCCS assessment.

			Mean RT (log)			Ratio Score (log)			RTV (log)			No. of errors		
	χ^2	*p*	*b*	*z*	*p*	*b*	*z*	*p*	*b*	*z*	*p*	*b*	*z*	*p*
Step 1														
Period	164.9	**<.001**	-0.74	-19.62	**<.001**	0.25	3.80	**<.001**	-0.83	-12.44	**<.001**	-0.21	-3.02	**.003**
ED	9.08	.028	0.09	2.44	.015	-0.13	-2.04	.042	0.09	1.39	.166	0.08	1.22	.224
SSCCS	3.03	.385	0.03	0.25	.803	0.18	1.43	.153	0.13	1.20	.231	-0.03	-0.22	.824
Gender	6.32	.097	-0.23	-1.19	.234	-0.57	-2.87	**.004**	-0.04	-0.21	.831	-0.17	-0.84	.399
Age	10.90	**.012**	0.04	3.25	**.001**	-0.01	-1.10	.271	0.02	2.28	.022	-0.01	-1.13	.259
Step 2														
Period x ED	7.89	.048	0.04	0.17	.864	0.40	1.60	.110	0.38	1.74	.083	-0.15	-0.59	.558
Period x SSCCS	11.67	.001	-0.19	-2.55	**.011**	0.19	1.47	.142	0.05	0.35	.729	0.14	0.99	.321
ED x SSCCS	1.41	.703	0.07	0.98	.328	0.04	0.28	.781	0.16	1.22	.223	-0.05	-0.34	.735
Step 3														
Period x ED x SSCCS	12.76	.005	-0.87	-1.73	.084	1.63	3.37	**.001**	-0.33	-0.75	.454	0.82	1.62	.104
Intercept	4.29	.232	-0.30	-0.88	.378	0.68	1.96	.050	-0.10	-0.33	.744	0.50	1.37	.170

Notes. ED = Ego depletion manipulation; SSCCS = State self-control capacity scale assessment. p-values in bold are significant at a Šidák-corrected level at $\alpha = .0127$.

Table 10

Means, standard deviations, inferential statisitics, and intercorrelations of MSIT outcomes of the first dual-task session.

		DC		*CC*									
	n	*M*	*SD*	*M*	*SD*	*r*	*t*	*p*	*d*	1.	2.	3.	4.
1. Latency of incongruent trials	188	785.0	132.7	773.0	125.5	.60	1.42	.158	0.10	-			
2. Difference Score	188	166.8	57.4	171.4	56.7	.61	-1.25	.214	-0.09	.45***	-		
3. Ratio Score	188	.12	.04	.13	.04	.57	-2.00	.047	-0.15	-.02	.87***	-	
4. Reaction Time Variability	188	0.48	0.15	0.46	0.15	.19	1.32	.188	0.10	.69***	.34***	.02	-
5. Number of Errors	188	7.15	6.25	6.51	4.79	.55	1.63	.105	0.12	.00	.07	.08	.16**

Notes. DC = depletion condition of the ego depletion manipulation; CC = control condition of the ego depletion manipulation. p-values in bold are significant at the Šidák-corrected significance level of α = .010.
* = p < .05, ** = p < .01, *** = p < .001

4.3.1.2.2 Results of the hierarchical random-effects multilevel models

The results of our hierarchical multivariate related-outcomes multilevel models show the variability and complexity of analyzing reaction time data. Here, we want to propose the use of multilevel analysis of raw reaction-time data as an alternative analysis method in research on ego depletion (Baayen & Milin, 2010). To streamline the presentation of the results, we will report the advantages of this analysis in the discussion only.

We computed hierarchical random-effects multilevel models for reaction time and for errors in the MSIT. For reaction times, we first fitted an unconditional model without any predictor, in which we determined that 80.6 % of total variance in reaction time was attributable to within-person variability. We then included the main effects of period, ego depletion, and SSCCS assessment, as well as the control variables of gender, age, trial number, reaction time on prior trial, incongruent trial, congruency of the prior trial, and whether the current trials were correctly responded to or not in the first step and two- and three-way interactions of interest in the next steps. Finally, we tested whether treating the non-lagged Level 1 predictors as random improved the model fit, which was indeed the case ($\chi^2(4) = 2846.2$, $p < .001$); random effects were therefore allowed for the intercept and the slopes. The final model (Table 5) accounted for 31.3% of reaction-time variance in the MSIT trials.

As indicated in Table 5, we first included only the main effects of the experimental manipulations without controlling for other variables; this choice was made in light of the large number of control variables. While the ego depletion manipulation did not have any effect on reaction time, period did: Participants performed better in the second compared to the first dual-task session (which took place two weeks earlier), indicating practice effects. When the control variables were entered in the second step, both the practice effect and the ego depletion manipulation were significant: Participants in the depletion condition responded, on average, more slowly to incongruent trials (M = 741.2 ms, SE_M = 12.3) as compared to participants in the control condition (M = 730.5 ms, SE_M = 12.3), mirroring the results of the hierarchical multivariate related-outcomes multilevel model of the aggregated reaction time outcome. In the third step, we entered the two-way interactions between period and the experimental manipulations to test for sequence effects of our crossover design; as no such effects were found for ego depletion, the order with which the depletion and control conditions were presented within subjects did not affect ego depletion. The interaction between ego depletion and incongruent trials was also significant: The reaction time difference was slightly less pronounced in the depletion (z = 41.78, p < .001) compared to the control condition (z = 43.35, p < .001). However, this small difference in reaction times is not enough to assert that ego depletion is greater for incongruent trials.

We were intrigued to find that so many control variables were significant, a finding that replicates prior evidence in studies on cognitive tasks outside of research on ego depletion. First, trial number was negatively associated with reaction time, indicating that participants showed practice effects not only between the two periods but also within the task. Second, participants responded slower to incongruent than congruent trials, demonstrating the need for cognitive control in the face of the incompatible information presented in congruent trials. Third, trial-to-trial adaptation was also found in the MSIT, as reflected by the interaction between the incongruence of the current and the prior trial: While the incongruence of the prior trial did not influence reaction time on a current congruent trial (z = 1.18, p = .236), it had a highly significant effect on a current incongruent trial (z = 7.21 p < .001). Participants responded more quickly to the incongruent trials when the prior trial was incongruent (M = 792.3 ms, SE_M = 12.3) than when the prior trial was congruent (M = 807.4 ms, SE_M = 12.3). We also included the three-way interaction between this interaction and the ego depletion manipulation to test whether initial depletion may have influenced trial-to-trial adaption in the

MSIT, which was not the case, as exerting initial self-control before the second task did not significantly influence trial-to-trial adaptation in the second task.

For the model of the error rates, the same predictors were included as in the above description of the hierarchical multilevel random-effects logistic regression. These results did not substantially differ from the results of the reaction time model. Participants still showed improved performance within and between the dual-task paradigm, and participants in the depletion condition made more errors than participants in the control condition, although this difference was not significant.

Taken together, although our method of investigating ego depletion curtailed researcher methodological freedom and increased the number of variables influencing self-control performance, the results of our multilevel analysis of non-aggregated reaction time data replicated the results produced by a multivariate related-outcomes multilevel model, with neither producing convincing evidence for the ego depletion effect.

Table 11

Fixed effects estimates (top) and variance estimates (bottom) of reaction time as function of period, ego depletion, and SSCCS assessment.

					CI_{95}	
Fixed effects	Estimate	*SE*	*z*	*p*[a]	Lower	Upper
Step 1						
Period	-93.67	12.19	-7.68	<.001	-117.57	-69.77
ED	13.26	12.19	1.09	.277	-10.64	37.16
SSCCS assessment	7.28	12.19	0.60	.550	-16.61	31.18
Step 2						
Period	-76.13	4.35	-17.49	<.001	-85.66	-67.60
ED	10.65	4.34	2.46	.014	2.16	19.16
SSCCS assessment	-6.00	21.75	-0.28	.783	-48.62	36.63
Gender	-18.14	35.17	-0.52	.606	-87.08	50.80
Age	7.05	2.02	3.50	<.001	3.10	11.00
Trial number	-0.86	0.04	-23.31	<.001	-0.93	-0.78
Lagged: RT on prior trial	0.17	0.01	43.22	<.001	0.17	0.18
Incongruent trial	187.85	4.11	45.74	<.001	179.80	195.90
Lagged: Prior trial incongruent	-10.23	1.87	-5.48	<.001	-13.89	-6.57
Correct trial	-35.54	12.68	-2.80	.005	-60.39	-10.69
Step 3						
Period x ED	7.79	43.58	0.18	.858	-77.63	93.21

Period x SSCCS assessment	-22.53	8.47	-2.66	.008	-39.13	-5.92
ED x SSCCS assessment	10.81	8.47	1.28	.202	-5.80	27.74
ED x Incongruent trial	-7.19	3.32	-2.16	.031	-13.70	-0.67
SSCCS assessment x Incongruent trial	12.72	8.12	1.57	.117	-3.19	28.64
Incongruent trial x Prior trial incongruent	-19.77	4.02	-4.92	<.001	-27.64	-11.90
Step 4						
Period x ED x SSCCS assessment	-193.67	86.27	-2.24	.025	-362.76	-24.58
ED x Incongruent trial x Prior trial incongruent	$\chi^2(1, N = 55{,}726) = 0.56$			.454		
Intercept	490.99	61.33	8.01	<.001	370.79	611.19
			Random parameters			
ED	1449.22	174.61	8.30	<.001	1144.40	1835.22
Trial number	0.19	0.03	7.21	<.001	0.14	0.24
Incongruent trials	1278.96	159.70	8.01	<.001	1001.31	1633.60
Correct trials	12940.38	1541.76	8.39	<.001	10245.48	16344.13
Intercept	12444.61	2287.52	5.44	<.001	8679.94	17842.09

Notes. ED = Ego depletion manipulation; SSCCS = State self-control capacity scale assessment. Chi-square tests reflect omnibus tests.

[a]All *p*-values are two-tailed except in the case of the random parameters, where one-tailed *p*-values are used (because variances are constrained to be non-negative).

Table 12

Fixed effects estimates (top) and variance estimates (bottom) of errors as function of period, ego depletion, and SSCCS assessment.

					CI_{95}	
Fixed effects	*OR*	*SE*	*z*	*p*[a]	Lower	Upper
Step 1						
Period	0.82	0.07	-2.32	.020	0.69	0.97
ED	1.05	0.09	0.57	.568	0.89	1.25
SSCCS assessment	0.94	0.08	-0.71	.478	0.79	1.11
Step 2						
Period	0.84	0.05	-3.06	.002	0.74	0.94
ED	1.06	0.06	0.98	.325	0.94	1.19
SSCCS assessment	0.94	0.10	-0.60	.548	0.75	1.16
Gender	0.80	0.14	-1.26	.206	0.56	1.13
Age	1.00	0.01	0.02	.987	0.98	1.02
Trial number	0.99	0.01	-6.35	<.001	0.99	1.00
Lagged: RT on prior trial	1.45	0.12	4.47	<.001	1.23	1.70
Incongruent trial	14.50	1.83	21.21	<.001	11.3	18.53
Lagged: Prior trial incongruent	0.81	0.04	-4.79	<.001	0.74	0.88
Step 3						
Period x ED	0.91	0.20	-0.41	.680	0.59	1.41
Period x SSCCS assessment	1.12	0.13	0.95	.344	0.89	1.41

ED x SSCCS assessment	0.94	0.11	-0.54	.590	0.75	1.18
ED x Incongruent trial	0.64	0.14	-2.01	.044	0.42	0.99
SSCCS assessment x Incongruent trial	1.14	0.28	0.51	.607	0.70	1.85
Incongruent trial x Prior trial incongruent	0.33	0.12	-3.00	.003	0.16	0.68
Step 4						
Period x ED x SSCCS assessment	1.45	0.67	0.83	.406	0.60	3.51
ED x Incongruent trial x Prior trial incongruent	$\chi^2(1, N = 55{,}726) = 0.74$			.390		
Intercept	0.002	0.001	-8.17	<.001	0.001	0.008
	Random parameters					
ED	0.13	0.03	3.98	<.001	0.08	0.21
Trial number	0.00	0.00	3.64	<.001	0.00	0.00
Incongruent trials	0.29	0.09	3.15	.002	0.15	0.53
Intercept	0.10	0.09	1.14	.253	0.02	0.55

Notes. ED = Ego depletion manipulation; SSCCS = State self-control capacity scale assessment. Chi-square tests reflect omnibus tests.

[a]All *p*-values are two-tailed except in the case of the random parameters, where one-tailed *p*-values are used (because variances are constrained to be non-negative).

4.3.2 Reactivity of the SSCCS assessment

The second research question was whether the SSCCS is a suitable manipulation check for assessing perceived ego depletion in the first task without inducing this state. To test this question, we included the SSCCS assessment in the analyses above. Since the evidence reported by Job and colleagues (2010) was based on a between-subject analysis of error rates, we first examined the results of the between-subject analyses reported in Table 2. As was the case in the study by Job and colleagues (2010), none of the main effects of implicit theory on the five outcomes were significant. However, Job and colleagues (2010) found a significant interaction, in that performance deteriorated only when participants' implicit theories about willpower were primed to be limited (i.e., performance did not deteriorate when participants' implicit theories were primed to not be limited). The results of our analyses were heterogeneous, as indicated by the interaction in Table 2 and the simple main effects (Aiken & West, 1991) of the respective interaction in Table 7. Out of the five interactions between the ego depletion and SSCCS assessment manipulation, one was non-significant, two were significant, and two were almost marginally significant. For the errors outcome used by Job and colleagues (2010), the interaction in the present study was almost marginally significant, but the pattern of simple main effects differed from that found by Job and colleagues (2010): In the present study, participants who received the SSCCS manipulation had almost the same number of errors as participants who did not receive it did, with participants in the depletion condition without the SSCCS making almost marginally significantly more mistakes than participants in the control condition without the SSCCS. In turn, the interaction between ego depletion and SSCCS assessment for the difference or ratio score of the MSIT was significant, and simple main effects analysis revealed that participants in the depletion condition with the SSCCS showed smaller interference effects than participants in the control condition with the

SSCCS, indicating better performance due to completing the SSCCS after initial depletion. For reaction time for incongruent trials, this difference in the SSCCS condition was in the opposite direction, although it failed to reach significance (p = .114). Again, these results demonstrate the variability of results as a function of the chosen outcome.

In light of the clear role played by the outcome measures used, an analysis of the reaction time data and of the number of errors using a hierarchical random-effects multilevel models may provide clearer evidence for the influence of SSCCS assessment. As indicated in Tables 5 and 6, the interaction between ego depletion and SSCCS assessment failed to reach significance for both outcomes: Participants in the depletion condition responded slower to the MSIT trials of the second task than participants in the control condition, and this effect was more pronounced when participants received the SSCCS (z = 8.38, p = .004 vs. z = 1.19, p = .275); however, this interaction failed to reach significance. To further test whether completing the SSCCS induces ego depletion, we hierarchically included the fatigue manipulation check and its interaction with ego depletion and SSCCS assessment in the final hierarchical random-effects multilevel models of reaction times (Step 4 in Table 5). Neither the main effect of fatigue (b = 3.71, SE = 5.02, z = 0.74, p = .460) nor its interaction with ego depletion (b = 7.41, SE = 6.51, z = 1.14, p = .255) or SSCCS assessment (b = 10.34, SE = 10.00, z = 1.03, p = .301) were significant. However, the three-way interaction was significant (b = 37.79, SE = 12.76, z = 2.96, p = .003). Only participants in the depletion condition who completed the SSCCS showed a significant positive association between fatigue and reaction times (b = 24.25, SE = 8.76, z = 2.77, p =.006). This finding indicates that only the combination of an initial task requiring self-control and completing the SSCCS led to a possible overuse of self-control.

