European Journal of Personality, *Eur. J. Pers.* (2019)
Published online in Wiley Online Library (wileyonlinelibrary.com) **DOI**: 10.1002/per.2208

Let There Be Variance: Individual Differences in Consecutive Self-control in a Laboratory Setting and Daily Life

MARIO WENZEL*, ZARAH ROWLAND, DANIELA ZAHN and THOMAS KUBIAK

Institute of Psychology, Johannes Gutenberg University Mainz, Germany

Abstract: The large body of research used to support ego-depletion effects is currently faced with conceptual and replication issues, leading to doubt over the extent or even existence of the ego-depletion effect. By using within-person designs in a laboratory (Study 1; 187 participants) and an ambulatory assessment study (Study 2; 125 participants), we sought to clarify this ambiguity by investigating whether prominent situational variables (such as motivation and affect) or personality traits can help elucidate when ego depletion can be observed and when not. Although only marginal ego-depletion effects were found in both studies, these effects varied considerably between individuals, indicating that some individuals experience self-control decrements after initial self-control exertion and others not. However, neither motivation nor affect nor personality traits such as trait self-control could consistently explain this variability when models were applied that controlled for variance due to targets and the depletion manipulation (Study 1) or days (Study 2) as well as for multiple testing. We discuss how the operationalization and reliability of our key measures may explain these null effects and demonstrate that alternative metrics may be required to study the consequences of the consecutive exertion of self-control. © 2019 European Association of Personality Psychology

Key words: self-control; self-regulation; ego depletion; individual differences

A lot has been written about ego depletion, the notion of impaired consecutive self-control due to limited self-control resources or limited motivation to self-control, and its surrounding crisis. This discussion has grown from initial foundational research (Muraven, Tice, & Baumeister, 1998; Baumeister, Bratslavsky, Muraven, & Tice, 1998), to the success years, culminating in an average effect size of $d = 0.62$ in a meta-analysis of 83 studies and 198 effect sizes (Hagger, Wood, Stiff, & Chatzisarantis, 2010), to the current replication and conceptual crisis. This crisis began with the reanalysis of the previously mentioned meta-analysis, which found null effects when correcting for small-study effects (Carter & McCullough, 2013; Carter & McCullough, 2014). Although the correction methods have been criticized (e.g. Cunningham & Baumeister, 2016), these null effects (or better, the effect sizes equivalent to zero; Wellek, 2010) were confirmed by a large multi-lab replication project (Hagger et al., 2016).

As a result, there is uncertainty as to whether or to what extent the ego-depletion phenomenon exists. Friese, Loschelder, Gieseler, Frankenbach, and Inzlicht (2018) provided an excellent summary of the current strengths and shortcomings in ego-depletion research, namely, (1) the absence of reverse ego-depletion effects in the literature, in that studies should have been published that show improved performance after initial self-control efforts if the true effect is equivalent to zero; (2) the particularly strong small-study effects in research on moderators and mediators of ego depletion; and (3) the small number of studies investigating ego depletion in daily life. In the present study, we empirically addressed these shortcomings highlighted by Friese et al. (2018) by employing a crossover design and an ambulatory assessment study. This design allowed us to investigate state and trait moderators of consecutive self-control, with the goal of explaining the heterogeneity of the current evidence regarding ego depletion.

REVERSE EGO DEPLETION

Research on ego depletion has focused mainly on decreased self-control performance after initial self-control execution. Although an important aspect of consecutive self-control research as well, only a few studies have directly examined improved consecutive self-control performance. For example, Savani and Job (2017) conducted four dual-task experiments and found significant reversed ego-depletion effects among Indian participants. The authors explain that this might possibly be due to cultural differences in implicit theories on the limited availability of self-control, with participants in India being less likely to believe that self-control is limited and more likely to believe that self-control exertion is energizing than participants in the USA. Moreover, Converse and DeShon (2009) built on the theory of learned industriousness

*Correspondence to: Mario Wenzel, Psychologisches Institut, Johannes Gutenberg-Universität Mainz, Binger Str. 14-16, D-55122 Mainz, Germany. E-mail: wenzelma@uni-mainz.de

This article earned Open Data and Open Materials badges through Open Practices Disclosure from the Center for Open Science: https://osf.io/tvyxz/wiki. The data and materials are permanently and openly accessible at https://osf.io/43fmu/. Author's disclosure form may also be found at the Supporting Information in the online version.

Handling editor: Erika Carlson
Received 5 November 2018
Revised 7 May 2019, Accepted 8 May 2019

(Eisenberger, 1992) and showed that more than one initial task can lead to increases instead of decreases in self-control performance in one's final task. According to Converse and DeShon (2009), this effect can be explained by the fact that exerting self-control does not necessarily lead to aversive states that individuals want to avoid but instead can also be reinforced by the experience of exerting effort, resulting in increased effort and performance. Although contradictory evidence exists (Vohs, Baumeister, & Schmeichel, 2013) that did not find improved self-control performance after multiple initial self-control tasks, Converse and DeShon (2009) also connect their findings to the counteractive self-control theory (Fishbach, Zhang, & Trope, 2010). According to this theory, individuals are not helpless when confronted with temptations but may instead actively reinforce the strength of their self-control goal to shield themselves from the allure of temptations. However, evidence for this effect in ego-depletion research is scarce, with only initial evidence provided by Wenzel, Zahn, Rowland, and Kubiak (2016). In this 2016 study, participants had to indicate how well they wanted to perform in the second task, which served as a measure of goal strength. The results demonstrated that initial self-control efforts led to stricter goal setting before the second task compared with the control condition but that participants in the depletion condition were less able to follow through with their intentions.

However, despite of the interesting and promising evidence, this line of research suffers from the same central problem that research on ego depletion suffers from in general: low statistical power. None of the three studies that explicitly investigated reversed ego-depletion effects (Converse & DeShon, 2009; Vohs et al., 2013; Savani & Job, 2017)—which included four to eight relevant effect sizes—exceeded an estimated power of 5% according to *p*-curve analyses (Simonsohn, Nelson, & Simmons, 2014), with a total estimated power of 5%, 90% confidence interval (CI) [5%, 8%], and no evidential value (the full results can be accessed under https://osf.io/43fmu/). This signals the need for experimental designs with high statistical power to understand why some individuals can increase self-control efforts after initial self-control executions and why others cannot. In the present study, we use a crossover design whereby participants first perform either the control or depletion condition, then 2 weeks later perform the other condition. This not only allows for a single within-person design with high statistical power but also allows ego depletion to be investigated on the within-person level. Thus, by using the crossover design, we can analyse individual differences in ego depletion, that is, how many individuals perform better, worse, or equally well in the depletion compared with the control condition.