Table 13

Simple main effects analysis of the interaction between ego depletion and SSCCS assessment of the ANOVAs reported in Table 8.

	No SSCCS							SSCCS						
	DC		CC		Simple main effect			DC		CC		Simple main effect		
	M	SE_M	*M*	SE_M	*z*	*p*	*d*	*M*	SE_M	*M*	SE_M	*z*	*p*	*d*
Latency of incongruent trials	0.939	0.003	0.943	0.003	1.01	.317	.21	0.949	0.003	0.942	0.003	1.58	.114	.26
Difference Score	0.450	0.021	0.439	0.021	0.48	.635	.14	0.397	0.021	0.508	0.021	3.65	**<.001**	.40
Ratio Score	0.467	0.022	0.450	0.022	0.54	.594	.15	0.405	0.022	0.522	0.022	3.75	**<.001**	.41
Reaction Time Variability	0.401	0.022	0.361	0.022	1.24	.216	.23	0.339	0.023	0.351	0.023	0.37	.707	.13
Number of Errors	0.463	0.025	0.392	0.026	1.96	.052	.29	0.395	0.026	0.415	0.026	0.56	.576	.16

Notes. DC = depletion condition of the ego depletion manipulation; CC = control condition of the ego depletion manipulation. p-values in bold are significant at the Šidák-corrected significance level of α = .010.

4.4 Discussion

We set out to investigate the robustness of the ego depletion effect, both at one time point using a typical dual-task paradigm, and at two time points using a crossover design that combined two dual-task sessions with a two-week wash-out period. Both the between-subjects analysis at one time point and the within-subjects analysis at two time points provided limited evidence for the existence of a substantial ego depletion effect. Effect sizes of the ego depletion manipulation on the five MSIT outcomes at the first dual-task session ranged from d = -.34 for the ratio score to d = .15 for the number of errors. Of these effect sizes, only the effect for the ratio score was significant, and in the opposite direction to that hypothesized, in that participants in the depletion condition performed better, not worse, on their second task. For the within-subject analysis of the crossover design, ego depletion exhibited a marginally significant effect on reaction times in the hierarchical random-effects multilevel model, in that participants performed significantly worse in the second task when their first task required self-control. We did not provide effect sizes for the multilevel models, as there is no general consensus on appropriate effect sizes in multilevel models (Peugh, 2010). However, Selya, Rose, Dierker, Hedeker, and Mermelstein (2012) advocate using Cohen's f^2 as a measure of local effect sizes in multilevel models. This measure has the advantage that it can be converted to Cohen's d for dichotomous variables, such as our ego depletion manipulation. Using this method, we computed an effect size of f^2 = .0009, which equals d = .06. Taken together, both the effect sizes of the between- and within-subject analyses are much smaller than the mean effect size of d = .62 reported by Hagger and colleagues (2010). Our effect sizes were more in line with the regression-based estimate of g = .003 reported by Carter and colleagues (2015), which was computed using precision effect estimation with a standard error technique, as well as with the effect size of d = .04 found by Hagger et al. (2016) in their replication study on ego depletion.

Our results on the influence of both the ego depletion and SSCCS assessment manipulations on the five outcomes of the MSIT demonstrate how much flexibility researchers have (Simmons, Nelson, & Simonsohn, 2011) in analysing Stroop-like tasks in ego depletion research. The choice of outcomes not only affected the size of the ego depletion effect and manipulation check reactivity, but also their direction. This outcome variability may partly explain the ambiguous results and small-study effects found in the ego depletion literature, as researcher are able to pick the dependent variable that suits their hypotheses the best.

To reduce this variability, we proposed using random-effects multilevel modeling to analyze raw reaction-time data in research on ego depletion (Baayen & Milin, 2010). In addition to limiting the number of outcomes from five to two (with the two being reaction time and errors) in the MSIT task, this technique allows researchers to investigate the temporal dependencies between successive trials in the MSIT. On the one hand, successive trials are often positively correlated, with trial performance being more similar the closer the trials are to each other (i.e., positive autocorrelation). On the other hand, successive trials also demonstrate trial-by-trial adaptation: Research on the Stroop task has shown that the Stroop effect (similar to the MSIT difference) is weaker for trials that take place directly after an incompatible trial (Kerns et al., 2004) due to the fine tuning of information processing that participants engage in in response to the detected conflict in the prior incompatible trial. A multilevel analysis of reaction time and errors allows researchers to investigate and quantify those adaptation processes. A second major advantage of multilevel analysis is that it allows researchers to model random effects, in that the slope of each subject may differ, resulting in a variance component that indicates the degree of variability of the fixed effect of that slope. In ego depletion research, modeling random effects allows researchers to estimate the interindividual variation in ego depletion effects, which can then be used to compute each participant's individual level of ego depletion. As indicated by the significant random effect of the ego depletion manipulation depicted in Table 5, performance differed highly between the depletion and control condition. Approximately 95% of the participants were within -63.95 and 85.28 ms ($10.65\ ms \pm \sqrt{1449.22}$ x 1.96) of the typical value for ego depletion, illustrating that some individuals responded slower and others faster in the depletion compared to the control condition. Although the inclusion of the control variables and interactions reduced variability from z = 13.36 to z = 8.30, the random effect of ego depletion was still significant in the final model, indicating evidence of further moderators of ego depletion that were not included in the final model. Using the crossover design, future research should investigate trait moderators of ego depletion, such as trait self-control (Imhoff, Schmidt, & Gerstenberg, 2011) and implicit theories about willpower (Job, Bernecker, Miketta, & Friese, 2015). This future research should also explore the role of state moderators such as self-awareness and self-monitoring (Alberts, Martijn, & de Vries, 2011; Wan & Sternthal, 2008), mindfulness (Friese, Messner, & Schaffner, 2012), and motivation (Muraven & Slessareva, 2003; Inzlicht & Schmeichel, 2012) to explain this within-subjects variability in ego depletion effects, and to do so with high statistical power.

A large portion of the sample performed better in their depletion condition, indicating gains in self-control performance after initial self-control execution. This phenomenon can be explained by the counteractive self-control theory (Fishbach, Zhang, & Trope, 2010). According to this theory, self-control demands do not always lead to failure, but can instead in fact highlight one's self-control goals. Initial evidence for this effect in ego depletion research was reported by Wenzel, Zahn, Rowland, and Kubiak (2016). In this study, participants in the depletion condition set themselves stricter goals regarding performance in the second task than participants in the control condition, but were less able to follow through with their more ambitious intentions when faced with increased self-control demands (Wenzel et al., 2016). Some participants may, therefore, show stronger intentions to control their impulses in the second task because the self-control demands in the first task trigger self-control goals that are still activated before and during the second task. Thus, future research would profit from not only focusing on limitations of self-control in settings using consecutive tasks, but also by exploring why some individuals even excel when performing consecutive self-control tasks.

This is the first study to investigate trial-to-trial adaption processes in ego depletion research. As is often found in research on cognitive control using the Stroop task (e.g., Lorist & Jolij, 2012), participants perform better on incongruent trials when the prior trial was also incongruent. In our study, we found the same effect. However, this adaptation effect to an incongruent trial did not interact with the ego depletion manipulation: Participants in the depletion condition showed neither reduced nor enhanced trial-to-trial adaptation after an incongruent trial. This further shows that an ego depletion manipulation might not be as impactful in limiting self-control performance as suggested by the medium-to-large effect size reported in the meta-analysis by Hagger and colleagues (2010). However, future research might still incorporate not only aggregated measures of self-control performance, but also analysis methods such as multilevel modeling to investigate processes on a more fine-grained level that may be relevant to research in social psychology.

A frequently mentioned problem in the ego depletion literature is that measures assessing control variables (e.g., ego depletion manipulation check) or motivation are not validated. In this study, we used the SSCCS to assess ego depletion and the QAM to assess task motivation. Regarding the latter, contrary to both MAPM as the most popular alternative explanation of the SMSC (Inzlicht & Schmeichel, 2012) and to previous evidence (e.g., Muraven & Slessareva, 2003), we found no significant performance differences in the ego depletion conditions due to differences in task motivation. Although we found differences for the ego depletion manipulation—participants rated themselves as less likely to succeed and were more afraid to perform in the first task of the depletion condition—none of the four subscales revealed significant differences in how motivated participants were to complete the second task. Moreover, although participants with high levels of challenge and low levels of anxiety regarding the second task performed better in general, none of the interactions of the four task motivation scales with the ego depletion manipulation were significant (not presented in the results section), indicating an influence of motivation independent of the ego depletion conditions.

As for the SSCCS assessment, our results for the hierarchical random-effects multilevel models gave evidence of reactivity, as expected by previous research on the influence of lay theories about willpower on the ego depletion effect (Job et al., 2010). While neither participants who completed a parallel form of the affect questionnaire nor participants in the control condition who completed the SSCCS showed significant associations between fatigue and reaction times, participants in the depletion condition who completed the SSCCS showed a positive correlation between those measures. These results indicate that, in the present study, only the combination of an initial self-control task and the presentation of limited self-control via the SSCCS led to a possible overuse of self-control. In light of this finding, measures that assess ego depletion more directly than measures of perceived task difficulty or fatigue should be used cautiously, as they reduce some participants' performance in the second task of the depletion condition. Future research on ego depletion might therefore modify or develop explicit depletion manipulation checks with reactivity issues in mind; they could also use other ways of assessing ego depletion, such as implicit measures (Friese, Hofmann, & Wänke, 2008) or goal deviation (Wenzel et al., 2016, see Chapter 5).

There were a number of limitations to the present study. One problem of crossover studies is that carry-over effects may occur, with a treatment effect carrying over from one period into the following one. A carry-over effect can be tested for by interacting the treatment with the sequence effect. In our data, the interaction between the ego depletion manipulation and the period was neither significant in the hierarchical multivariate related-outcomes multilevel model nor in the hierarchical random-effects multilevel model, indicating no carry-over effect. However, the main effect of period was significant in all models and for all outcomes, in that participants performed better in the second dual-task session two weeks later, independent of the order of the experimental conditions. This finding is especially important to keep in mind when investigating ego depletion, since it represents a much more optimistic view on self-control in laboratory settings: Not only was our effect size for the ego depletion manipulation on reaction times (d = 0.06) lower than the d = 0.20 defined by Cohen (1988) as indicating a small effect size, but participants also showed practice effects within a task that remained stable for two weeks. The local effect size of this practice effect was f = .19, which is larger than the ego depletion effect for reaction time (f = .03). Thus, although small significant differences could be found between the performance in the second task after completing a task that required or did not require self-control, this effect was overshadowed by the performance gains due to practice. On the other hand, self-control performance may decrease over longer periods of time, either during a long task (drawing on the mental fatigue literature; e.g., Boksem & Tops, 2008; Gergelyfi, Jacob, Olivier, & Zénon, 2015), or when repeatedly engaging in the same self-control behavior, such as dieting or smoking cessation (drawing on the literature on habit formation; Neal, Wood, & Drolet, 2013). Connecting ego depletion research with those other fields of research to investigate the boundaries and conditions of the limits of self-control in more detail represents a promising future research avenue.

Another problem of our study was that we recruited a convenience sample that consisted mainly of students of psychology. Since males and older participants tended to have slower responses to MSIT, with similar accuracy, basic socio-demographic variables influenced reaction times more than our ego depletion manipulation. Assessing socio-demographics variables in a more representative sample could not only provide evidence for the generalizability of ego depletion effects, but might also explain some of the high observed variability in reported ego depletion effects.

A third problem was that the wash-out period was not long enough to rule out practice effects. Although we interpret the lingering practice effects that we observed as an important result, rather than as a problem that invalidates our ego depletion manipulation, a longer wash-out period might have had led to slightly higher ego depletion effects.

Finally, the tasks that we used for the dual-task paradigm were the same as those used in Hagger et al.'s (2016) replication project on the ego depletion effect. This project was criticized by Baumeister and Vohs (2016) for relying on computerized tasks without context, and for not establishing a habit in the first task. However, in the present study, all participants completed 50 practice trials in which they had to react to every word containing the letter e; participants in the depletion condition were then switched to the complex ruleset. In light of this methodological measure, we do not think our results can be easily dismissed by assuming that the first task in the depletion condition did not establish a behavioral pattern.

To conclude, we believe that studying the limitations of self-control is an important endeavor for psychological research. However, in relation to recent evidence (Hagger et al., 2016), we generally found only small effect of ego depletion, which was influenced by methodological choices such as the type of dependent variable. The strong effect of such methodological choices may partially explain the heterogeneous results that have been reported in the ego depletion literature. In light of these findings, the study of ego depletion should be connected to research on cognitive control and to research using ambulatory assessment, allowing researchers to investigate those limitations with rigorous methods. This connection allows researchers to not only examine the fine-grained underlying processes, but to capture behavior directly in the context of daily life. We hope that future research will explore new pathways outside the typical dual-task paradigm to gain a more diverse and detailed understanding of the limits of self-control.

Chapter 5:

Goal Deviation as an Alternative Method of Capturing Limited Self-Control[5]

5.1 Introduction

Self-control is usually framed as the ability to control one's behaviour and its failures are seen as an important factor in individual and societal problems (Muraven, Tice, & Baumeister, 1998). In the last two decades, the SMSC has been a prominent theory to explain the limits of self-control (e.g., Baumeister, 2001). According to this model, overriding a dominant response draws on a limited self-regulatory resource. In this state called *ego depletion*, people are not able to spend further self-regulatory resources, resulting in subsequent reduced self-control. There is a plethora of evidence for this ego depletion effect: For example, a meta-analysis by Hagger, Wood, Stiff, and Chatzisarantis (2010) reported a medium-to-large pooled effect size of d = .62 for the ego depletion effect.

However, the mechanisms underlying the ego depletion effect observed in subsequent efforts in self-control are widely debated (Kurzban, Duckworth, Kable, & Myers, 2013). Ego depletion, for example, does not always occur: Recent research found evidence for moderating factors counteracting ego depletion, such as intrinsic motivation (Moller, Deci, & Ryan, 2006) or lay theories about self-control (Job, Dweck, & Walton, 2010). To acknowledge this evidence, several alternative accounts have been proposed, in which motivation plays a crucial role such as the MAPM, in which ego depletion is explained by shifts in motivation and attention, or justifications-based accounts, in which people employ justifications that permit

[5] This study was published in 2016 in a peer-reviewed journal: Wenzel, M., Zahn, D., Rowland, Z., & Kubiak, T. (2016). The benefits of self-set goals: Is ego depletion really a result of self-control failure? *PLOS ONE, 11*, e0157009. doi:10.1371/journal.pone.0157009

violating goals they endorse to resolve self-control conflict (see Section 1.1 for more information).

Recently, Carter and McCollough (2013; 2014) raised concerns that the pooled effect size of $d = 0.62$ found by Hagger and colleagues (2010) may be inflated by small-study effects such as publication bias. Using two correction methods to account for publication bias, they found an average effect size of $d = 0.25$ and $d = -0.10$, respectively. Subsequently, Carter, Kofler, Forster, McCollough (2015) conducted a meta-analysis using only laboratory tasks that relate directly to self-control and ego depletion and included non-published effect sizes and found a pooled effect size of $g = 0.43$ that was reduced to zero when accounted for small-study effects. Recently, Hagger and colleagues (2016) published the results of a large and rigorous multi-lab replication project of ego depletion that only found a non-significant mean effect size of $d = 0.04$, providing further evidence that the ego depletion effect may not be as robust as reported in the last nearly 20 years.