MODERATORS OF EGO DEPLETION

A plethora of moderators have been investigated in ego-depletion research, with more than 100 moderator studies (Friese et al., 2018). Here, we will not repeat prior overviews of this wide array of moderators (Loschelder & Friese, 2016) but will focus only on the most prominent candidate moderators.

Moderators can be separated into state and trait variables. The two most prominent state moderators are motivation and affect. Initial evidence has found that the ego-depletion effect disappears when individuals are provided with sufficient incentives to perform well in the second task (Muraven & Slessareva, 2003), and Boucher and Kofos (2012) demonstrated that even the idea of an incentive was enough to counteract the ego-depletion effect. Comparable effects were observed when the second task was presented in a supportive manner that highlighted the individual's autonomy, to facilitate intrinsic motivation (Moller, Deci, & Ryan, 2006; Muraven, Gagné, & Rosman, 2008). Similarly, a mild induction of positive mood as an incentive can also decrease subsequent self-control limitations (Tice, Baumeister, Shmueli, & Muraven, 2007; Wenzel, Conner, & Kubiak, 2013), although contradictory evidence exists (Nealis, van Allen, & Zelenski, 2016). The role of motivation as a moderator of ego depletion is reflected in the observation that all alternative accounts of explaining ego depletion have been built in some way or another on motivational processes such as motivation shifts (Inzlicht & Schmeichel, 2012), opportunity costs (Kurzban, Duckworth, Kable, & Myers, 2013), justification (De Witt Huberts, Evers, & De Ridder, 2014), or a labour/leisure trade-off (Kool & Botvinick, 2014). However, these alternative accounts act on the assumption that the ego-depletion effect exists, and thus, evidence for the respective model demonstrates that an observed ego-depletion effect in the control depletion condition disappears in the experimental depletion condition. Given the uncertainty regarding the existence and extent of the ego-depletion effect in general, the empirical confirmation of these alternative accounts is therefore also called into question.

There are many trait moderators, such as construal level (Agrawal & Wan, 2009), self-control (Imhoff, Schmidt, & Gerstenberg, 2014), action orientation (Gröpel, Baumeister, & Beckmann, 2014), or implicit theories about willpower (Job, Dweck, & Walton, 2010; Job, Bernecker, Miketta, & Friese, 2015). However, few studies focusing on specific moderators have not compared multiple moderators, and most have a low estimated power, clearly indicating a need for additional research. Moreover, this evidence is sometimes ambiguous: while DeWall, Baumeister, Stillman, and Gailliot (2007) found that individuals with high levels of trait self-control showed reduced ego-depletion effects, Imhoff et al. (2014) reported the opposite, somewhat paradoxical, effect.

The crossover design employed in the present study whereby every participant performed both the depletion and control condition avoids averaging responses within experimental conditions, making it well suited to investigating state and trait moderators with high statistical power. Moreover, this study design allows for a direct examination of why a given individual performs worse after an initial self-control exertion, with one possible explanation being lower levels of self-reported motivation and affect when performing in the depletion as opposed to the control condition.

DOI: 10.1002/per

Additionally, besides the demographic variables of gender (Meece & Painter, 2008; Nolen-Hoeksema & Corte, 2004), which is related to several aspects of momentary self-control, and age (Dahm et al., 2011), for which early evidence suggests that only young individuals under 25 years are susceptible to depletion effects, we included a number of self-control relevant traits (i.e. self-control, impulsivity, behavioural inhibition system, and implicit theories about willpower) as the most promising candidates for domain-specific moderators. We additionally included the Big Five personality traits as general moderators in order to test whether the self-control-related trait moderators can explain specific variance in consecutive self-control beyond the explanatory power of the Big Five personality traits.

EVIDENCE FOR EGO DEPLETION IN DAILY LIFE

Despite these questions regarding the robustness of the ego-depletion effect, the possible existence of moderators that may help researchers to understand this ambiguous evidence, and alternative explanations such as small-study effects and publication bias, ego-depletion research also suffers from relying almost solely on the dual-task paradigm and from employing artificial tasks. Artificial tasks do not replicate the self-control demands that individuals experience in daily life, through which many people clearly demonstrate a widespread capacity for self-control across different domains and over a relatively long time frame (e.g. Hughes, Keely, & Naud, 2004; De Vet, Nelissen, Zeelenberg, & De Ridder, 2013). Although Baumeister and colleagues originally set out to conduct research on ego depletion to explain self-control failures observed in daily life (in areas such as weight control, binge eating, impulse buying, violence, and substance abuse; Baumeister & Heatherton, 1996), this line of research mainly rested on the dual-task paradigm. Friese et al. (2018) and Baumeister, Tice, and Vohs (2018) both provide succinct overviews of the few studies that have tested the basic premise of ego depletion (i.e. impaired self-control performance after prolonged exertion of self-control) in ecologically valid settings outside of the dual-task setting. However, few such studies exist, and evidence is contradictory, with some studies providing evidence for real-world ego-depletion effects (e.g. Dai, Milkman, Hofmann, & Staats, 2015; Sievertsen, Gino, & Piovesan, 2016) and others not (e.g. O'Connell et al., 2008; Randles, Harlow, & Inzlicht, 2017). Moreover, aside from a recent study—which found that subjective depletion over two 12-hour nursing shifts was not associated with perceived work demands or physical energy expended but with perceived control and reward associated with the work—studies investigating which and when individuals are more prone to ego depletion in daily life do not exist (Johnston et al., 2018).

THE PRESENT STUDY

The present study contributes to filling this gap by extending our laboratory findings of the short-term depletion effect and its moderators to the everyday lives of individuals. We conducted two studies, a laboratory study (Study 1) and an ambulatory assessment study (Study 2), to investigate state and trait moderators of consecutive self-control and thereby explain the heterogeneity of the current evidence regarding ego depletion.

STUDY 1

We affirm that we have reported all measures, manipulations, and exclusions in the present study, regardless of whether they support our hypotheses. Both studies were approved by the applicable ethics committee. The data reported in Study 1 were used in another publication (Wenzel, Lind, Rowland, Zahn, & Kubiak, in press), which investigated (1) the usefulness of the crossover design for studying the ego-depletion effect, (2) the robustness of the ego-depletion effect, (3) whether the difficulty and fatigue mediate the effect of the ego-depletion effect, and (4) whether an overt assessment of individuals' perceived limited self-control can induce ego depletion as shown in prior research (Job et al., 2010). The last research question was examined by an additional experimental between-person factor, whereby one-half of the sample received the German version (Bertrams, Unger, & Dickhäuser, 2011) of the State Self-Control Capacity Scale (SSCCS; Twenge, Muraven, & Tice, 2004), which directly assesses subjective self-control limitations, and the other half a parallel form of the affect questionnaire. As the fourth factor was not of interest for the present study, all reported analyses here are controlled for the SSCCS manipulation.

Method

Participants

To detect an ego-depletion effect size equivalent to that reported by Carter and McCullough (2014) (i.e. d = 0.25, f = 0.125), our goal was a 95% power with α = .05, a conservative correlation among the repeated measures of .60, and an expected attrition rate of no more than 10%.[1] To achieve this, a total of 191 university students of psychology were recruited through flyers, mailing lists, social networks, bulletins, and direct approaches. As compensation for participating, participants received partial course credit and had the chance to win one of 13 €5 vouchers (approximately $US6) for a local ice cream parlour if they completed both lab sessions. Four participants were excluded because of technical problems, leading to a final sample size of 187 students (168 women, age M = 23.0 years, SD = 5.3).

[1]Power analysis for mixed models is not straightforward and currently relies on simulation studies. However, a conservative approach is to use each observation and run a senstivity power analysis with g*Power (Faul, Erdfelder, Lang, & Buchner, 2007) by indicating 'Exact' ('linear multiple regression: random model') as statistical test. For example, the sensitivity analysis for the multi-variable model of state moderators on $RT_{incongruent}$ in Study 1 showed that a very small effect size of r = .03 could be detected with an α level of .05, a power of .95, 35 813 observations, and 22 predictors. However, some researchers argue that the values that g*Power presents are too optimistic and are likely overestimating power (e.g., Giner-Sorolla, 2018). To accompany for this, the sensitivity with only a quarter of observations yielded a detectable effect size of r = .06.

DOI: 10.1002/per

Design
The complete design of this study was a randomized, 2 (depletion manipulation) × 2 (period) × 2 (order) × 2 crossover study, with two experimental factors of interest: The within-person factor of ego depletion (depletion versus control condition) and the within-person factor of period (first versus second lab session). The between-person factor order (first versus second depletion condition) was implemented to control for effects due to the order of the experimental conditions. Using block randomization, participants were randomly and nearly evenly assigned to the experiment conditions, with $n_{\text{depletion condition first}} = 47$ and $n_{\text{depletion condition second}} = 46$ for the SSCCS condition and $n_{\text{depletion condition first}} = 48$ and $n_{\text{depletion condition second}} = 46$ for the no-SSCCS condition.

Procedure
Participants signed up to take part in a study consisting of two lab sessions that were 2 weeks apart. After signing informed consent forms, participants indicated their baseline affect and then completed a computer version (Sripada, Kessler, & Jonides, 2014) of the letter-e task (Baumeister et al., 1998) as the first task in the first lab session (task described in the section). This task was used as the ego-depletion manipulation whereby one version either required self-control (complex rules condition) or not (simple rules condition). Subsequently, participants completed questionnaires on the depletion manipulation, affect, and task motivation, which were followed by the second task, a multi-source interference task (MSIT) (Bush, Shin, Holmes, Rosen, & Vogt, 2003; task is described in the section). The first lab session ended with participants rating the task difficulty, task motivation, and affect. This procedure was repeated in the second lab session, but participants who had previously been in the control condition were assigned to the depletion condition and vice versa. In between the two lab sessions, participants completed trait questionnaires online.

Materials
First task: computerized letter-e task. We used a computer version (Sripada et al., 2014) of the letter-e task (Baumeister et al., 1998), which is as effective in invoking performance decrements as the pen-and-paper version (Arber, Ireland, Feger, Marrington, Tehan, & Tehan, 2017). The letter-e or crossing-out letters task was chosen because it showed the highest averaged effect size as the depleting task in the meta-analysis by Hagger et al. (2010). Participants in the control condition were instructed to press the letter ‘g’ on the keyboard as quickly and accurately as possible whenever the letter ‘e’ appeared in a displayed single word. Participants in the depletion condition were shown the same words but with the additional instruction to refrain from pressing the letter ‘g’ when the letter ‘e’ was next to or one letter away from a vowel. Following this additional rule requires self-control, because participants need to assess when and when not to react to the letter ‘e’, controlling their impulse to solely react to the letter ‘e’. After 20 practice trials, during which participants were instructed to react to each letter ‘e’ to establish a behavioural pattern in the depletion condition, participants completed 150 trials, which lasted approximately 7.5 minutes in total.

As a manipulation check of the ego-depletion manipulation, we used three items to assess, respectively, the three categories established by Hagger et al. (2010): difficulty [three items: ‘difficult’, ‘hard’, and ‘easy’ (reversed)], effort [three items: ‘effortful’, ‘laborious’, and ‘undemanding’ (reversed)], and fatigue [three items: ‘depleting, ‘exhausting’, and ‘refreshing’ (reversed)]. Each item was scored on a Likert scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). As shown before (Wenzel et al., in press), an exploratory factor analysis demonstrated only two factors with eigenvalues larger than 1: a difficulty factor (‘difficult’, ‘hard’, ‘easy’, and ‘undemanding’) and a fatigue factor (‘depleting’ and ‘exhausting’).

Regarding reliability of the manipulation check, we followed suggestions by Shrout and Lane (2012) to not report Cronbach’s α, given that the random measurement error is affected by both variations between and within individuals in repeated measures. Instead, we separated the reliabilities between and within individuals. Scores for both the between-person reliabilities (R_{KRN}) for task difficulty and fatigue and the within-person reliabilities (R_{CN}) were acceptable to good (Table 1).
Second task: multi-source interference task. We used the MSIT (Bush, Shin, Holmes, Rosen, & Vogt, 2003; Shamosh & Gray, 2007; Wenzel, Kubiak, & Conner, 2014) as the second task, given that the Stroop task showed the largest averaged effect size as the dependent task in the meta-analysis by Hagger et al. (2010). The MSIT is a similar interference task and was specifically designed by Bush et al. (2003) to produce larger interference effects than the Stroop task. Participants were instructed to indicate as quickly and as accurately as possible the one character out of three characters that differed from the other identical ones, regardless of its position. Control trials were ‘1xx’, ‘x2x’, and ‘xx3’, such that the target was always a number and the distractors were always the letter ‘x’. The target was also larger than the distractors and placed such that the target character ‘2’ appeared in the second place in control trial ‘x2x’. In interference trials (112, 131, 313, 221, 232, 233, 311, 322, 331, and 332), the distractors were also numbers, and the target was smaller in size than the distractors. Additionally, the target did not always match its position. In total, participants completed 150 trials (50 congruent and 100 incongruent trials) in each second task, with a fixed, interspersed order to hold trial-to-trial-adaptations constant across participants (Lorist & Jolij, 2012).