Thus, since assessing self-control failure by comparing the performance between the experimental and control condition may be limited and the ego depletion may not be reliably replicable, alternative ways of assessing self-control failures in ego depletion research could help clarify how strongly self-control is limited in laboratory research (Converse & DeShon, 2009). A situation requiring self-control is characterized by incompatible motivations, where doing something desired with positive short-term consequences conflicts with a long-term goal (Hofmann, Friese, & Strack, 2009; Fujita, 2011). However, ego depletion usually is assessed by a declined performance in the experimental compared to the control condition or to a baseline but only rarely in comparison to a self-set goal. Therefore, the group differences may not reflect a self-control failure, that is a limited capacity to control one's behavior, but may also reflect a reduced motivation to control one's behaviors. To clarify this question, it is necessary to assess the personal goals of the participants with regard to their performance in the second task directly before the second task. In this modified dual-task procedure, ego depletion could not only be operationalized by the performance difference in the second task between experimental and control task but also whether participants in the experimental condition experience more difficulties to adhere to their goal than participants in the control condition. Moreover, this approach enables testing the assumption of the MAPM (Inzlicht & Schmeichel, 2012) that people shift away from wanting to control themselves after initial depletion which should be reflected in participants setting a more lenient goal in the experimental condition compared to controls. Thus, the aim of the present research was to test the different assumptions made by the SMSC and the MAPM in order to disentangle the processes underlying ego depletion. The SMSC would not predict an effect of an initial act of self-control on control-related goals (but possibly on the adherence to those goals), whereas the MAPM would assume decreased control-related goal strength in participants with versus without an initial self-control task.

5.2 Experiment 1

5.2.1 Method

Both Experiment 1 and 2 were approved by the ethics committee of the Landesärztekammer Rheinland Pfalz (State Chamber of Medicine Rhineland-Palatinate; 837.069.13 (8754-F)). Written informed consent was obtained in both experiments.

5.2.1.1 Participants and Design

We recruited 136 students (84 female, age M = 24.0 years, SD = 3.5) through bulletins and flyers. Each participant received 5 EUR (approx. US$7) as compensation for participating and another 5 EUR as a bonus for completing all experiments. Based on outlier analyses, one participant was excluded due to performance issues in the first task ($z > 3$), six participants due to at least 33% errors in the first task, and one participant due to technical problems. The inclusion of the outliers did not substantially change the results, in that none of the significant results became insignificant and none of the insignificant results were significant without the exclusion of the outliers. Furthermore, a z-test (difference between a coefficient from a model with all participants and a model excluding outliers divided by the square root of the sum of the squared standard errors) did not reveal any significant deviations, $|z| < 0.22$, ps > .831. The outlier analyses resulted in a final sample size of 128 students (80 female, age $M = 23.8$ years, $SD = 3.4$). In a between-subject design (ego depletion), participants were randomly assigned to the depletion ($n = 63$) or the control condition ($n = 65$). We aimed for 128 participants to achieve a power of .80 for a t-test for the difference between two independent measures with a medium effect size ($d = 0.50$) and an $\alpha = .05$ and a power of .80 for a within-between interaction in a mixed ANOVA with a small effect size ($d = 0.25$) and an $\alpha = .05$.

5.2.1.2 Procedure

Experiment 1 was the third of three combined experiments, which were approved by the local ethics committee. The other two experiments were realized one and two weeks before the third experiment and participants were finally debriefed after this experiment. Participants were told not to eat anything for four hours until the experiment started. Prior to the experimental procedure, participants signed an informed consent form and indicated their baseline affect. Then, they completed the first task, which included our ego depletion manipulation. Afterwards, they rated its difficulty and completed the affect questionnaire again. The investigators explained the second task as a product test of cookies and asked the participant how many cookies he or she would like to eat. The investigator made a note of the desired amount, put twice as many cookies on a platter and said that he or she had brought twice as many cookies in case the participant wanted to eat more. Finally, participants had the opportunity to taste the cookies for four minutes and rated the cookies quality and their affect subsequently.

5.2.1.3 Ego depletion manipulation

To manipulate ego depletion, participants in the depletion condition performed the Multi-Source Interference Task (MSIT; Bush, Shin, Holmes, Rosen, & Vogt, 2003), which was developed to maximally stress the conflict-monitoring system (Botvinick, Braver, Barch, Carter, & Cohen, 2001). It was previously used in ego depletion research (Shamosh & Gray, 2007; Wenzel, Kubiak, & Conner, 2014). In the MSIT, participants respond to the character that differs from the other two characters (e.g., "1" in trial "*221*"). The MSIT consists of control (e.g., *xx3*) and interference trials (e.g., *221*, see Bush et al., 2003, for more information). Since, in control trials, the distractors are always the letters x and the place of the target always corresponds to the key participants have to press, there is no response conflict and, thus, control trials do not require self-control. Interference trials comprise several sources of interference, in that the distractors are also numbers and the place of the target does not correspond to the key participants have to press (e.g., *221*). Hence, since only interference trials require self-control, the difference between mean latency on interference trials and control trials serves as a measure of self-control performance.

Participants in the control condition also performed a MSIT but with control trials only. Since responses to control trials are usually faster than to interference trials, the duration of the MSIT in the experimental condition (M = 285.5 s, SD = 9.4) was longer than in the control condition (M = 248.7 s, SD = 7.0), $t(117) = 24.39$, $p < .001$, $d = 4.48$. However, including duration in the analyses presented below did not change the results.

In each MSIT trial, a white fixation cross appeared on a black background for 500 ms, followed by the stimulus for a maximum of 2000 ms. The participants responded by pressing predefined keys on a response pad (Cedrus RB730; Cedrus Corporation, San Pedro, CA, USA). Trials hitting the response deadline as well as error trials were excluded from subsequent statistical analyses. The whole experiment consisted of 180 trials, with 36 control and 144 interference trials, presented in the same predetermined random order because we were primarily interested in individual differences.

5.2.1.4 Measures

For the cookie task, participants received twice as many cookies as they wanted and tasted them for four minutes. Afterwards, they rated the cookies' quality with three items (taste, consistency, and appearance) on a scale from 1 ("very good") to 6 ("not good at all"). Reliability was acceptable to good, Cronbach's $\alpha = .79$. Participants also indicated how often and how willingly they eat cookies in general on a scale ranging from 1 ("not at all") to 7 ("very"). Reliability was good, with a Spearman-Brown reliability estimate $r = .74$ (Eisinga, Te Grotenhuis, & Pelzer, 2012). Finally, participants responded if they had paid attention not to eat too many cookies and if they had eaten more than they wanted. Our measures of self-control were the amount of cookies eaten by the participants and the deviation of this amount from the amount participants wanted to eat.

To control for group differences in regard to perceived difficulty of the first task, participants rated the three items "difficult", "frustrating", and "exhaustive" with a 1 ("not at all") to 5 ("extremely") scale. Reliability was marginally acceptable, with a Cronbach's α of .59.

5.2.1.5 Data Analysis

A mixed 2 x 2 ANOVA was conducted with the ego depletion manipulation as between-subject factor and cookies desired vs. cookies eaten as within-subject factor. The interaction between both factors indicates whether our experimental manipulation exhibits any effect on our three interesting variables: cookies eaten, cookies desired and the deviation. If the interaction was significant, simple main effects analysis (Aiken & West, 1991) was performed in order to clarify the effect of the experimental manipulation on those measures.

5.2.2 Results

5.2.2.1 Preliminary analyses

In order to check whether our manipulation of the first task worked as intended, an independent samples t-test with the depletion measure as the outcome revealed as expected that participants in the depletion condition ($M = 2.09$, $SD = 0.71$) perceived the first task as significantly more difficult than participants in the control condition ($M = 1.68$, $SD = 0.57$), $t(126) = 3.66$, $p < .001$, $d = 0.65$. Moreover, participants in the depletion and control condition did not differ in regard to how they rated the quality of the cookies, $t(125) = -0.38$, $p = .702$, $d = 0.07$, nor how much they enjoy eating cookies in general, $t(125) = -0.95$, $p = .342$, $d = 0.17$. Taken together, these results indicate that our ego depletion manipulation worked as intended and that there were no group differences regarding cookie quality and general preference.

5.2.2.2 Cookie task

We computed a 2 (between-subject: ego depletion) x 2 (within-subject: goal deviation: cookies desired vs. cookies eaten) mixed ANOVA. The interaction between ego depletion and deviation from the goal was significant, $F(1, 126) = 5.52$, $p = .019$, $\eta_p^2 = .04$. This indicates that participants in the depletion and control condition differed in regard to the amount of cookies they wanted to eat and they actually ate and, thus, simple main effects analysis was performed in order to interpret these differences (Figure 6).

5.2.2.2.1 Ego depletion and cookies desired

First, an independent samples t-test on the number of cookies desired showed that participants in the depletion condition wanted to eat significantly fewer cookies ($M = 9.33$, $SD = 7.46$) than participants in the control condition ($M = 12.92$, $SD = 12.29$), $t(126) = -1.99$, $p = .049$, $d = -0.35$. Thus, the motivation to control one's eating behaviour was not reduced but increased after initial depletion.

5.2.2.2.2 Ego depletion and cookies eaten

Furthermore, an independent samples t-test on the number of cookies eaten revealed that participants in the depletion condition ($M = 11.78$, $SD = 9.71$) did not eat more cookies than participants in the control condition ($M = 13.02$, $SD = 9.77$), $t(126) = 0.72$, $p = .474$, $d = 0.13$, indicating no evidence for the typical ego depletion effect.

5.2.2.2.3 Ego depletion and goal deviation

However, a paired t-test showed on the number of cookies eaten and desired indicated that only participants in the control condition ate the desired amount of cookies, $t(64) = -0.11$, $p = .911$, $d = -0.02$ with dependency-corrected effect size (Morris & DeShon, 2002), whereas participants in the depletion condition ate significantly more cookies than they wanted, $t(62) = 4.41$, $p = .001$, $d = 0.67$. Importantly, these results did not substantially change when the analyses were controlled for gender and for how often and how willingly the participants eat cookies in general.[6]

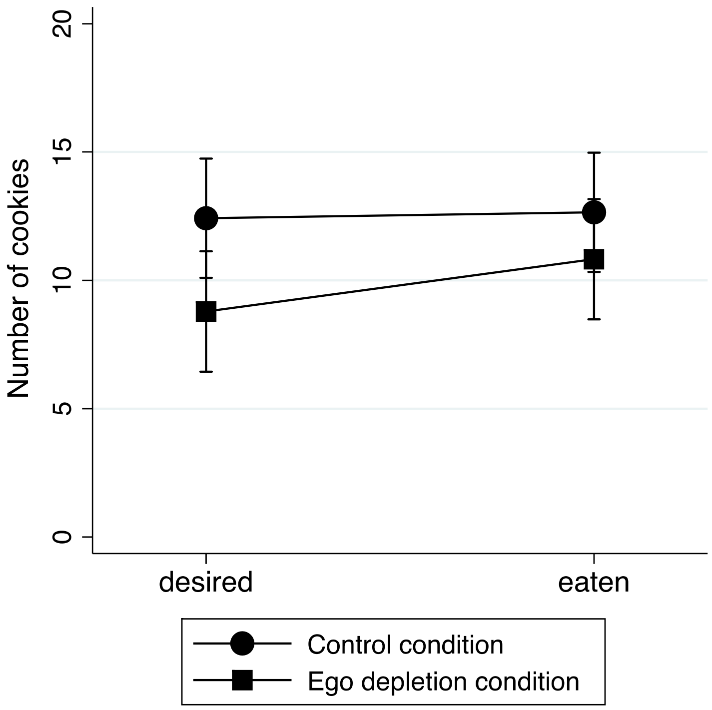

Figure 6. Mean number of cookies desired and eaten as a function of experimental condition (depletion vs. control condition). Error bars represent standard errors.

5.2.3 Discussion

To summarize, although participants in the depletion condition did not eat more cookies than controls, only participants in the depletion but not in the control condition ate more than they wanted. Moreover, the motivation to control one's eating behaviour was not reduced after initial depletion, as expected by motivational accounts of ego depletion. By contrast, it was notably increased, in that participants in the experimental condition wanted to

[6] However, when gender was entered as a factor that could interact with ego depletion and goal deviation, the three-way interaction was significant, $F(1, 122) = 5.25$, $p = .022$, $\eta_p^2 = .04$. Post-hoc analyses revealed that the two-way interaction was only significant for men, $F(1, 44) = 5.22$, $p = .022$, $\eta_p^2 = .11$, but not women, $F(1, 76) = 0.70$, $p = .402$, $\eta_p^2 = .01$.

eat fewer (not more) cookies than controls. Thus, other than assumed by the MAPM (Inzlicht & Schmeichel, 2012), participants in the depletion condition set themselves a stricter instead of a more lenient standard than participants in the control condition, indicating stronger intentions to self-control. Next, we conducted Experiment 2 in order to investigate the influence of motivation on self-control performance in more detail.

5.3 Experiment 1

The aim of Experiment 2 was threefold. First, since the results of Experiment 1 were quite surprising, in that we found neither evidence for the SMSC or the MAPM but found that goal strength is higher not lower in the experimental compared to the control condition, we deemed it necessary to conduct a conceptual replication in order to (a) provide further evidence for this effect and (b) generalize the findings of Experiment 1 using an experimental design suitable for many typical self-control tasks, increasing external validity. Since these tasks usually are more artificial than tasks and actions occurring in daily life, such as eating, people may not know how they can or would like to perform in them. Thus, we split the second task in the dual-task procedure in two parts and asked the participants in between how well they want to perform in the second part. Importantly, it might be that splitting the task may induce ego depletion since performing the first task may drain the self-regulatory resource, leading to worse performance in the second part. However, there is evidence that performing a task twice in rapid succession does not lead to ego depletion (Dewitte, Bruyneel, & Geyskens, 2009; Wenzel et al., 2013).

Second, we wanted to separate goal content from motivation, since individuals may set a goal without the necessary motivation to stick to it. We, thus, included a motivation factor, where participants in the motivation condition received an external reward (chance to win a lottery), to investigate whether a mild induction of external motivation could counteract the deviation from the self-set goal.

Third, we sought to assess motivational and attentional shifts in self-control, as proposed the MAPM, more elaborately. To this end, we developed a modified Stroop task designed to capture attention to neutral and motivational distractors. The standard Stroop task comprises two different types of trials: control trials, where the meaning of the word corresponds to the colour the word is presented in, and interference trials, where the meaning of the word does not depict the color. We extended this by two additional types that add a distracting background, which was either neutral (e.g., a table) or motivational (e.g., money) and, thus, captured attention to neutral or rewarding stimuli. If ego depletion is due to shifts in motivation and attention to reward, as proposed by the MAPM, participants with initial depletion should pay more attention to motivational distractors, leading to increased latency and, thus, larger Stroop effects.

5.3.1 Method

5.3.1.1 Participants and Design

We recruited 100 students (87 female, age M = 22.8 years, SD = 6.2) through a bulletin. Participants received partial course credit as compensation for participating. Outlier analyses did not reveal substantial outliers. Experiment 2 followed a 2 (ego depletion) x 2 (motivation) between-subject experimental design. We aimed for approximately 100 participants to achieve a power of .80 for an ANOVA with a comparable effect size typically found in ego depletion research of d = 0.62 (Hagger et al., 2010) and an α = .05.