The MSIT characters were printed in white (font: Arial; size: 18 and 24 pt; and style: bold) on a black background and were displayed in the centre of a computer screen positioned 90 cm in front of participants’ eyes. Participants had 2 seconds to complete a trial before hitting the response deadline, and a white plus sign appeared for 500 milliseconds before each trial. The first trial (0.67% of all trials), trials completed in less than 200 milliseconds (0.0%), and any trials with a reaction time (RT) above or below 3 *SD*s from each individual’s mean RT (1.47%) were excluded from all subsequent analyses. Applying these rules excluded 2.12%

DOI: 10.1002/per

Table 1. Means, standard deviations, range, reliability, and zero-order correlations of trait moderators in Study 1

Dependent variables	*M*	*SD*	Min	Max	*α*	1	2	3	4	5	6	7	8	9	10	11
1. Age (years)	23.0	5.3	18	52	—	—										
2. Gender (0 = male; 1 = female)	0.89	0.31	0	1	—	−.07	—									
3. SC	55.1	15.3	19.2	94.2	.83	.05	.07	—								
4. Impulsivity	35.7	12.4	8.9	71.1	.77	−.01	.03	−.60***	—							
5. Behavioural inhibition system	31.3	19.1	0.0	95.2	.82	.08	−.16*	.11	.04	—						
6. Behavioural activation system	29.3	14.3	2.6	76.9	.82	.00	−.13	−.01	−.10	.36***	—					
7. Implicit theories about SC	51.2	12.0	18.3	93.3	.77	.01	−.03	−.33***	.08	−.32***	.04	—				
8. Openness	66.4	13.7	27.1	95.8	.75	.09	−.09	−.15*	.09	.09	−.08	−.12	—			
9. Conscientiousness	68.1	14.6	22.9	97.9	.86	−.07	.13	.67***	−.57***	.00	−.17*	−.15*	−.24**	—		
10. Extraversion	62.2	13.4	18.8	91.7	.80	−.07	.01	.16*	.03	.14	−.31***	−.23**	.00	.23**	—	
11. Agreeableness	62.2	12.0	29.2	95.8	.77	−.12	.26*	.07	.03	−.11	−.11	.01	.09	.00	.23**	—
12. Neuroticism	73.0	15.3	2.1	89.6	.83	−.09	.07	−.36***	.14	−.51***	−.01	.35***	.05	−.18*	−.38***	−.17*

Note: SC, self-control.
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

of the data. The first trial was excluded because participants usually respond slower to the first trial in a task, which was the case for the second task as well, $z = -12.81$, $p < .001$, $f = 0.48$.

State moderators. The mean and standard error for each condition of the ego-depletion manipulation, as well as the reliabilities of the state measures, are shown in Table 1.

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 (Steyer, Schwenkmezger, Notz, & Eid, 1997). This instrument assesses momentary affect on three dimensions: valence (items: ‘good’ and ‘bad’), energetic arousal (items: ‘awake’ and ‘tired’), and calmness (items: ‘nervous’ and ‘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, more energy, or more calmness, respectively.

Motivation. Because motivation is currently a popular explanation for reduced performance in the second task of a dual-task set-up (e.g. Inzlicht & Schmeichel, 2012), participants were asked to complete the short form of the Questionnaire on Current Motivation (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 achievement situations: probability of success, anxiety, challenge, and interest. Anxiety, challenge, and interest showed very good between-person reliabilities, whereas probability of success was only acceptable. The within-person reliabilities were also only acceptable, with the only exception being challenge ($R_{CN} = .38$).[2]

Trait moderators. The mean, standard deviation, range, internal consistencies, and intercorrelations of the trait measures are shown in Table 1.

Self-control. To assess trait levels of self-control, participants completed the German short version (Bertrams & Dickhäuser, 2009) of the Brief Self-Control Scale (BSCS; Tangney, Baumeister, & Boone, 2004). The BSCS consists of 13 items (1 = *not at all* to 5 = *very much*) and showed good internal consistency in our data (Cronbach’s $\alpha = .83$). Higher values of the averaged scale represent higher levels of self-reported self-control.

Impulsivity. Participants completed 15 items assessing impulsivity using the German version of the short version (Meule, Vögele, & Kübler, 2011) of the Barratt Impulsiveness Scale (BIS-15; Spinella, 2007). The BIS-15 has three subscales: attentional, motivational, and non-planning. For this study, we used the averaged BIS-15 across all 15 items, which demonstrated acceptable internal consistency, Cronbach’s $\alpha = .77$. Items were answered on a 4-point scale (1 = *rarely or never* to 4 = *always or almost always*), with higher averaged scores reflecting more impulsivity.

[2]Within-person reliability (R_{CN}) is more sensitive to the number of items, which may explain the lower values.

DOI: 10.1002/per

Regulation motivation. The German version (Strobel, Beauducel, Debener, & Brocke, 2001) of the behavioural inhibition system (BIS)/behavioural approach system (BAS) questionnaire (Carver & White, 1994) was used to assess the BIS and the BAS, using an aggregated score for the BAS subscales for fun seeking, drive, and reward responsiveness. Twenty-four items were rated on a 4-point scale (1 = *very true for me* to 4 = *very false for me*), with higher averaged scores representing stronger levels of the BIS and BAS.

Implicit theories about self-control limitations. To assess implicit theories about self-control limitations, we translated 12 items into German that were used in prior research (Job et al., 2010). An example item for a limited theory is 'After a strenuous mental activity your energy is depleted and you must rest to get it refueled again', while an example item for a non-limited theory is 'Your mental stamina fuels itself; even after strenuous mental exertion you can continue doing more of it'. All items were answered on a 6-point scale (1 = *strongly agree* to 6 = *strongly disagree*). Items indicating limited theory were reverse coded, such that higher averaged scores reflected greater agreement with a limited theory of self-control.

Big Five personality traits. The Big Five personality traits (i.e. openness, conscientiousness, extraversion, agreeableness, and neuroticism) were assessed using the 60-item German version (Borkenau & Ostendorf, 2008) of the NEO Five-Factor Inventory (McCrae & Costa, 2004). All five personality traits were measured using 12 Likert-type items per scale, ranging from 1 (*strongly agree*) to 5 (*strongly disagree*), and showed good internal consistencies.

Statistical analyses

Data preparation

The following five measures can be derived from the MSIT as a measure of self-control performance: (1) mean RT on interference trials ($RT_{incongruent}$); (2) difference between mean RT on interference and control trials ($M_{RT\_incongruent} - M_{RT\_congruent}$; MSIT effect); (3) ratio score ($M_{RT\_incongruent} - M_{RT\_congruent}$) ($M_{RT\_incongruent} + M_{RT\_congruent}$); (4) RT variability on interference trials (i.e. the sum of the sigma and tau variability parameters using ex-Gaussian modelling; Sripada et al., 2014); and (5) the number of errors. The ratio score was highly correlated with the difference score ($r > .90$), and the RT variability score aggregates the data, and thus, both outcomes were discarded from the multilevel analyses.

We aimed at computing hierarchical generalized random-effects mixed models to model the non-normally distributed reactions times by choosing a gamma distribution as the probability distribution of the dependent variable given its good fit with RT data (Baayen & Milin, 2010) and to avoid problems associated with transformation of variables (Lo & Andrews, 2015). However, the models with crossed random effects did not converge, and we, thus, transformed the data to an inverse Gaussian distribution, which showed the best fit with RT data (Lo & Andrews, 2015). To that end, we first computed the inverse of the RT multiplied by the response deadline (2000 milliseconds) and then divided the inversed data by the maximum of the inversed RTs (.0040323). We finally 'swapped' the inversed distribution by subtracting the value from the sum of the minimum and maximum value, such that higher values indicate slower RTs. This procedure provides a better fit of the inverse Gaussian distribution, leading to a good normal fit, with −0.27 for skewness and 2.74 for kurtosis. The transformed RT variable can also be interpreted in the same way as changes of the RT in milliseconds, easing interpretation of the results.

Given the differing scales of the state and trait moderators, we used the percent of maximum possible scores (POMP; Cohen, Cohen, Aiken, & West, 1999) to make the coefficients more comparable. The POMP can be calculated by subtracting the minimum of a respective outcome from the outcome, dividing this difference by the difference between the maximum and the minimum of a respective outcome, and finally multiplying it by 100 so that the POMP score ranges between 0 and 100. This procedure maintains the distribution of the variables while making the results easier to interpret and compare with each other.

Analysis of the manipulation check and moderator variables

To investigate manipulation check and state moderator differences due to the depletion manipulation, we first subjected the outcomes to 2 (depletion versus control) × 2 (period: first versus second lab session) × 2 (SSCCS versus no SSCCS assessment) mixed models, using the mixed command with the restricted maximum likelihood option in Stata 15 (Stata Corporation, College Station, TX). Underspecifying the random structure by not including the depletion manipulation as random slopes and the targets as random intercepts typically increases false positive rates (Schielzeth & Forstmeier, 2009). The reason is that including nested or crossed random slopes can lead to larger CIs (and non-significance) because the standard errors of the depletion manipulation interactions increase. The model without random slopes estimates one common slope for each moderator and for each condition of the depletion manipulation. However, when including random slopes, each individual can now have their own slope; the fixed effect is then the average of those slopes (for each moderator and for each condition). Increased standard errors then make the group average less certain, leading to wider CIs that more likely include zero. Thus, the mixed model (Equation (1)) allowed for random participant (*P*) intercept variance, which was crossed with the depletion manipulation (*C*), which allows that the mean difference in, for example, motivation, varies across participants, for the *i*th participants in the *k*th depletion condition:

$$Y_{ik} = \beta_0 + \mu_i^P + \left(\beta_1 + \mu_i^{PxC}\right)C_{ik} + \beta_2 SSCCS_{ik} + \beta_3 Period_{ik} + \beta_4 C_{ik} SSCCS_{ik} + \beta_5 C_{ik} Period_{ik} + \beta_6 SSCCS_{ik} Period_{ik} + \beta_7 C_{ik} SSCCS_{ik} Period_{ik} + \varepsilon_{ik} \quad (1)$$

We, then, used the contrast command to compute an analysis of variance-style test of the main effect of the depletion manipulation. As a measure of effect size for the ego-depletion measure, we found the *t*-value that corresponded

DOI: 10.1002/per

to the p-value of the z-value Stata 15 reports for the depletion manipulation and then divided this t-value by the square root of our sample size (n = 187), which resulted in the within-person effect size d_z (Lakens, 2013). However, this effect size is not directly comparable with effect sizes derived from between-person design. There are various ways to solve this problem, but we found that the best way to compare results from different experimental setups would be an effect size that is not affected by those differences. This is Cohen's d or d_{av} (Lakens, 2013), which divides the mean difference by the averaged standard deviation from the depletion and control condition. Given that prior research on ego depletion relied almost solely on between-person designs, with only a few exceptions (Francis, Milyavskaya, Lin, & Inzlicht, 2018; Wenzel et al., in press), we focused on d_{av}, which can be better compared with the effect sizes reported of the multi-lab replication projects or meta-analyses (e.g. Hagger et al., 2010).

Analysis of the ego-depletion effect

To investigate consecutive self-control performances due to our depletion manipulation, we employed multilevel analyses on the three MSIT outcomes $RT_{incongruent}$, MSIT effect, and errors. In comparison with analysing the aggregated data, this design allowed us to better model the underlying RT distribution as well as to separate out variance due to random factors. As, in the current study, participants were crossed (not nested) with the depletion manipulation and with the targets, whereas the targets were nested in the depletion manipulation, we modelled this by computing mixed models with crossed random effects (see Judd, Westfall, and Kenny, 2017, for more information about the differences between nested and crossed random factors). To be transparent and to illustrate the importance of including these variance components as crossed random factors and random slopes, we also computed models without crossed random effects and indicate differences from the main analyses.

For example, for the model predicting the $RT_{incongruent}$, we included the fixed effects of the depletion manipulation and controlled for period, SSCCS assessment, trial number, error, and congruency, and error in and RT of the prior trial. The random part consisted of the random intercepts of the participants (P), the random intercepts of the targets (T), the random slope of the depletion manipulation across participants (PxC), the random intercept–slope covariance between the participants and the depletion manipulation (P,PxC), as well as the residual variance (ε). Equation (2) shows the model for the ith participant to the jth target in the kth depletion condition.

$$\begin{aligned} Y_{ik} = {} & \beta_0 + \mu_i^P + \mu_j^T + \left(\beta_1 + \mu_i^{PxC}\right)C_{ijk} + \beta_2 SSCCS_{ijk} \\ & + \beta_3 Period_{ijk} + \beta_4 TrialType_{ijk} + \beta_5 Error_{ijk} \\ & + \beta_6 priorError_{ijk} + \beta_7 priorRT_{ijk} + \beta_8 Trialnumber_{ijk} \\ & + \beta_9 C_{ijk} Period_{ijk} + \mu_i^{P,PxC} + \varepsilon_{ijk} \end{aligned} \tag{2}$$

We included a relatively large number of level 1 variables such as the prior RT to reduce the large amount of residual variance that can be observed when analysing RT on the trial level. For example, Gratton, Coles, and Donchin (1992) reported that interference in the Stroop task is reduced following incongruent trials compared with congruent trials, evidencing trial-to-trial adaptation processes. The models for the MSIT effect additionally included the two-way interaction between the depletion manipulation and the trial type (interference versus control). A Bernoulli-based mixed logistic regression was used to predict the binary error variable (0 = no error; 1 = error).