5.3.1.2 Procedure

After signing an informed consent form, participants indicated their baseline affect and completed the first task where all participants copied a text by hand but participants in the depletion condition had to substitute each letter e with the number 3, which constituted our ego depletion manipulation. This task resembles a computerized version of the crossing-out-letter task, which demonstrated the largest effect sizes in prior research on ego depletion (Hagger et al., 2010). In both tasks, participants need to cross out or replace a letter according to certain rules. Since substituting is slower, the text in the control condition was extended to achieve a comparable duration of the first task, which was nearly the case, $t(98) = 1.72$, $p = .088$, $d = 0.35$. Afterwards, participants rated the difficulty of the first task and completed the affect questionnaire again. Then, participants performed the first part of the Stroop task which was followed by the question how well the participants wanted to perform in the second part of the Stroop task. Participants in the external motivation condition were additionally informed that they would take part in a lottery for a $70 voucher if they complete the second part of the Stroop task as good as the first part. Finally, participants rated their affect again.

5.3.1.3 Measures

5.3.1.3.1 Stroop task

Stroop stimuli were shown in a green or red font on a black background. Each trial began with a white fixation cross for 800 ms, followed by the stimulus for a maximum of 2000 ms. The participants responded by pressing predefined keys on the keyboard (g for green words and h for red words). All error trials including trials hitting the response deadline were excluded from subsequent statistical analyses. The first and second part of the Stroop task consisted of 70 trials each; with 18 control trials, 18 interference trials, 17 neutral distraction interference trials, and 17 motivational distraction interference trials in a predetermined random order since we were primarily interested in individual differences.

5.3.1.3.2 Goal strength

Goal strength for performing well in the second part of the Stroop task was assessed by the single item "How well would you like to perform in the second part of task compared to the first part?" on a 7-point scale ranging from -3 ("worse") to 3 ("better"). On average, participants wanted to perform better than in the first part (M = 1.14, SD = 1.13), with three participants indicating a negative goal strength and 71 participants a positive goal strength.

5.3.1.3.3 Depletion manipulation check

To achieve a better check for our depletion manipulation, we kept the item "exhaustive" and replaced the items "difficult" and "frustrating" with "taxing", assessed on a 1 ("not at all") to 7 ("extremely") scale. Reliability of the average score was good, with a Spearman-Brown reliability estimate of .79 for the first task and .81 for the second (Eisinga et al., 2012).

5.3.2 Results

5.3.2.1 Preliminary analyses

Participants in the depletion condition perceived the first task as more exhaustive and taxing (M = 4.12, SD = 1.08) than participants in the control condition (M = 3.31, SD = 1.43), as evidenced by a significant independent samples t-test on the depletion manipulation check, $t(98) = 3.18$, $p = .002$, $d = 0.64$. For the second task, this effect did not reach significance, $t(98) = 1.56$, $p = .121$, $d = 0.32$. This indicates that our ego depletion manipulation worked as intended.

5.3.2.2 Stroop task

As in Experiment 1, we first investigated whether our ego depletion manipulation exhibited any effects on Stroop performance, goal strength and the goal deviation. Afterwards, we report the effects of the motivation manipulation on those outcomes.

5.3.2.2.1 Ego depletion and goal strength

First, we investigated whether participants differed in the goal they set themselves regarding how well they wanted to perform in the second part of the Stroop task. To this end, we computed a 2 (ego depletion) x 2 (motivation) ANOVA on goal strength, which yielded a significant main effect of ego depletion, $F(1, 96) = 4.84$, $p = .030$, $\eta_p^2 = .05$. Participants in the depletion condition reported significantly stronger goal intentions ($M = 1.41$, $SD = 0.98$) than controls ($M = 0.92$, $SD = 1.21$), which mirrors the results of Experiment 1.

5.3.2.2.2 Ego depletion and Stroop performance

Next, in order to test for ego depletion effects, we conducted a 2 (ego depletion) x 2 (motivation) ANOVA on Stroop performance in the second task. As illustrated in

Table 14, which shows the simple main effect using an independent samples t-test of the ego depletion factor on three self-control measures of the Stroop task, participants in the depletion condition did not significantly differ in any of the performance measures in both parts the Stroop task compared to controls. As in Experiment 1, we did not found evidence for the typical ego depletion effect, reflected in a difference between experimental and control group.

Table 14

Means and standard deviations of the central variables as a function of ego depletion condition.

	Experimental Condition		Control Condition				
	M	*SD*	*M*	*SD*	*t*(98)	*p*	*d*
Stroop effect, part 1	18.41	56.79	11.82	46.59	0.64	.526	0.13
Stroop effect, part 2	34.39	31.64	28.80	35.13	0.84	.406	0.17
Incongruent trials, part 1	483.87	72.67	484.91	94.30	-0.06	.951	-0.01
Incongruent trials, part 2	453.45	61.40	442.65	57.66	0.91	.367	0.18
Errors, part 1	3.98	3.21	3.75	3.90	0.33	.744	0.07
Errors, part 2	2.41	2.01	2.10	2.11	0.75	.454	0.15

N = 100. p-values are two-tailed.

5.3.2.2.3 Ego depletion and goal deviation

In order to investigate goal deviation, we conducted a 2 (ego depletion) x 2 (motivation) ANCOVA on the latency of incongruent trials in the second part of the Stroop task and included goal strength (centered) and its interactions as continuous variables and differences in the latency of incongruent trials in the first part as a control variable. We focused on latency of incongruent trials only since the quicker adaptation to congruent trials in the second part of the Stroop task hampers the interpretation of the Stroop effect. As indicated in Table 15, goal strength interacted significantly with ego depletion, $p = .008$. Simple slope analysis (Aiken & West, 1991) revealed that participants in the depletion condition responded significantly slower to incongruent trials, $b = 14.51$, $SE = 6.51$, $t(91) = 2.23$, $p = .028$, indicating that they were significantly less able to follow through with their intentions. In contrast, performance of participants in the control condition was not influenced by goal strength, $b = -8.41$, $SE = 5.29$, $t(91) = -1.59$, $p = .116$. Thus, although participants in the depletion condition wanted to perform better than before in the second part of the Stroop task compared to controls, they could not keep up with their more ambitious goal.

Table 15

Analysis of covariance (ANCOVA) between ego depletion condition, manipulation condition and goal strength.

	F	*df*	*p*	η_p^2
Ego depletion	1.80	1	.184	.02
Motivation	0.02	1	.894	.00
Goal strength	0.46	1	.501	.01
Ego depletion x motivation	0.41	1	.523	.01
Ego depletion x goal strength	7.36	1	.008	.08
Motivation x goal strength	7.31	1	.008	.07
Ego depletion x motivation x goal strength	1.84	1	.178	.02
Latency incongruent trials, part 1	73.08	1	< .001	.45

N = 100. *p*-values are two-tailed.

5.3.2.3 Ego depletion and motivation

As illustrated in Table 15, the motivation condition and the interaction with ego depletion did not exhibit a significant effect on Stroop performance in the 2 x 2 (ego depletion) x 2 (motivation) ANCOVA. However, like the ego depletion condition, the motivation condition significantly interacted with goal strength, in that participants with an external reward could keep up with their more ambitious goals, $b = -8.63$, $SE = 5.45$, $t(91) = -1.58$, $p = .117$, whereas controls could not, $b = 13.83$, $SE = 6.32$, $t(91) = 2.19$, $p = .031$. Although the three-way interaction between ego depletion, motivation, and goal strength was not significant, $F(1, 91) = 1.84$, $p = .178$, $\eta_p^2 = .02$, simple slope analysis revealed the predicted pattern: When participants had the chance to take part in the lottery, both participants in the depletion and control condition performed better the higher they intentions were, although only the simple slope for participants in the control condition reached significance, $b = -2.85$, $SE = 8.37$, $t(91) = -0.34$, $p = .734$ and $b = -14.19$, $SE = 6.42$, $t(91) = -2.21$, $p = .030$. Without external motivation, performance of participants in the control condition was not significantly influenced by goal strength, $b = -2.85$, $SE = 8.88$, $t(91) = -0.32$, $p = .749$, but participants in the experimental condition showed difficulties keeping up with their intention, $b = 31.19$, $SE = 9.48$, $t(91) = 3.29$, $p = .001$. The same analysis with the number of errors showed a similar pattern: Both ego depletion, $F(1, 91) = 8.07$, $p = .005$, $\eta_p^2 = .08$, and motivation, $F(1, 91) = 4.08$, $p = .046$, $\eta_p^2 = .04$, interacted with goal strength significantly. Thus, only participants in the depletion condition demonstrated difficulties in keeping up with their intentions as in Experiment 1, especially without external reward.[7]

5.3.2.4 Attentional and motivational processes

Inzlicht and Schmeichel (2012) hypothesized in their MAPM that decreased performance after initial self-control may be due to increased attention to rewarding stimuli, which interferes with task demands. To investigate those attentional shifts in our modified Stroop task, we computed a multilevel model with latency as the outcome and simultaneously entering depletion and motivation conditions, trial type (congruent vs. incongruent vs. incongruent with neutral distractor vs. incongruent with motivational distractor), and part of the Stroop task as the predictors using the mixed command in Stata 13 (Stata Corporation, College Station, TX, USA). This model revealed a main effect for trial type, $\chi^2(1, N = 13{,}335) = 150.24$, $p < .001$. Responses to congruent trials were the fastest ($M = 441.45$ ms, $SE =$

[7] The results did not substantially change when the analyses were controlled for gender.

6.49), followed by incongruent trials (450.08 ms, *SE* = 6.49), incongruent trials with neutral distractor (*M* = 467.08 ms, *SE* = 6.53), and incongruent trials with motivational distractor (*M* = 478.02 ms, *SE* = 6.53). All Šidák-adjusted group comparisons were significant. Moreover, the part of the Stroop task also was significant, $\chi^2(1, N = 13{,}335) = 302.48$, $p < .001$, with participants responding faster in general in the second part. Importantly, the part of the Stroop task also interacted with the trial type, $\chi^2(3, N = 13{,}335) = 26.06$, $p < .001$. Analysis of marginal means (see Figure 7) indicated that whereas in the first part of the Stroop task, congruent trials, incongruent trials and incongruent trials with neutral distractor did not differ significantly from each other, *t*s(13,203) < 1.99, *p*s > .747 (Šidák-adjusted), mean latency on incongruent trials with motivational distractors was significantly higher compared to the other trial types, most importantly incongruent trials with neutral distractors, $t(13{,}203) = 4.98$, $p < .001$. In the second part of the Stroop task, this difference disappeared, $t(13{,}203) = -0.31$, $p = 1.00$. However, none of the experimental conditions influenced this two-way interaction significantly, χ^2s$(1, N = 13{,}335) < 3.63$, *p*s < .304, which demonstrates that not only participants with initial depletion paid more attention to rewarding stimuli.

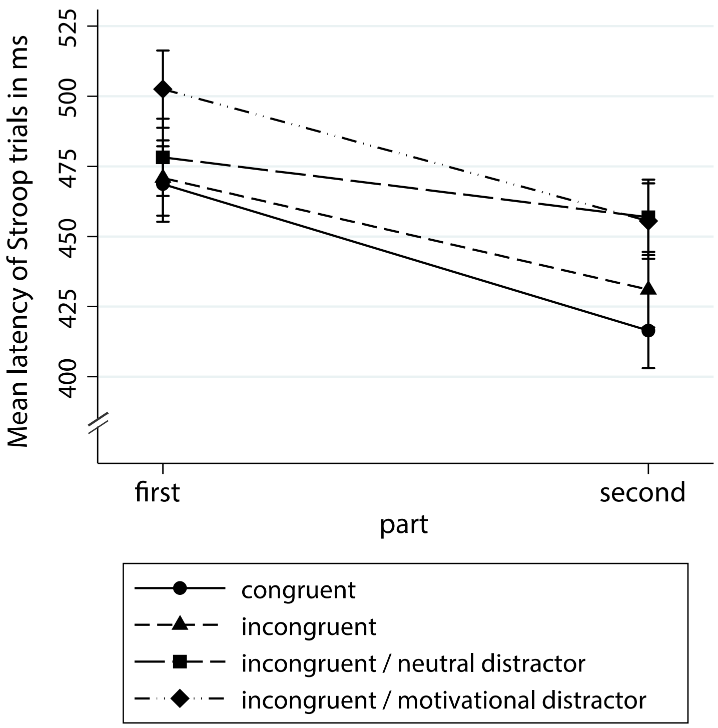

Figure 7. Mean latency of Stroop trials as a function of Stroop task part and trial type. Error bars represent standard errors.

5.4 General Discussion

We investigated the roles of motivation and goal strength in ego depletion in two experiments. We did not find evidence for ego depletion, in that participants with a first requiring self-control in Experiment 1 ate as many cookies and performed equally well in the Stroop task in Experiment 2 than participants with a first task that did not require self-control. However, participants in the depletion condition in Experiment 1 and 2 reported stronger intentions to control themselves than controls. Importantly, participants in the depletion condition in Experiment 1 could not keep up with their goal and ate more than they wanted, whereas participants in the control condition did not deviate significantly from their self-set goal. This effect was conceptually replicated in Experiment 2, in that only participants in the experimental condition and without an additional external reward performed the worse the stricter their goal was.

Thus, we did not find strong support for the MAPM by Inzlicht and Schmeichel (2012). Whereas other research shows that initial depletion increases approach motivations (Schmeichel, Harmon-Jones, & Harmon-Jones, 2010), participants in the depletion condition did not want to eat more, but significantly less cookies (Experiment 1) and wanted to perform better than controls (Experiment 2). In contrast to Schmeichel and colleagues (2010), asking people how much they would like to eat taps into both, approach motivation and avoidance motivation (controlling an unwanted behaviour), which could partially explain the different results. However, in Experiment 2, we also found that both groups demonstrated an increased approach motivation at the beginning of the second task, which diminished over the course of the task. Hence, attention to motivational stimuli may not be a proxy for goal strength and the notion of a specific increase after initial depletion could not be confirmed.

Our results are in line with predictions made by the counteractive self-control theory (Fishbach, Zhang, & Trope, 2010). According to this theory, temptations do not always lure people into failure but may also increase the strength of the goal to control one's own behaviour in this situation. Thus, initial depletion may have triggered self-control goals that are still activated before and during the second task, resulting in stronger intentions to control the behaviour. In turn, the activated self-control goal led to exerting greater control in the second task, which may have counteracted the typical ego depletion effect. However, participants in the depletion condition were less able to follow through with their more ambitious intentions when faced with increased temptation. Thus, it remains unclear whether ego depletion may involve some sort of self-control failure or whether participants in the depletion condition may set themselves higher standards that are a bit too optimistic. To separate these two explanations, future research should combine research in ego depletion with counteractive self-control theory: For instance, goals could be assessed *before* both tasks in order to examine whether performance in the second task deviates from the baseline goal strength after an initial self-control task, suggesting some sort of ego depletion effect, or whether their goal striving is too optimistic, indicated by a performance in the second task that is in line with the baseline goal but not with an increased goal strength after the first task.

A potential limitation of this study was that we did not find ego depletion effects in both experiments. The absence of a significant ego depletion effect may indicate that our experimental manipulation failed, which could hamper the interpretability of our analyses. However, we not only found similar effect sizes in both experiments compared to Carter and colleagues (2015) but we could also show that there are other processes and performance limitation in the dual-task design except for a group difference in the second task. That is, participants in the depletion condition set themselves stricter goals than participants in the control condition, which is not in line with the assumption that initial acts of self-control lead to decreased motivation to control oneself. This effect remained of similar size when participants' habitual preference for cookies was controlled for in Experiment 1. Moreover, we could find evidence for the predicted shifts in attention towards reward stimuli in Experiment 2; these shifts, however, were independent from our ego depletion manipulation: Even participants in the control condition showed larger Stroop latencies in trials with motivational distractor compared to neutral ones. Finally, our effect sizes in Experiment 1 ($d = -0.08$) and Experiment 2 (around $d = 0.20$) are close to the bias-corrected pooled effect size of $d = 0.25$ and $d = -0.10$ reported by Carter and McCollough (2013; 2014). Thus, our research indicates that assessing the deviation from self-set goals presents an interesting alternative way of assessing self-control failure.

Furthermore, we chose a cookie tasting task in Experiment 1 since consuming sugary food is often tempting and controlled by individuals and individuals can easily report how much they want to eat of it. Although the participants reported that they had paid attention not to eat too many cookies, future research could limit the sample to individuals that restrict their eating behaviour or could use a task where individuals are asked to eat radishes while avoiding cookies in order to increase the self-control conflict.

Another limitation refers to our manipulation check of ego depletion. In the first experiment, we assessed task difficulty, although commonly used in ego depletion research (Hagger et al., 2010), instead of depletion, which we changed in the second experiment to items that better reflect depletion. However, given the importance of assessing initial depletion, developing a scale that can validly differentiate between difficult and depleting tasks would be of utmost importance to advance self-control research.

Although we did not find the typical ego depletion effects, our results suggest that initial acts of self-control involve deviation from a self-set goal. Whether this deviation can be seen as a self-control failure is still unclear. Clearly, more research is needed that compares the assumptions of the resource and motivational approaches, but we believe that assessing ego depletion as a deviation from a self-set goal is a promising method in this endeavor.

Chapter 6:

General Discussion

6.1 Overview and Integration of the Present Findings

Ego depletion is a concept that is increasingly popular for researchers as well as lay people, as indicated by the large body of scientific work on the topic and by the best-selling books about willpower and ego depletion (e.g., Baumeister & Tierney, 2011). The aim of this dissertation was to explore important open questions in research on ego depletion. How does positive affect influence self-control both in laboratory and ambulatory settings? Can the effect of positive affect be used to compare the diverging predictions made by competing explanatory models of ego depletion? How large and robust is the ego depletion effect? Do ego depletion effects differ when assessed between- or within-subjects? Can the subjective experience of ego depletion be measured without inducing this state? Is ego depletion accompanied by shift in goal strength regarding the motivation towards the second task? In this dissertation, I took several approaches to these questions by extending prior research results and introducing novel statistical methods for the field as well as a new way of assessing limited self-control.

The purpose of Chapter 2 was to test the different assumptions about the effect of positive affect on consecutive self-control performance made by the SMSC and the CMT. Both theoretical frameworks differ in their predictions with regard to the impact of affect and task characteristics on self-control deterioration within a dual-task paradigm. Whereas the SMSC predicts decrements in self-control performance whenever both tasks require a limited resource, under the cognitive control perspective of the CMT, decrements should only occur when people switch to a different response conflict in the second task. Moreover, only the CMT predicts an interaction between task switching and positive affect. In Chapter 2, I, thus, investigated this interaction within a dual-task paradigm and found evidence that favored a cognitive control interpretation of the results based on the CMT. Positive affect only benefitted consecutive self-control performance if response conflicts in the two tasks were different (resisting sweets followed by a Stroop task). If they were the same (two consecutive Stroop tasks), positive affect impaired self-control performance. These effects were partially replicated in the second study that also examined negative affect, which did not affect self-control performance.

The findings of Chapter 2 are in line with another study the author of this dissertation published (Wenzel, Kubiak, & Conner, 2014) and which was part of his diploma thesis. The aim of this study was to test the differential effects of positive affect on consecutive self-control performance in an improved experimental design that is able to investigate whether the differential effects of positive affect on task performance in the second task is mediated by attentional shifts. The design was improved on important points that have not been addressed in the two studies in Chapter 2. First, perceived difficulty was only assessed for the first task in the two studies in Chapter 2, so it could not be tested whether performing the same task twice with different stimuli was easier and, therefore, yielded in a better Stroop performance. Second, the effect of positive affect on the stability-flexibility-balance was tested only by the Stroop effect in the second task but without a direct measure of attention to the new self-control cues of the second task. Third, participants in the stability condition performed the same task twice rather than similar tasks. Thus, the second task did not contain new (distracting) information, which hampers the interpretation of a stability-reducing effect of positive affect due to increased distractibility.

Therefore, a dual-task experiment with a cognitive reaction time task was employed that allows assessing self-control performance and attention to self-control demands separately. In the second task of this dual-task experiment, participants performed the Multi-Source

Interference Task (MSIT; Bush et al., 2002), which was also used in Chapter 4. This standard MSIT was modified by adding color as an additional source of interference to test for attentional processes between the two tasks in the dual-task paradigm. In the color version of the MSIT (cMSIT), the characters in control trials are always black; the characters in interference trials are always green and red with two characters presented in one color and the third character in the other color. Consequently, color can be congruent or incongruent to the target dimension. Hence, attention to color by the prolonged reaction time latency to color-incongruent trials can be compared to color-congruent trials. As in Chapter 2, the content of the first self-control task was manipulated to establish three conditions of conflict similarity. Participants in the *stability condition* performed the standard MSIT without color and, thus, had to ignore the distracting color as a new information, which requires stable self-control. Participants in the flexibility condition performed a modified version of the cMSIT, in which participants had to respond to the color of the character that differed from the color on the remaining two characters. Consequently, participants had to switch from responding to the different color to the different character, which requires flexible self-control. Participants in the *control condition* did not perform any type of the MSIT but were instructed to copy a text about psychology by hand into the textbox under the text. This task does not require self-control to perform and allowed to compare performance in the second task without practicing a similar task in the first task. Finally, positive affect was manipulated as a between-subjects factor with two conditions. Participants in the *neutral affect condition* watched a film clip of a lecture from the US American physicist Richard Feynman about the string theory, whereas participants in the *positive affect condition* watched a cartoon of funny moments in the life of a four-fingered, yellow colored family.

The findings of this study by Wenzel and colleagues (2014) show that participants under positive affect in the flexibility condition, who had to switch to a new response dimension, were less attentive to distracting information compared to neutral affect, leading to better performance. In contrast, participants under positive affect in the stability condition, who did not have to switch, were more attentive to distracting information compared to participants under neutral. These findings highlight the opposite effects of positive affect on consecutive self-control depending on task similarity and, thus, replicate the results found in Chapter 2 in an improved design. Taken together, the two studies in Chapter 2 and the additional study (Wenzel et al., 2014) demonstrate the differential effect of positive affect on ego depletion, in that it can either improve or impair self-control performance, depending on whether two tasks are dissimilar, and thus require flexible releasing and switching, or similar, which requires stable maintenance.

Chapter 3 builds on this research in order to explore if the flexibility-enhancing effects of positive affect can be found in everyday life as well. In a daily diary study, 297 participants completed a 13-day daily diary that included measures of positive affect, desire, and habit strength as well as three self-control strategies (i.e., monitoring, distraction, and stimulus control). The findings in Chapter 3 indicate specific effects of positive affect on self-control strategies: Individuals with higher positive affect were most successful when following a strategy of distraction (e.g., thinking about something else), particularly when faced with strong tempting desires. These results reinforce the idea that positive affect is associated with both cognitive flexibility and distractibility, which may help people distract them from tempting desires, and provides evidence that the effects found in laboratory research can be found in daily life as well, demonstrating the internal and ecological validity of the flexibility-enhancing effect of positive affect.

The flexibility-enhancing effect found in the two studies presented in Chapter 2 cannot be explained by the popular SMSC since if positive affect replenishes a limited resource, positive affect should always have a positive effect on self-control performance after initial depletion. Although the detrimental effect of positive affect on similar tasks that require stability and not flexibility can be explained by the CMT, both the SMSC and CMT assume the existence of ego depletion effects of a considerable size. However, the ego depletion effect may not be as robust as previously shown (Hagger et al., 2010). Evidence for this problem is already is seen in Chapter 2 in which one depletion condition failed to demonstrate performance difference in comparison to the control condition. The purpose of Chapter 4 was, thus, to build on a crossover design that enables high statistical power with only a single study in order to pursue three aims: (a) to investigate the robustness of the ego depletion effect between and within subjects; (b) to compare ego depletion across a range of established measures of a self-control task, thus allowing us to judge the influence of the researcher's choice of analysis method; and (c) to estimate the reactivity effects of an explicit manipulation check of ego depletion. The findings in Chapter 4 demonstrated that (a) the within-subjects ego depletion effect only had small effect sizes, (b) there was substantial heterogeneity in the results depending on the outcome measure, and (c) there was significant reactivity for our explicit ego depletion manipulation check, resulting in worse performance after initial performance.

The small effect sizes found in Chapter 4 are in line with a recent meta-analysis on ego depletion controlling for publication bias (Carter et al., 2015) and with the results of a pre-registered multilab replication project of ego depletion (Hagger et al., 2016). However, the results in Chapter 4 were also similar in regard to the large variability in the ego depletion effect size, in that a large portion of the sample performed worse after initial depletion compared to no depletion but an almost equally large portion performed better. This finding indicates the existence of trait or state influences that affect self-control performance after previously exerting self-control.

In light of the small ego depletion effect sizes and the large variability found in Chapter 4, Chapter 5 introduces a design that presents an alternative way of capturing limited self-control in consecutive self-control demands. Typically, limitations of self-control are inferred from performance differences – either between-subjects as in Chapter 2 and 3 or within-subjects as in Chapter 4. However, as outlined before, inferring from self-control performance is problematic since reduced performance could be the result of various factors such as self-control capacity or motivation, which renders it difficult to directly separate those competing explanations. Thus, Chapter 5 added to the dual-task paradigm the assessment of goal strength between the two tasks, in that participants indicate the strength of a goal in regard to the second task (e.g., how many cookies they would like to eat) that, then, can be compared to their actual performance in this task. Following, the assumptions made by the SMSC and the MAPM can directly be compared: If the SMSC holds true, participants exerting initial self-control would show equally large goal strength that they cannot keep up with, leading to performance decrements in the depletion compared to the control condition. If the MAPM holds true, goal strength would be reduced in the depletion compared to the control condition, which, in turn, mediates the performance decrements. In two dual-tasks experiments, Chapter 5 did not provide evidence for any of the predictions: Participants who initially exerted self-control set themselves a stricter instead of a more lenient goal than controls, in that they chose to eat less cookies (Experiment 1) or wanted to perform better (Experiment 2). Moreover, only participants without an initial self-control task could adhere to their self-set goal, whereas participants in the ego depletion condition in both experiments could not follow through with their more ambitious intentions. Taken together, the findings neither support the SMSC nor the MAPM but highlight the importance of goals in ego depletion research.

To sum up, the present studies of this dissertation reflect the ambiguous evidence regarding the ego depletion effect found in previous research efforts (Carter et al., 2015; Hagger et al., 2016). Not only indicate the results in Chapter 4 that the ego depletion is only small in size with a large interindividual variability but neither resource (SMSC) nor motivational accounts (MAPM) provide a good fit in explaining ego depletion when it can be observed. However, the findings also demonstrate the usefulness of the CMT in explaining ego depletion due to switch costs. Moreover, the results further demonstrate the usefulness of assessing goal strength that can be compared to the actual performance, providing a measure that captures the very essence of self-control – whether participants are able to achieve a goal they set themselves by shielding it from distracting influences and by employing strategies that facilitate goal achievement.

6.2 Limitations and Implications for the Study of Ego Depletion

Due to the large number of relatively diverse studies in this dissertation (four studies with six experiments), limitations and implications of the results of each study have been discussed separately in the respective chapter and will not be repeated here. Instead, I will outline the overall implications for future research on ego depletion in the following section.

The results throughout this dissertation exemplify that neither the assumption of a limited resource made by the SMSC nor the assumption of motivational shifts leading to attentional shifts away from control cues made by the MAPM could be supported in this paper. Although none of the studies directly assessed a candidate of a limited self-control resource and only Chapter 5 consisted of a direct test of the MAPM – which leaves enough room for more opportunities for research testing the assumption – the evidence is relatively clear: Performance decrements due to initial acts of self-control cannot be consistently observed in the dual-task paradigm due to small effect sizes, ranging from d = -0.08 to d =

0.20 in the study with the highest statistical power (Chapter 4). On the one hand, when the ego depletion effect appeared (Chapter 2), both the SMSC and the MAPM have difficulties explaining why positive affect would have differential effects on consecutive self-control performance, depending on task similarity. If positive affect replenishes the limited resource (SMSC) or if positive affect acts as a reward signal and, thus, counteracting the motivational and attentional shifts (MAPM), positive affect should not only counteract ego depletion when tasks are dissimilar but should not impair performance when tasks are similar. The CMT provides an explanation for these results that is based on cognitive control models (e.g., Botvinick et al., 2001). However, although a transfer of the CMT from processes on a trial-to-trial basis to processes on a task-to-task basis is promising, open questions remain. For example, Chapter 2 found perseveration in two dissimilar tasks, although the two tasks involved highly similar forms of response conflicts and, thus, may not require different control processes. Future research should, therefore, test the application of the CMT for between-task processes more elaborately. Moreover, future research should also investigate whether adaptation to the self-control demands of the first task can explain the ambiguous results for the existence and size of the ego depletion effect (e.g., Carter et al., 2015) as well as for the large variability with- and between-subject found in Chapter 4 as well as in previous research (e.g., Hagger et al., 2010).

On the other hand, when the ego depletion effect did not appear in Chapter 5, both the SMSC and the MAPM could not explain why goal strength was increased instead of equal (SMSC) or decreased (MAPM). As indicated in the discussion in Chapter 5, the counteractive self-control theory provides an explanation for the increased goal strength based on the assumption that initial self-control efforts may have triggered self-control goals that are still activated before and during the second task, resulting in stronger intentions to control the behavior. This fits in with the CMT, which, in turn, can explain why some participants may not be able to follow through with their more ambitious intentions when faced with increased temptation: Although participants indicated greater goal strength, not only the self-control goals but also the adapting processes to the self-control demands in the first task may still be activated, leading to perseveration in the second task and counteracting increased intentions to perform well.

Taken together, the findings in this dissertation highlight the importance of assessing underlying processes of ego depletion in order to gain a deeper understanding of the boundaries and conditions of ego depletion. Future research should adopt the crossover design introduced to ego depletion research in Chapter 4 or similar within-subjects design to solve a striking discrepancy in ego depletion research: Whereas ego depletion is conceptualized as an intraindividual process, prior research mainly focused on between-subjects comparisons which are not suited to investigate and explain processes that occur within individuals.

6.3 Future Directions

In this section, I will outline future research avenues for studying self-control limitations both in laboratory and ambulatory settings. I will focus on three particularly promising areas: (a) the investigation of ego depletion and mental fatigue in a common framework, (b) the investigation of why some individuals perform better and others worse when experiencing a self-control conflict, and (c) the investigation of self-control limitation in the everyday life of individuals.