Moderation analysis of the ego-depletion effect

To investigate the moderators of ego depletion, we computed univariable and multi-variable models for each outcome in order to show the direct association of a moderator as well as the association controlled for the influence of the other moderators. We separated the models addressing state and trait moderators for the multi-variable analyses because the associations between trait moderators might be mediated by the state moderators. In the following section, the term *moderators* refers to both state and trait moderators.

$$\begin{aligned} Y_{ik} = {} & \beta_0 + \mu_i^P + \mu_j^T + \left(\beta_1 + \mu_i^{PxC}\right)C_{ijk} + \beta_2 SSCCS_{ijk} \\ & + \beta_3 Period_{ijk} + +\beta_4 TrialType_{ijk} + \beta_5 Error_{ijk} \\ & + \beta_6 priorError_{ijk} + \beta_7 priorRT_{ijk} + \beta_8 Trialnumber_{ijk} \\ & + \beta_9 C_{ijk} Period_{ijk} + \beta_{10} Moderator_{ijk} \\ & + \beta_{11} C_{ijk} Moderator_{ijk} + \mu_i^{P,PxC} + \varepsilon_{ijk} \end{aligned} \tag{3}$$

The models for the MSIT effect additionally included the three-way interaction between the depletion manipulation, the moderator, and the trial type (interference versus control) and all of the two-way interactions and main effects.

Moderators in all models were person mean centred for state moderators and grand mean centred for trait moderators. Given the vast number of moderators, we applied the false discovery rate to adjust the 95% intervals for multiple testing (Benjamini & Yekutieli, 2005). In this procedure, the α level used to construct the CI is divided by the number of tests and then multiplied by the rank of the p-value of the respective tests. For example, there were 19 relevant coefficients for the moderation of consecutive self-control: seven state moderators and 12 trait moderators. Thus, the moderator with the second lowest p-value is compared with an α level of .05/19 * 2 = 0.005.

RESULTS

Analysis of the manipulation check and moderator variables

As indicated in Table 2, the 2 (depletion versus control) × 2 (period: first versus second lab session) × 2 (SSCCS versus no SSCCS assessment) mixed models show that participants in the depletion condition reported both higher task difficulty and fatigue of the first task, with a relatively small 95% CI that did not include zero and with large effect sizes. Both groups showed only small differences in the mood and motivation scales, except for the probability of success scale, such

DOI: 10.1002/per

Table 2. Reliabilities, group means, standard errors, and inferential statistics of state moderators in Study 1

Dependent variables	R_{KRN}	R_{CN}	DC M	DC SE_M	CC M	CC SE_M	M_{DC-CC}	SE_{DC-CC}	95% CI	d_z	d_{av}	$\beta^{\dagger}$
1. Difficulty	.44	.81	57.2	1.54	22.0	1.13	35.2	1.60	**[32.1, 38.3]**	1.66	1.90	.28
2. Fatigue	.54	.73	47.0	1.77	29.6	1.66	17.4	1.92	**[13.6, 21.2]**	0.66	0.74	.35
2. MDMQ valence	.36	.84	62.6	1.23	66.4	1.23	−3.8	1.47	**[−6.6, −0.9]**	−0.19	−0.22	.30
3. MDMQ energetic arousal	.54	.77	50.2	1.57	53.1	1.57	−2.9	1.74	[−6.3, 0.6]	−0.12	−0.13	.39
4. MDMQ calmness	.57	.79	69.0	1.43	70.7	1.43	−1.7	1.54	[−4.7, 1.3]	−0.08	−0.09	.43
5. QCM probability of success	.28	.68	65.6	1.31	84.0	0.96	−18.5	1.42	**[−21.2, −15.7]**	−0.99	−1.17	.22
6. QCM anxiety	.81	.55	35.1	1.77	28.7	1.72	6.3	1.37	**[3.7, 9.0]**	0.34	0.26	.68
7. QCM challenge	.69	.47	56.9	1.38	54.1	1.38	2.9	1.16	**[0.6, 5.1]**	0.18	0.15	.64
8. QCM interest	.82	.52	34.5	1.66	35.5	1.57	−1.0	1.21	[−3.4, 1.4]	−0.06	−0.04	.72

Note: The 95% CI in bold do not include zero. *DC*, depletion condition; *CC*, control condition; MDMQ, Multidimensional Mood Questionnaire; QCM, Questionnaire on Current Motivation; R_{KRN}, between-person reliability; R_{CN}, within-person reliability (change from session to session); 95% CI, 95% confidence interval; d_z, within-person effect size; d_{av}, between-person comparable effect size.
†Test–retest reliability (controlled for order and State Self-Control Capacity Scale manipulation).

that participants reported lower anticipated success in the depletion than in the control condition. Taken together, these results demonstrate that our manipulation worked as intended by employing a depletion condition that felt more difficult and fatiguing but not did not impact mood and motivation. The full models can be found in the online supplementary materials (https://osf.io/43fmu/?view_only=faf2593c9c4c4e9b8c71c9b2718bd1d0).

Analysis of the ego-depletion effect

An in-depth analysis of the ego-depletion effect in our cross-over design can be found in the other publication (Wenzel et al., in press).[3] For the three MSIT outcomes $RT_{incongruent}$, MSIT effect, and errors, the effect sizes were very small, ranging from $d_{av} = -0.03$ to $d_{av} = 0.13$. However, the random slope of the depletion manipulation was considerable: for example, for $RT_{incongruent}$, the fixed effect of the depletion manipulation was $b = 6.44$ milliseconds, 95% CI [−0.20, 13.09], and the random slope was $var = 1827.6$, 95% CI [1438.6, 2321.9]. Although the mean difference was, thus, very small, there was large between-person variance, such that approximately 95% of the participants were within −77.35 and 90.23 milliseconds of the typical value for ego depletion ($6.44 \pm 1.96 \times \sqrt{1827.6}$). To further illustrate this variability, we computed how many participants showed an improved performance by estimating best linear unbiased prediction (BLUP; Robinson, 1991). This demonstrated that 81 participants (43.3%) showed a BLUP lower than −6.44 milliseconds and, thus, an improved performance due to initial self-control. The BLUP figures can be found in the online supplementary materials (https://osf.io/43fmu/?view_only=faf2593c9c4c4e9b8c71c9b2718bd1d0).

The results for the MSIT effect and the errors showed similar large between-person variance, with approximately 95% of the participants who were within −78.18 and 60.12 milliseconds of the fixed depletion effect for the MSIT effect of $b = -9.03$ milliseconds and within −0.76 and 1.17 of the fixed depletion effect for the errors of $b = 0.21$.

Finally, we checked how reliable the three MSIT outcomes were across the two lab situations by aggregating the data and computing multiple regressions with the performance of the second lab session as the outcome and of the first lab session as the predictor, controlled for the SSCCS manipulation and the order of depletion and control condition. The results indicated that only the mean $RT_{incongruent}$ showed a good test–retest reliability ($\beta = 0.86$), whereas the reliability for the MSIT effect ($\beta = 0.57$) and the number of errors ($\beta = 0.53$) was only acceptable.

To sum up, these results demonstrate, on the one hand, very small and inconsistent ego-depletion effects. On the other hand, they also demonstrate large between-person variability, evidencing moderating influences that were not covered by the control variables such as period or prior RT, which were included in the model.

Moderation analysis of the ego-depletion effect

To investigate statistical moderation, we computed the univariable models outlined in formula (2) as well as multi-variable models including either all state or all trait moderators. Tables 3 and 4 show that the models did not reveal consistent moderation of ego depletion across the outcomes and moderators. Regarding state moderators, the impact of the depletion manipulation on the $RT_{incongruent}$ and on errors did not show significant moderation. The exceptions were the three affect scales and the motivation subscales challenge and interest, where the 95% CIs of the two-way interactions with the depletion manipulation did not include zero, such that participants who reported feeling more well, awake, or calm or being less challenged and less interested in the first task demonstrated a lower MSIT effect, even after adjusting the 95% CI for multiple comparisons. However, in the multi-variable analysis, only the moderation of energetic arousal, calmness, and interest remained significant, as long as the comparisons in the multi-variable analysis were not adjusted for multiple comparisons. With regard to trait moderators, only a few trait moderators were statistically

[3]The results of the between-person analysis of the ego-depletion effect with the data from the first laboratory session only can be found on OSF (https://osf.io/43fmu/?view_only=faf2593c9c4c4e9b8c71c9b2718bd1d0).

DOI: 10.1002/per

Table 3. Fixed- and random-effects estimates of reaction time on incongruent trials in the MSIT, MSIT effect, and errors as a function of consecutive self-control and state moderators in Study 1

	$RT_{incongruent}$		MSIT effect ($RT_{incongruent} - RT_{congruent}$)		Errors	
Fixed effects	*b* (*SE*)	95% CI	*b* (*SE*)	95% CI	*b* (*SE*)	95% CI
Univariable analyses						
MDMQ—valence	0.78 (1.24)	[−1.64, 3.29]	**−0.88†,‡ (0.25)**	[−1.37, −0.40]	0.020 (0.011)	[−0.002, 0.042]
MDMQ—energetic arousal	0.70 (1.04)	[−1.33, 2.74]	**−0.87†,‡ (0.21)**	[−1.27, −0.46]	0.018 (0.010)	[−0.001, 0.037]
MDMQ—calmness	0.48 (1.17)	[−1.81, 2.77]	**−0.87†,‡ (0.23)**	[−1.32, −0.42]	0.019 (0.011)	[−0.002, 0.040]
QCM—probability of success	1.89 (1.28)	[−0.61, 4.39]	0.39 (0.25)	[−0.10, 0.90]	0.006 (0.012)	[−0.018, 0.029]
QCM—anxiety	−1.53 (1.33)	[−4.13, 1.07]	0.17 (0.26)	[−0.35, 0.69]	0.011 (0.012)	[−0.013, 0.036]
QCM—challenge	−2.09 (1.56)	[−5.14, 0.96]	**0.93†,‡ (0.31)**	[0.32, 1.54]	0.008 (0.014)	[−0.020, 0.036]
QCM—interest	−0.53 (1.50)	[−3.47, 2.42]	**0.98†,‡ (0.30)**	[0.39, 1.56]	−0.004 (0.014)	[−0.031, 0.023]
Multi-variable analyses						
MDMQ—valence	0.71 (1.87)	[−2.95, 4.37]	0.07 (0.38)	[−0.66, 0.81]	0.005 (0.017)	[−0.029, 0.039]
MDMQ—energetic arousal	0.09 (1.34)	[−2.55, 2.72]	**−0.63† (0.27)**	[−1.16, −0.10]	0.012 (0.013)	[−0.013, 0.037]
MDMQ—calmness	0.14 (1.49)	[−2.78, 3.07]	**−0.68† (0.30)**	[−1.27, −0.09]	0.011 (0.014)	[−0.016, 0.039]
QCM—probability of success	1.80 (1.32)	[−0.79 4.40]	0.37 (0.27)	[−0.15, 0.89]	0.009 (0.012)	[−0.015, 0.033]
QCM—anxiety	−0.90 (1.35)	[−3.55, 1.76]	0.12 (0.27)	[−0.41, 0.66]	0.014 (0.013)	[−0.011, 0.039]
QCM—challenge	−2.24 (1.82)	[−5.79, 1.32]	0.54 (0.37)	[−0.17, 1.26]	0.013 (0.017)	[−0.020, 0.046]
QCM—interest	0.36 (1.73)	[−3.04, 3.76]	**0.68 (0.35)**	[0.00, 1.36]	−0.013 (0.016)	[−0.044, 0.019]
Random effects (multi-variable)	*Var* (*SE*)	95% CI	*Var* (*SE*)	95% CI	*Var* (*SE*)	95% CI
Depletion manipulation	1796 (220)	[1412, 2283]	2449 (280)	[1957, 3064]	0.218 (0.063)	[0.124, 0.383]
Participants	7098 (755)	[5763, 8744]	7961 (840)	[6473, 9790]	0.421 (0.070)	[0.305, 0.583]
Depletion manipulation × Participants	−703 (295)	[−1280, −125]	−971 (352)	[−1661, −280]	−0.002 (0.051)	[−0.101, 0.097]
Target	2042 (583)	[1167, 3572]	1839 (497)	[1083, 3125]	—	—
Residual	15 348 (115)	[15 123, 15 575]	17 934 (110)	[17 720, 18 151]	$\pi^2/3$	—

Note: Estimates in bold do not include zero in their 95% CI. MSIT, multi-source interference task; $RT_{incongruent}$ = reaction time on incongruent trials; MSIT effect, the difference between reaction time on interference and control trials; Errors = error in a trial of the MSIT; 95% CI, 95% confidence interval; MDMQ, Multidimensional Mood Questionnaire; QCM, Questionnaire on Current Motivation.
†Estimates do not include zero in their 95% CI when crossed effects are removed from the model.
‡Estimates still do not include zero in their 95% CI after adjusting for multiple comparisons via the false discovery rate.

significant, although none of these trait moderators were consistent across the univariable and multi-variable analyses or across the three MSIT outcomes.