6.3.1 Ego depletion and mental fatigue

A promising line for further research pertains to combining research on ego depletion with the mental fatigue literature. Mental fatigue is commonly defined as a state that arises after prolonged periods of cognitive activity and that is characterized by an unfocused mental state (e.g, bored and distracted) and low arousal negative mood (e.g., weariness and tiredness) (Hockey, 2013). Furthermore, it is characterized by decreased levels of task commitment combined with reluctance to continue with the present task, leading to performance declines (Boksem, Mijman, & Lorist, 2006; Boksem & Tops, 2008). Thus, it is apparent that ego depletion and mental share large similarities: Both concepts involve continuous effort, which is accompanied by negative feelings, including subjective measures of fatigue (Hagger et al., 2010). Moreover, both ego depletion and mental fatigue assume performance decrements that are due to incapacity of a limited resource in older accounts (Muraven et al., 1998 for ego depletion; Gopher, 1986; Kahneman, 1973; Wickens, 1984 for mental fatigue) and that are challenged by accounts that do not rely on a limited resource (MAPM and CMT for ego depletion; Hockey, 1997; Hockey, 2013 for mental fatigue). This supported by the evidence that both ego depletion and mental fatigue can be counteracted with sufficient task-motivation (Muraven & Slessareva, 2003; Boksem et al., 2006). However, although ego depletion has been described as a form of fatigue (e.g., Muraven & Baumeister, 2000), proponents of resource accounts such as the SMSC state that "ego depletion is not just fatigue" (Vohs, Glass, Maddox, & Markman, 2010, p. 166). Vohs and colleagues (2010) provide evidence for this predication by distinguishing the effects of ego depletion from sleep deprivation: Conducting a study using the dual-task paradigm, the authors demonstrated that participants who exerted initial self-control showed more aggression in the second task compared to participants without initial self-control exertion, indicating the typical ego depletion effect. However, participants that were deprived of sleep for 24 hours to induce the state of mental fatigue showed similar levels of performance compared to rested participants. Although this might seem to indicate that sleepiness did not affect self-control performance, it is important to note that the operationalization of mental fatigue via sleepiness by Vohs and colleagues (2010) is problematic. Sleep researchers view mental fatigue and sleepiness as distinct states (Hossain, Reinish, Kayumov, Bhuiya, & Shapiro, 2003) that motivates different things: Whereas mental fatigue motivates individuals to disengage form a boring and laborious task to a wide array of activities that can provide change and interest, sleepiness motivates individual only to sleep.

Although aligning ego depletion with mental fatigue would allow the study of self-control to build on a strong foundation of past findings that have been demonstrated that performance decrements due to mental fatigue are the product of a decision to withdraw further effort rather than some cognitive incapacity (Hockey, 2013), it is important to note that ego depletion and mental fatigue typically appear at different times in experimental setups: While mental fatigue appears after prolonged acts of cognitive control (e.g., Hockey & Earle, 2006), ego depletion is almost exclusively studied in short initial tasks. However, although motivation might be the common underlying process, as assumed by the MAPM and the OCM, Chapter 2 demonstrates that the assumptions derived by the CMT provide an alternative explanation of ego depletion that does not render self-control as volatile as the assumption that participants are quickly demotivated to exert self-control in laboratory tasks due to short acts of self-control. Thus, future research should combine evidence from ego depletion, mental fatigue, and cognitive control models such as the CMT in order to gain a better understanding of the boundaries and conditions of limited self-control. One example of such future research could be to extend the dual-task paradigm to more than two tasks. For example, participants could perform 12 twelve consecutive dual-task setups for a total of 90 minutes where the even task is always the same task in order to compare self-control performance between the experimental conditions. The odd task, then, can either be a control task, a similar task compared to the same even task or a dissimilar one. Combined with a motivation manipulation, this setup would allow to (a) disentangle effects of adaptation, practice, and mental fatigue effects, (b) investigate the influence of motivation on those effects, and (c) mapping performance limitations on a broader timeline.

6.3.2 Moderators of ego depletion

Research on ego depletion did not only report a large body of evidence on the ego depletion effect but also on moderators of this effect. For example, trait influences were found for self-control (Imhoff, Schmidt, & Gerstenberg, 2011) and implicit theories about willpower (Job, Bernecker, Miketta, & Friese, 2015). For state influences, research reported evidence of moderating effects for self-awareness and self-monitoring (Alberts, Martijn, & de Vries, 2011; Wan & Sternthal, 2008), mindfulness (Friese, Messner, & Schaffner, 2012), and motivation (Muraven & Slessareva, 2003; Inzlicht & Schmeichel, 2012). However, these studies are characterized by the same limitations, in particular low statistical power. Up to date, there is no meta-analysis or replication project on the moderators listed above but the large within-subjects variability in ego depletion effects in Chapter 4 as well as the heterogeneity reported in the meta-analyses (Hagger et al., 2010; Carter et al., 2015) and replication project (Hagger et al., 2016) indicate that moderators are present and stubstantial. Thus, future research could use the crossover design employed in Chapter 4 to test the various state and trait moderators in order to explain why self-control performance in consecutive task is limited in some individuals but not in others.

6.3.3 Ego depletion in everyday life

Research on ego depletion was set out to investigate self-control phenomena such as an unhealthy diet, procrastination, prejudice, or taking drugs that have wide individual and societal implications. In order to approach these everyday behaviors, which occur at least in some individuals every day, the dual-task paradigm was developed to capture limited self-control effects in laboratory tasks that could be transferred to everyday behavior associated with self-control. However, as tempting as this research might seem to investigate self-control with high internal validity, it is not clear to which extend laboratory tasks such as the Stroop task reflect the self-control challenges individuals face in their daily lives. Specifically, most definitions of self-control hypothesize that a situation requiring self-control is characterized by incompatible motivations, where doing something desired with positive short-term consequences conflicts with a long-term goal (Hofmann, Friese, & Strack, 2009). However, most research on ego depletion assumes that motivation to control oneself in laboratory tasks such as watching a video and controlling emotions along the way or performing the Stroop task is high without checking it. This is especially important for models that explain ego depletion due to motivational shifts such as the MAPM or the OCM since these models assume that motivation to control oneself or have-to goals are high in the beginning and decline in the course of the experiment. However, if the laboratory tasks do not conflict with long-term goals, performance differences in conditions cannot be attributed to limited self-control or to self-control itself but rather to a lack of motivation to perform well in an artificial task. Thus, future research on ego depletion or, more general, on the limits of self-control in daily life should incorporate more strongly settings that directly draw on situations in daily life where individuals indicate to exert self-control; for example, dieting (Geisler, Kleinfeldt, & Kubiak, 2016), procrastination (Reinecke, Hartmann, & Eden, 2014), prejudice (Inzlicht, McKay, & Aronson, 2006), and smoking (Heckman et al. 2012), as examples for the self-control failures listed in the beginning of this section. These behaviors are difficult to experimentally manipulate – from a technical as well as from an ethical point of view – and, thus, laboratory research should be complemented by the use of ambulatory assessment.

A good example for this research avenue is a study by Hofmann, Baumeister, Foerster, and Vohs (2012) in which participants had to report their desires six times a day by indicating the domain in which a desire occurred, the strength of this desire, if it conflicted with goals and if it was enacted or resisted. Importantly, the majority of desires were unproblematic, which highlights the difficulty to assess specific behaviors such as dieting or smoking without assessing whether these behavior conflict with long-term goals (i.e., weight loss or smoking cessation). Moreover, Hofmann, Vohs, and Baumeister (2012) reported evidence for the ego depletion effect in daily life, in that participants were less likely to resist a conflicted desire, the more frequent and recent prior acts of self-control had been. Clearly more research is needed but these studies demonstrate the usefulness of ambulatory assessment as a complementary method to laboratory research to study the boundaries and conditions of self-control and its limitations in daily life.

6.4 Concluding remarks

This dissertation made several contributions to research on self-control. In one line of research, Chapter 2 and 3 demonstrated both in a laboratory (Chapter 2) and ambulatory setting (Chapter 3) that positive affect influenced self-control – conceptualized as a trade-off between stability and flexibility – in favor of heightened flexibility. This lead to differential effects of positive affect on self-control performance, depending on whether the heightened flexibility is beneficial for the actual task at hand (i.e., by facilitating to switch to self-control demands) or harmful (i.e., by hampering to shield from distracting information).

These results also shed light on another line of research – whether performance decrements in consecutive self-control tasks can be understood as the consequence of a limited resource or as a consequence of switch costs. Chapter 2 demonstrated that the differential effects of positive affect cannot be explained by a limited-resource account (i.e., SMSC) but can be better explained by a cognitive control model (i.e., CTM).

In a third line of research, the ambiguous results regarding the size of the ego depletion effect reported in recent meta-analyses (Hagger et al., 2010; Carter et al., 2015) as well as in a preregistered replication project (Hagger et al., 2016) was put to test in Chapter 4 using a crossover design that enables researcher to study limited self-control performance within individual with high statistical power. Since the null effects of this crossover design are rather sobering in regard to the existence of ego depletion, Chapter 5 aimed at ending on a high note by providing a research that is able to capture limited self-control that is associated with goals set by the participants themselves instead of inferred by performance decrements.

Taken together, despite the ambiguous evidence regarding the ego depletion effect, the findings of this dissertation will hopefully contribute to newly inspired research that uses rigorous methods and research designs that are more suited to capture the underlying processes as well as the behaviors research on ego depletion was set out to investigate originally.

References

Abelson, R. P., & Prentice, D. A. (1997). Contrast tests of interaction hypotheses. *Psychological Methods, 2*, 315–328. doi:10.1037/1082-989X.2.4.315

Aiken, L.S. & West, S.G. (1991). *Multiple regression: Testing and interpreting interactions*. Thousand Oaks: Sage Publications.

Alberts, H., Martijn, C., Greb, J., Merckelbach, H., & 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

Alberts, H., Martijn, C., Nievelstein, F., Jansen, A., & de Vries, N. (2008). Distracting the self: Shifting attention prevents ego depletion. *Self and Identity, 7*, 322–334. doi:10.1080/15298860801987583

Alberts, H. J. E. M., Martijn, C., & de Vries, N. K. (2011). Fighting self-control failure: Overcoming ego depletion by increasing self-awareness. *Journal of Experimental Social Psychology, 47*, 58–62. doi:10.1016/j.jesp.2010.08.004

American Psychology Association (2012). Stress in America: Our health at risk. Retrieved July 28, 2016, from http://www.apa.org/news/press/releases/stress/2011/final-2011.pdf

Aspinwall, L. G. (1998). Rethinking the role of positive affect in self-regulation. *Motivation and Emotion*, 22, 1-32. doi:10.1023/A:1023080224401

Baayen, R.H. & Milin, P. (2010). Analyzing reaction times. *International Journal of Psychological Research, 3*, 12–28.

Baldwin, S.A., Imel, Z.E., Braithwaite, S., & Atkins, D.C. (2014). Analyzing multiple outcomes using multivariate multilevel models. *Journal of Consulting and Clinical Psychology, 82*, 920-930. doi:10.1037/a0035628

Baumeister, R. F. (1997). Evil: Inside human violence and cruelty. New York: Freeman.

Baumeister, R. F. (2001). 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. (2014). Self-regulation, ego depletion, and inhibition. Neuropsychologia, 65, 313-319. doi:10.1016/j.neuropsychologia.2014.08.012

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., & Heatherton, T. F. (1996). Self-regulation failure: An overview. *Psychological Inquiry, 7*, 1-15. doi:10.1207/s15327965pli0701_1

Baumeister, R. F, Heatherton, T. F, & Tice, D. M. (1994). *Losing control: How and why people fail at self-regulation*. San Diego, CA: Academic Press.

Baumeister, R. F., Masicampo, E. J., & DeWall, C. N. (2009). Prosocial benefits of feeling free: Disbelief in free will increases aggression and reduces helpfulness. *Personality and Social Psychology Bulletin, 35*, 260–268. doi:10.1016/j.concog.2012.04.004

Baumeister, R.F., & Tierney, J. (2011). *Willpower: Rediscovering the greatest human strength*. New York: Penguin Press.

Baumeister, R. F. & Vohs, K. D. (2007). Self-regulation, ego depletion, and motivation. *Social and Personality Psychology Compass, 1*, 1-14. doi:10.1111/j.1751-9004.2007.00001.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

Berkman, E. T., Falk, E. B., & Lieberman, M. D. (2011). In the trenches of real-world self-control: Neural correlates of breaking the link between craving and smoking. *Psychological Science, 22*, 498-506. doi:10.1177/0956797611400918

Bertrams, A. & Dickhäuser, O. (2009). Messung dispositioneller Selbstkontroll-Kapazität: Eine deutsche Adaptation der Kurzform der Self-Control Scale (SCS-K-D) [Measuring dispositional self-control capacity: A German adaptation of the short form of the Self-Control Scale (SCS-K-D)]. *Diagnostica, 55*, 2-10. doi:10.1026/0012-1924.55.1.2

Boksem, M.A.S. & Tops, M. (2008). Mental fatigue: Costs and benefits. *Brain Research Reviews, 59*, 125-139. doi:10.1016/j.brainresrev.2008.07.001

Boksem, M. A. S., Meijman, T. F., & Lorist, M. M. (2006). Mental fatigue, motivation and action monitoring. *Biological Psychology, 72*, 123–132. doi: 10.1016/j.biopsycho.2005.08.007

Bolger, N., & Laurenceau, J.-P. (2013). *Intensive longitudinal methods: An introduction to diary and experience sampling research*. New York: Guilford.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. New York, NY: John Wiley & Sons.

Botvinick, M., Nystrom, L.E., Fissell, K., Carter, C.S., & Cohen, J.D. (1999). Conflict monitoring versus selection-for-action in anterior cingulate cortex. *Nature*, *402*, 179–181. doi:10.1038/46035

Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*, 624-652. doi:10.1037/0033-295X.108.3.624

Bruyneel, S. D., Dewitte, S., Franses, P. H., & Dekimpe, M. G. (2009). I felt low and my purse feels light: Depleting mood regulation attempts affect risk decision making. *Journal of Behavioral Decision Making, 22*, 153–170. doi:10.1002/bdm.619

Burger, N., Charness, G., & Lynham, J. (2010). Field and online experiments on self-control. *Journal of Economic Behavior and Organization, 77,* 393-404. doi:10.1016/j.jebo.2010.11.010

Bush, G., Shin, L.M., Holmes, J., Rosen, B.R., & Vogt, B.A. (2003). The Multi-Source Interference Task: Validation study with fMRI in individual subjects. *Molecular Psychiatry, 8*, 60–70. doi: 10.1038/sj.mp.4001217

Carter, E.C., Kofler, L.M., Forster, D.E., & McCullough, M.E. (2015). A series of meta-analytic tests of the depletion effect: Self-control does not seem to rely on a limited resource. *Journal of Experimental Psychology: General, 144*, 796-815. doi:10.1037/xge0000083

Carter, E.C. & McCullough, M.E. (2013). Publication bias and the limited strength model of self-control: has the evidence for ego depletion been overestimated? *Behavioral and Brain Sciences, 36*, 683–684. doi: 10.1017/S0140525X12003196

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

Carver, C. S. (2003). Pleasure as a sign you can attend to something else: Placing positive feelings within a general model of affect. *Cognition & Emotion, 17*, 241-261. doi:10.1080/02699930302294

Chiew K. S. & Braver T. S. (2011). Positive affect versus reward: emotional and motivational influences on cognitive control. *Frontiers in Psychology, 2*, 1-10. doi:10.3389/fpsyg.2011.00279.

Ciarocco, N.J., Twenge, J.M., Muraven, M., Tice, D.M. (2004). Measuring state self-control: Reliability, validity, and correlations with physical and psychological stress. Monmouth, NJ: Monmouth University; unpublished

Clarkson, J. J., Hirt, E. R., Jia, L., & Alexander, M. B. (2010). When perception is more than reality: The effects of perceived versus actual resource depletion on self-regulatory behavior. Journal of Personality and Social Psychology, 98, 29-46. doi: http://dx.doi.org/10.1037/0022-3514.57.2.261.

Cleary, T. J. (2006). The development and validation of the Self-Regulation Strategy Inventory—Self-Report. *Journal of School Psychology, 44*, 307-322. doi:10.1016/j.jsp.2006.05.002

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple regression/correlation analysis for the behavioral sciences*, 3rd ed. Hillsdale: Erlbaum.

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

Crescioni, A. W., Ehrlinger, J., Alquist, J. L., Conlon, K. E., Baumeister, R. F., Schatschneider, C., & Dutton, G. R. (2011). High trait self-control predicts positive health behaviors and success in weight loss. *Journal of Health Psychology, 16*, 750–759

Dahm, T.; Neshat-Doost, H.T., Golden, A., J., Horn, E., Hagger, M., & Dalgleish, T. (2011). Age shall not weary us: Deleterious effects of self-regulation depletion are specific to younger adults. *PLoS One, 6*(10): e26351. doi:10.1371/journal.pone.0026351

Dang, J., Dewitte, S., Mao, L., Xiao, S., & Shi, Y. (2013). Adapting to an initial self-regulatory task cancels the ego depletion effect. *Consciousness and Cognition, 22*, 816–821. doi:10.1016/j.concog.2013.05.005

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:DOI 10.1016/j.jesp.2010.05.023

Derrick, J.L. (2012). Energized by television: familiar fictional worlds restore self-control. *Social Psychology and Personality Science, 4*, 299–307. doi: 10.1177/1948550612454889

DeWall, C. N., Baumeister, R. F., Gailliot, M. T., & Maner, J. K. (2008). Depletion makes the heart grow less helpful: Helping as a function of self-regulatory energy and genetic relatedness. *Personality & Social Psychology Bulletin, 34*, 1653-1662. doi:10.1177/0146167208323981

Dewitte S., Bruyneel S., & Geyskens K. (2009). Self-regulating enhances self-control in subsequent consumer decisions involving similar response conflicts. *Journal of Consumer Research*, *36*, 394-405. doi:10.1086/598615

De Witt Huberts, J. C., Evers, C., & De Ridder, D. T. D. (2012). License to sin: Self-licensing as a mechanism underlying hedonic consumption. *European Journal of Social Psychology, 42*, 490–496. doi:10.1002/ejsp.861

De Witt Huberts, J. C., Evers, C., & De Ridder, D. T. D. (2014). "Because I am worth it": A theoretical framework and empirical review of a justification-based account of self-regulation failure. *Personality and Social Psychology Review, 18*, 119-138. doi: 10.1177/1088868313507533

Dreisbach, G. (2006). How positive affect modulates cognitive control: the costs and benefits of reduced maintenance capability. *Brain and Cognition, 60*, 11-19. doi:10.1016/j.bandc.2005.08.003

Dreisbach, G., & Goschke, T. (2004). How positive affect modulates cognitive control: reduced perseveration at the cost of increased distractibility. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *30*, 343-353. doi:10.1037/0278-7393.30.2.343

Dvorak, R.D. & Simons, J.S. (2009). Moderation of resource depletion in the self-control strength model: Differing effects of two modes of self-control. *Personality and Social Psychology Bulletin, 35*, 572–583. doi:10.1177/0146167208330855

Eisinga, R., Te Grotenhuis, M., & Pelzer, B. (2012). The reliability of a two-item scale: Pearson, Cronbach or Spearman-Brown? *International Journal of Public Health, 58*, 637–642. doi:10.1007/s00038-012-0416-3

Ender, P. E. (2013). *sgmediation: Sobel-Goodman mediation tests*. UCLA Academic Technology Services Statistical Consulting Group. Retrieved from http://www.ats.ucla.edu/stat/stata/ado/analysis/.

Fedorikhin, S., & Patrick, V. M. (2010). Positive mood and resistance to temptation: The interfering influence of elevated arousal. *Journal of Consumer Research, 37*, 698-711. doi:10.1086/655665

Fishbach, A., & Labroo, A. A. (2007). Be better or be merry: How mood affects self-control. *Journal of Personality and Social Psychology*, *93*, 158–173. doi:10.1037/0022-3514.93.2.158

Fishbach, A., Zhang, Y., & Trope, Y. (2010). Counteractive evaluation: Asymmetric shifts in the implicit value of conflicting motivations. *Journal of Experimental Social Psychology, 46*, 29-38. doi:10.1016/j.jesp.2009.09.008

Fredrickson, B. L. (2004). The broaden-and-build theory of positive emotions. *Philosophical Transactions-Royal Society of London Series B Biological Sciences*, *359*, 1367–1378. doi:10.1098/rstb.2004.1512

Fredrickson, B. L., & Branigan, C. (2005). Positive emotions broaden the scope of attention and thought-action repertoires. *Cognition & Emotion, 19*, 313-332. doi:10.1080/02699930441000238

Freund, P. A., Kuhn, J.-T.,& Holling, H. (2011). Measuring current achievement motivation with the QCM: Short form development and investigation of measurement invariance. *Personality and Individual Differences, 51*, 629-634. doi:10.1016/j.paid.2011.05.033

Friese, M., Hofmann, W., & Wänke, M. (2008). When impulses take over: Moderated predictive validity of explicit and implicit attitude measures in predicting food choice and consumption behaviour. *British Journal of Social Psychology, 47*, 397-419. doi: 10.1348/014466607X241540

Friese, M., Messner, C., & Schaffner, Y. (2012). Mindfulness meditation counteracts self-control depletion. *Consciousness and Cognition, 21*, 1016–1022. doi:10.1016/j.concog.2012.01.008

Fujita, K. (2011). On conceptualizing self-control as more than the effortful inhibition of impulses. *Personality and Social Psychology Bulletin, 15*, 352-366. doi: 10.1177/1088868311411165

Kinder, D. C, & Block, J. (1989). The role of ego-control, ego-resiliency, and IQ in delay of gratification in adolescence. *Journal of Personality and Social Psychology, 57*, 1041-1050. doi:10.1037/0022-3514.57.6.1041

Funder, D. C, Block, J. H., & Block, J. (1983). Delay of gratification: Some longitudinal personality correlates. *Journal of Personality and Social Psychology, 44*, 1198-1213. doi:10.1037/0022-3514.44.6.1198.

Gable, P. A., and Harmon-Jones, E. (2008). Approach-motivated positive affect reduces breadth of attention. *Psychological Science, 19*, 476–482. doi:10.1111/j.1467-9280.2008.02112.x

Gailliot, 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

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

Gailliot, M., Peruche, B. M., Plant, E. A., & Baumeister, R. F. (2009). Stereotypes and prejudice in the blood: Sucrose drinks reduce prejudice and stereotyping. *Journal of Experimental Social Psychology*, *45*, 288-290. doi:10.1016/j.jesp.2008.09.003

Geisler, F. C. M., Kleinfeldt, A, & Kubiak, T. (2016). Restrained eating predicts effortful self-control as indicated by heart rate variability after ego depletion. *Appetite, 96*, 502-508. doi: 10.1016/j.appet.2015.10.020

Geisler, F. C. M., & Kubiak, T. (2009). Heart rate variability predicts self-control in goal pursuit. *European Journal of Personality, 23*, 623-633. doi:10.1002/per.727

Gergelyfi, M., Jacob, B., Olivier, E., & Zénon, A. (2015). Dissociation between mental fatigue and motivational state during prolonged mental activity. Frontiers in Behavioral Neuroscience, 9. doi: 10.3389/fnbeh.2015.00176

Gopher. D. & Sanders, A.F. (1984). S-Oh-R: Oh stages! Oh resources! In W. Printz & A.F. Sanders (Eds.). *Cognition and Motor Processes*. Berlin: Springer-Verlag.

Goschke, T. (2003). Voluntary action and cognitive control from a cognitive neuroscience perspective. In S. Maasen, W. Prinz,& G. Roth (Hrsg.), *Voluntary action: Brains, minds, and sociality* (pp.49-85). Oxford: Oxford University Press.

Gottfredson, M. R., & Hirschi, T. (1990). *A general theory of crime*. Stanford, CA: Stanford University Press

Hagger, M. S., & Chatzisarantis, N. L. (2013). The sweet taste of success: the presence of glucose in the oral cavity moderates the depletion of self-control resources. *Personality & Social Psychology Bulletin, 39*, 28-42. doi:10.1177/0146167212459912

Hagger, M. S., Chatzisarantis, N. L. D., Alberts, H., Anggono, C. O., Birt, A., Brand, R., … Rentzsch, K., … Zweinenberg, M. (2016). A multi-lab pre-registered replication of the ego-depletion effect. *Perspectives on Psychological Science, 11*, 546-573. doi:10.1177/1745691616652873

Hagger, M. S., Wood, C., Stiff, C., & Chatzisarantis, N. L. D. (2010). Ego-depletion and the strength model of self-control: A meta-analysis. *Psychological Bulletin, 136*, 496-525. doi:10.1037/a0019486

Halali, E., Bereby-Meyer, Y., & Meiran, N. (2014). Between self-interest and reciprocity: the social bright side of self-control failure. *Journal of Experimental Psychology. General, 143*, 745–754. doi:10.1037/a0033824

Hamaker, E. L. (2012). Why researchers should think "within-person": A paradigmatic rationale. In M. R. Mehl & T. S. Conner (Eds.), *Handbook of research methods for studying daily life* (pp. 43-61). New York: Guilford Press.

Harmon-Jones, E., Gable, P. A., & Price, T. F. (2012). The influence of affective states varying in motivational intensity on cognitive scope. *Frontiers in Integrative Neuroscience, 6*, 1-5. doi:10.3389/fnint.2012.00073

Hayter, A. J. (1986). The maximum familywise error rate of Fisher's least significant difference test. *Journal of the American Statistical Association, 81*, 1000-1004. doi:10.2307/2289074

Heckman, B. W., Ditre, J. W., & Brandon, T. H. (2012). The restorative effects of smoking upon self-control resources: A negative reinforcement pathway. *Journal of Abnormal Psychology, 121*, 244–249. doi:10.1037/a0023032

Herman, C. P., & Polivy, J. (1975). *Anxiety, restraint and eating behavior*. Journal of Abnormal Psychology, 84, 666-672.

Hockey, G. R. J. (1997). Compensatory control in the regulation of human performance under stress and high workload: A cognitive-energetical framework. *Biological Pschology, 45*, 73-93. doi:10.1016/S0301-0511(96)05223-4

Hockey, G. R. J. (2013). *The psychology of fatigue*. Cambridge University Press.

Hockey, G. R. J. & Earle, F. (2006). Control over the scheduling of simulated office work reduces the impact of workload on mental fatigue and task performance. *Journal of Experimental Psychology: Applied, 12*, 50-65. doi:10.1037/1076-898X.12.1.50

Hofmann, W., Baumeister, R. F., Foerster, G., & Vohs, K. D. (2012). Everyday temptations: An experience sampling study of desire, conflict, and self-control. *Journal of Personality and Social Psychology, 102*, 1318-1335. doi:10.1037/a0026545

Hofmann, W., Deutsch, R., Lancaster, K., & Banaji, M. R. (2009). Cooling the heat of temptation: Mental self-control and the automatic evaluation of tempting stimuli. *European Journal of Social Psychology*, 40, 17–25. doi:10.1002/ejsp.708

Hofmann, W., Friese, M., & Strack, F. (2009). Impulse and self-control from a dual-systems perspective. *Perspectives on Psychological Science, 4*, 162-176. doi: 10.1111/j.1745-6924.2009.01116.x

Hofmann, W. & Fisher, R. R. (2012). How guilt and pride shape subsequent self-control. *Social Psychological and Personality Science, 3*, 682-690. doi:10.1177/1948550611435136

Hofmann, W., Vohs, K. D., & Baumeister, R. F. (2012). What people desire, feel conflicted about, and try to resist in everyday life. *Psychological Science, 23*, 582–588. doi:10.1177/0956797612437426

Holroyd, C. B., & Coles, M. G. H. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review, 109*, 679–709. doi:10.1037/0033-295X.109.4.679

Hossain, J. L., Reinish, L. W., Kayumov, L., Bhuiya, P., & Shapiro, C. M. (2003). Underlying sleep pathology may cause chronic high fatigue in shift-workers. Journal of Sleep Research, 12, 223–230. doi: 10.1046/j.1365-2869.2003.00354.x

Hox, J. (2010). Multilevel analysis: *Techniques and applications* (2nd ed.). New York, NY: Routledge.

Imhoff, R., Schmidt, A.F., & Gerstenberg, F. (2014). Exploring the interplay of trait self-control and ego depletion: Empirical evidence for ironic effects. *European Journal of Personality, 28*, 413-424. doi:10.1002/per.1899

Inzlicht, M., Gervais, W., & Berkman, E. (2015). Bias-correction techniques alone cannot determine whether ego depletion is different from zero: Commentary on Carter, Kofler, Forster, & McCullough, 2015. *Available at SSRN: http://ssrn.com/abstract=2659409 or http://dx.doi.org/10.2139/ssrn.2659409*

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

Inzlicht, M., McKay, L., & Aronson, J. (2006). Stigma as ego depletion: How begin the target of prejudice affects self-control. *Psychological Science, 17*, 262-269.

Inzlicht, M., & Schmeichel, B. J. (2012). What is ego depletion? Toward a mechanistic revision of the resource model of self-control. *Perspectives on Psychological Science, 7*, 450-463. doi:10.1177/1745691612454134

Inzlicht, M., Schmeichel, B. J., & Macrae, C. N. (2014). Why self-control seems (but may not be) limited. *Trends in Cognitive Sciences, 14*, 127-133. doi:10.1016/j.tics.2013.12.009

Isen, A. M. (2002). Missing in action in the AIM: Positive affect's facilitation of cognitive flexibility, innovation, and problem solving. *Psychological Inquiry, 13*, 57-65.

Job, V., Bernecker, K., Miketta, S., & Friese, M. (2015). Implicit theories about willpower predict the activation of a rest goal following self-control exertion. *Journal of Personality and Social Psychology, 109*, 694-706. doi: 10.1037/pspp0000042

Job, V., Dweck, C. S., & Walton, G. M. (2010). Ego depletion: Is it all in your head? Implicit theories about willpower affect self-regulation. *Psychological Science, 21*, 1686-1693. doi:10.1177/0956797610384745

Job, V., Walton, G. M., Bernecker, K., & Dweck, C. S. (2013). Beliefs about willpower determine the impact of glucose on self-control. *Proc Natl Acad Sci U S A, 110*, 14837-14842. doi:10.1073/pnas.1313475110

Kahneman. D. (1973). *Attention und effort*. Englewood Cliffs, NJ: Prentice-Hall.

Keppel, G. (1991). *Design and analysis: A researcher's handbook* (3rd. ed.). Englewood Cliffs, NJ: Prentice Hall.

Kelly, E. L., & Conley, J. J. (1987). Personality and compatibility: A prospective analysis of marital stability and marital satisfaction. *Journal of Personality and Social Psychology, 52*, 27-40.

Kerns, J. G., Cohen, J. D., MacDonald, A. W., III, Cho, R. Y., Stenger, V. A., & Carter, C. S. (2004). Anterior cingulate, conflict monitoring, and adjustments in control. *Science, 303*, 1023–1026. doi:10.1126/science.1089910

Knight, R. T., & Stuss, D. T. (2002). Prefrontal cortex: The present and the future. In D. T. Stuss & R. T. Knight (Eds.), *Principles of frontal lobe function* (pp. 573–597). New York: Oxford University Press.

Kool, W., & Botvinick, M. (2014). A labor/leisure trade-off in cognitive control. *Journal of Experimental Psychology: General, 143*, 131-141. doi: 10.1037/a0031048

Kool, W., McGuire, J. T., Rosen, Z. B., & Botvinick, M. M. (2010). Decision making and the avoidance of cognitive demand. *Journal of Experimental Psychology: General, 139*, 665-682. doi:10.1037/a0020198

Krause, M. S., & Howard, K. I. (2003). What random assignment does and does not do. *Journal of Clinical Psychology, 59*, 751–766. doi:10.1002/ jclp.10170.

Kubiak, T. & Stone, A.A. (2012). Ambulatory monitoring of biobehavioral processes in health and disease [Editorial to the Special Issue on Ambulatory Monitoring]. *Psychosomatic Medicine, 74*, 325-326. doi:10.1097/PSY.0b013e31825878da

Kühberger, A., Fritz, A., & Scherndl, T. (2014). Publication bias in psychology: a diagnosis based on the correlation between effect size and sample size. *PloS One, 9*, e105825. doi:10.1371/journal.pone.0105825

Kurzban, R., Duckworth, A., Kable, J., & Myers, J. (2013). An opportunity cost model of subjective effort and task performance. *Behavioral and Brain Sciences, 36*, 661–726. doi:10.1017/S0140525X12003196

Labroo, A. A., & Patrick, V. M. (2009). Psychological distancing: Why happiness helps you see the big picture. *Journal of Consumer Research, 35*, 800–809. doi:10.1086/593683

Lange, F., & Eggert, F. (2014). Sweet delusion. Glucose drinks fail to counteract ego depletion. *Appetite, 75*, 54-63. doi:10.1016/j.appet.2013.12.020

Magen, E. & Gross, J. J. (2007). Harnessing the need for immediate gratification: Cognitive reconstrual modulates the reward value of temptations. *Emotion, 7*, 415–428. doi:10.1037/1528-3542.7.2.415

Marcora, S. M., Staiano, W., and Manning, V. (2009). Mental fatigue impairs physical performance in humans. *Journal of Applied Physiology, 106*, 857–864. doi: 10.1152/japplphysiol.91324.2008

Martijn, C., Alberts, H., Merckelbach, H., Havermans, R., Huijts, A., & Vries, N. K. de. (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, L. L., &Tesser, A. (1989). Towards a motivational and structural theory of ruminative thought. In J. S. Uleman & J. A. Bargh (Eds.), *Unintended thought* (pp. 306-326). New York: Guilford Press.

McCall, A. L. (2004). Cerebral glucose metabolism in diabetes mellitus. *European Journal of Pharmacology, 490*, 147-158. doi:http://dx.doi.org/10.1016/j.ejphar.2004.02.052

Messier, C. (2004). Glucose improvement of memory: a review. *European Journal of Pharmacology, 490*, 33-57. doi:10.1016/j.ejphar.2004.02.043

Miller, E. K., & Cohen, J. D. (2001). An integrative theory of prefrontal cortex function. *Annual Review of Neuroscience*, *24*, 167-202. doi:10.1146/annurev.neuro.24.1.167

Mischel, W., Ebbesen, E. B., & Raskoff Zeiss, A. (1972). Cognitive and attentional mechanisms in delay of gratification. *Journal of Personality and Social Psychology*, *21*, 204–218. doi:10.1037/h0032198

Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H., … Caspi, A. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. Proceedings of the National Academy of Sciences, 108, 2693–2698, doi:10.1073/pnas.1010076108

Molden, D. C., Hui, C. M., Scholer, A. A., Meier, B. P., Noreen, E. E., D'Agostino, P. R., & Martin, V. (2012). Motivational versus metabolic effects of carbohydrates on self-control. *Psychological Science, 23*, 1137-1144. doi:10.1177/0956797612439069

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

Moreno, S. G., Sutton, A. J., Turner, E. H., Abrams, K. R., Cooper, N. J., Palmer, T. M., & Ades, a E. (2009). Novel methods to deal with publication biases: secondary analysis of antidepressant trials in the FDA trial registry database and related journal publications. *BMJ (Clinical Research Ed.), 339*, b2981. doi:10.1136/bmj.b2981

Morris, S. B., & DeShon, R. P. (2002). Combining effect size estimates in meta-analysis with repeated measures and independent-group designs. *Psychological Methods, 7*, 105–125. doi:10.1037/1082-989X.7.1.105

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:I0.I037//0033-2909.126.2.247

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

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

Navon, D. (1984). Resources--a theoretical soup stone? *Psychological Review, 91*, 216–234. doi:10.1037//0033-295X.91.2.216

Nealis, L. J., & van Allen, Z. M., & Zelenski, J. M. (2016). Positive Affect and Cognitive Restoration: Investigating the Role of Valence and Arousal. *PLoS ONE, 11*: e0147275. doi:10.1371/journal.pone.0147275

Neal, D. T., Wood, W., & Drolet, A. (2013). How do people adhere to goals when willpower is low? The profits (and pitfalls) of strong habits. *Journal of Personality and Social Psychology, 104*, 959–975. doi:10.1037/a0032626

Nordgren, L. F., & Chou, E. Y. (2011). The push and pull of temptation: The bidirectional influence of temptation on self-control. *Psychological Science*, *22*, 1386–1390. doi:10.1177/0956797611418349

Oaten, M., & Cheng, K. (2006). Improved self-control: The benefits of a regular program of academic study. *Basic and Applied Social Psychology, 28,* 1–16. doi:10.1207/s15324834basp2801_1

Quinn, J. M., Pascoe, A., Wood, W., & Neal, D. T. (2010). Can't control yourself? Monitor those bad habits. *Personality and Social Psychology Bulletin*, *36*, 499–511. doi:10.1177/0146167209360665

Peugh, J. L. (2010). A practical guide to multilevel modeling. Journal of School Psychology, 48, 85-112. doi:10.1016/j.jsp.2009.09.002

Prinsen, S., Evers, C., de Ridder, D. (2016). Oops I did it again: Examining self-licensing effects in a subsequent self-regulation dilemma. *Health and Well-Being, 8*, 104-126. doi:10.1111/aphw.12064

Reed, W. R., Florax, R. J. G. M., & Poot, J. (2015). A Monte Carlo analysis of alternative meta-analysis estimators in the presence of publication bias. *Economics Discussion Papers, 2015-9*. Retrieved from http://www.economicsejournal.org/economics/discussionpapers/2015-9

Reimers, S., & Stewart, N. (2007). Adobe Flash as a medium for online experimentation: A test of reaction time measurement capabilities. *Behavior Research Methods*, *39*, 365-370. doi:10.3758/BF03193004

Reinecke, L., Hartmann, T., & Eden, A. (2014). The guilty couch potato: The role of ego depletion in reducing recovery through media use. *Journal of Communication, 64*, 569-589. doi:10.1111/jcom.12107

Rheinberg, F., Vollmeyer, R., & Burns, B. D. (2001). FAM: Ein Fragebogen zur Erfassung aktueller Motivation in Lern- und Leistungssituationen [QCM: A questionnaire to assess current motivation in learning situations]. *Diagnostica, 47*, 57–66. doi: 10.1026//0012-1924.47.2.57

Rigoni, D., Kühn, S., Gaudino, G., Sartori, G., & Brass, M. (2012). Reducing self-control by weakening belief in free will. *Consciousness and Cognition, 21*, 1482-90. doi: 10.1016/j.concog.2012.04.004

Robinson, M. D., Schmeichel, B. J., & Inzlicht, M. (2010). A cognitive control perspective of self-control strength and its depletion. *Social and Personality Psychology Compass, 4*, 189-200. doi:10.1111/j.1751-9004.2009.00244.x

Ross, M., & Conway, M. (1986). Remembering one's own past: The construction of personal histories. In R. M. Sorrenlino & E. G. Higgins (Eds.). *Handbook of motivation and cognition: Foundations of social behavior* (pp. 122-144). New York: Ouilford Press.

Russell, M. A. H. (1971). *Cigarette smoking: Natural history of a dependence disorder*. British Journal of Medical Psychology, 44, 1-16.

Sanders, M. A., Shirk, S. D., Burgin, C. J., & Martin, L. L. (2012). The gargle effect: rinsing the mouth with glucose enhances self-control. *Psychological Science, 23*, 1470-1472. doi:10.1177/0956797612450034

Schimmack, U., & Rainer, R. (2002). Experiencing activation: Energetic arousal and tense arousal are not mixtures of valence and activation. *Emotion, 2*, 412–417. doi:10.1037//1528-3542.2.4.412

Schmeichel, B. J. (2007). Attention control, memory updating, and emotion 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., Harmon-Jones, C., & Harmon-Jones, E. (2010). Exercising self-control increases approach motivation. *Journal of Personality and Social Psychology, 99*, 162-173. doi:10.1037/a0019797

Schmeichel, B. J. & Vohs, K. D. (2009). Self-affirmation and self-control: Affirming core values counteracts ego depletion. Journal of Personality and Social Psychology, 13, 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

Segerstrom, S. C., & Nes, L. S. (2007). Heart rate variability reflects self-regulatory strength, effort, and fatigue. *Psychological Science, 18*, 275–281. doi:10.1111/j.1467-9280.2007.01888.x

Selya, A.S., Rose, J.S., Dierker, L.C., Hedeker, D., & Mermelstein R.J. (2012). A practical guide to calculating Cohen's f^2, a measure of local effect size, from PROC MIXED. Frontiers in Psychology. doi:10.3389/fpsyg.2012.00111

Shamosh, N. & Gray, J. (2007). The relation between fluid intelligence and self-regulatory depletion. *Cognition and Emotion, 21*, 1833–1843. doi: 10.1080/02699930701273658

Shoda, Y, Mischel, W., & Peake, P. K. (1990). Predicting adolescent cognitive and self-regulatory competencies from preschool delay of gratification: Identifying diagnostic conditions. *Developmental Psychology, 26*, 978-986. doi:http://dx.doi.org/10.1037/0012-1649.26.6.978

Shrout, P. E., & Lane, S. P. (2012). Psychometrics. In M. R. Mehl & T. S. Conner (Eds.), *Handbook of research methods for studying daily life* (pp. 302-321). New York: Guilford Press.

Šidák, Z. (1967). Rectangular confidence regions for the means of multivariate normal distributions. *Journal of the American Statistical Association, 62*, 626–633. doi:10.1080/01621459.1967.10482935.

Simmons, J.P., Nelson, L.D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science, 22*, 1359-1366. doi:10.1177/0956797611417632

Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2014). p-Curve and effect size: Correcting for publication bias using only significant results. *Perspectives on Psychological Science, 9*, 666–681. doi:10.1177/1745691614553988

Sripada, C., Kessler, D., & Jonides, J. (2014). Methylphenidate blocks effort-induced depletion of regulatory control in healthy volunteers. *Psychological Science, 22*, 1-8. doi:10.1177/0956797614526415

Sripada, C., Kessler, D., & Jonides, J. (2016). Sifting signal from noise with replication science. *Perspectives on Psychological Science, 11*, 576-578. doi:10.1177/1745691616652875

Stanley, T. D. (2008). Meta-regression methods for detecting and estimating empirical effects in the presence of publication selection. *Oxford Bulletin of Economics and Statistics, 70*, 103–127. doi:10.1111/j.1468-0084.2007.00487.x

Steyer, R., Schwenkmezger, P., Notz, P., & Eid, M. (1997). *Der Mehrdimensionale Befindlichkeitsfragebogen (MDBF).* [The multidimensional mood questionnaire - Manual]. Göttingen/Germany: Hogrefe.

Stillman, T.F., Tice, D.M., Fincham, F.D., & Lambert, N.M. (2009). The psychological presence of family improves self-control. *Journal of Social and Clinical Psychology, 28*, 498–529. doi:10.1521/ jscp.2009.28.4.498

Tangney, J.P., Baumeister, R.F., & Boone, A.L. (2004). High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. *Journal of Personality, 72*, 271-324. doi:10.1111/j.0022-3506.2004.00263.x

Thayer, R. E. (1989). *The biopsychology of mood and arousal.* New York: Oxford University Press.

Tice, D. M., Baumeister, R. F., Shmueli, D., & Muraven, M. (2007). Restoring the self: Positive affect helps improve self-control following ego depletion. *Journal of Experimental Social Psychology*, *43*, 379-384. doi:10.1016/j.jesp.2006.05.007

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

vanDellen, M. R., Shea, C. T., Davisson, E. K., Koval, C., & Fitzsimons (2014). Motivated misperception: Self-regulatory resources affect goal appraisals. *Journal of Experimental Social Psychology, 53*, 118-124. doi:10.1016/j.jesp.2014.03.007

van Veen, V., & Carter, C. S. (2002). The anterior cingulate as a conflict monitor: fMRI and ERP studies. *Physiology and Behavior, 77*, 477–482. doi:10.1016/S0031-9384(02)00930-7

Vohs, K. D., Baumeister, R. F., & Ciarocco, N. (2005). Self-control 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. (2004). Understanding self-regulation: An introduction. In R. F. Baumeister & K. D. Vohs (Eds.), *Handbook of self-regulation: Research theory, and applications* (pp. 1-9). New York, NY, US: Guilford Press.

Vohs, K. D., Glass, B. D., Maddox, W. T., & Markman, a. B. (2010). Ego depletion is not just fatigue: Evidence from a total sleep deprivation experiment. *Social Psychological and Personality Science*, *2*, 166–173. doi:10.1177/1948550610386123

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

Wan, E. W., & Sternthal, B. (2008). Regulating the effects of depletion through resource monitoring. *Personality and Social Psychology Bulletin, 34*, 47-60. doi:10.1177/0146167207306756

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

Wegner, D. M., Schneider, D. J., Carter, S. R., & White, T. L. (1987). Paradoxical effects of thought suppression. *Journal of Personality and Social Psychology, 53*, 5–13. doi:10.1037/0022-3514.53.1.5

Wenzel, M., Conner, T. S., & Kubiak, T. (2013). Understanding the limits of self-control: Positive affect moderates the impact of task switching on consecutive self-control performance. *European Journal of Social Psychology, 43*, 175-184. doi:10.1002/ejsp.1936

Wenzel, M., Kubiak, T., & Conner, T. S. (2014). Positive affect and self-control: Attention to self-control demands mediates the influence of positive affect on consecutive self-control. *Cognition and Emotion, 28*, 747-755. doi:10.1080/02699931.2013.851069

Westermann, R., Spiess, K., Stahl, G., & Hesse, F.W. (1996). Relative effectiveness and validity of mood induction procedures: a meta-analysis. *European Journal of Social Psychology*, *26*, 557-580. doi:10.1002/(SICI)1099-0992(199607)26:4<557::AID-EJSP769>3.0.CO;2-4

Wickens, C.D. (1984). Processing resources in attention. In R. Parasuraman & D.R. Davies (Eds.), *Varieties of Attention*, (pp. 63-102). New York, NY: John Wiley.

Wood, W., & Neal, D. T. (2007). A new look at habits and the habit-goal interface. *Psychological Review*, *114*, 843–863. doi:10.1037/0033-295X.114.4.843

Xu, X., Demos, K.E., Leahey, T.M., Hart, C.N., Trautvetter, J., Coward, P., et al. (2014). Failure to replicate depletion of self-control. *PLoS One, 9*. doi: 10.1371/journal.pone.0109950

Yik, M. S. M., Russell, J. A., & Barrett, L. F. (1999). Structure of self-reported current affect: Integration and beyond. *Journal of Personality and Social Psychology, 77*, 600–619. doi:10.1037/0022-3514.77.3.600