DISCUSSION

Although we found large between-person variability of our depletion manipulation on all three MSIT outcomes, the estimates of all state and trait moderators were rather uncertain, demonstrated by wide CIs. When adjusting for multiple comparisons, all CIs included zero, and, thus, we refrain from interpreting the few moderators that were significant without correction due to an increased likelihood of being false positives. We conclude that neither the specific self-control related traits nor any of the Big Five personality traits can sufficiently explain why some individuals show impaired or improved self-control performance after initial self-control efforts. However, although no single moderator was significant and residual variance was not reduced, some variance in the depletion manipulation could be attributed to the moderators, as indicated by a proportional reduction in the depletion manipulation slope variance (Peugh, 2010) of 1.7%, 1.6%, and 12.8% for state moderators of $RT_{incongruent}$, MSIT effect, and errors as well as 6.8%, 6.4%, and 28.5% for trait moderators. This demonstrates that some of the variance in the random slopes of the depletion manipulation across the participants could be explained by the moderators but not by a single impactful moderator.

One limitation of Study 1 was that, as is commonly performed in the field of ego-depletion research, it employed artificial tasks that do not connect well to the self-control demands present in daily life. In Study 2, we wanted to extend our research on individual differences in consecutive self-control to the daily life of individuals.

STUDY 2

We affirm that we reported all manipulations and exclusions in the present study. However, we did not report all measures, because this study was part of the Bonnie and Clyde study, which examined the effectiveness of a mindfulness intervention in influencing mindfulness, self-control, and affective processes (Rowland, Wenzel, & Kubiak, 2016). The data and statistical analyses of Study 2 are available at https://osf.io/43fmu/?view_only=faf2593c9c4c4e9b8c71c9b2718bd1d0. Please note that this dataset has been used for other publications or manuscripts, which investigated the following research questions: the influence of mindfulness on affective dynamics (Rowland, Wenzel, & Kubiak, in press), the efficacy of an ultra-brief mindfulness training, the influence of trait self-control on the emotion regulation network, and the importance of variability in emotion regulation strategies.

DOI: 10.1002/per

Table 4. Fixed- and random-effects estimates of reaction time on incongruent trials in the MSIT, MSIT effect, and errors as a function of consecutive SC and trait moderators in Study 1

	$RT_{incongruent}$		MSIT effect ($RT_{incongruent} - RT_{congruent}$)		Errors	
Fixed effects	*b* (*SE*)	95% CI	*b* (*SE*)	95% CI	*b* (*SE*)	95% CI
Univariable analyses						
Age	−0.35 (0.64)	[−1.61, 0.90]	**0.95 (0.46)**	[0.04, 1.86]	0.006 (0.012)	[−0.017, 0.028]
Gender (0 = male; 1 = female)	−2.97 (10.97)	[−24.46, 18.52]	5.55 (7.93)	[−9.99, 21.08]	−0.248[†] (0.188)	[−0.617, 0.121]
SC	0.31[†] (0.22)	[−0.12, 0.74]	−0.31 (0.16)	[−0.63, <.01]	**0.008[†] (0.004)**	[0.001, 0.016]
Impulsivity	−0.37[†] (0.27)	[−0.90, 0.16]	0.04 (0.20)	[−0.35, 0.43]	**−0.012[†] (0.004)**	[−0.021, −0.002]
Behavioural inhibition system	0.19[†] (0.18)	[−0.15, 0.54]	−0.08 (0.13)	[−0.33, 0.17]	0.005[†] (0.003)	[−0.001, 0.011]
Behavioural activation system	0.16 (0.24)	[−0.31, 0.62]	0.22 (0.17)	[−0.12, 0.55]	−0.001 (0.004)	[−0.009, 0.007]
Implicit theories about SC	−0.23 (0.28)	[−0.78, 0.33]	0.40 (0.20)	[−0.01, 0.80]	−0.009 (0.005)	[−0.018, 0.001]
Openness	−0.31[†] (0.25)	[−0.79, 0.17]	−0.07 (0.18)	[−0.42, 0.28]	0.001 (0.004)	[−0.007, 0.010]
Conscientiousness	0.25[†] (0.08)	[−0.21, 0.70]	0.05 (0.17)	[−0.27, 0.38]	0.007[†] (0.004)	[−0.001, 0.015]
Extraversion	0.08 (0.25)	[−0.41, 0.58]	0.22 (0.18)	[−0.14, 0.58]	0.000 (0.004)	[−.009, 0.009]
Agreeableness	0.53[†] (0.28)	[−0.02, 1.08]	−0.23 (0.20)	[−0.62, 0.18]	**−0.012[†] (0.005)**	[−0.021, −0.002]
Neuroticism	−0.19 (0.22)	[−0.62, 0.25]	0.17 (0.16)	[−0.14, 0.48]	−0.002 (0.004)	[−0.010, 0.005]
Multi-variable analyses						
Age	−0.14 (0.65)	[−1.41, 1.13]	**1.27[†] (0.48)**	[0.33, 2.21]	0.001 (0.011)	[−0.021, 0.024]
Gender (0 = male; 1 = female)	−10.37 (11.53)	[−32.98, 12.23]	10.67 (8.57)	[−6.11, 27.46]	−0.155 (0.192)	[−0.532, 0.222]
SC	0.06 (0.35)	[−0.63, 0.74]	**−0.67[†] (0.26)**	[−1.18, −0.16]	0.005 (0.006)	[−0.007, 0.016]
Impulsivity	−0.36 (0.38)	[−1.09, 0.38]	−0.09 (0.28)	[−0.64, 0.46]	−0.010[†] (0.006)	[−0.022, 0.002]
Behavioural inhibition system	0.34 (0.24)	[−0.14, 0.81]	−0.19 (0.18)	[−0.54, 0.16]	0.005[†] (0.004)	[−0.003, 0.013]
Behavioural activation system	−0.04 (0.29)	[−0.60, 0.52]	**0.54[†] (0.21)**	[0.12, 0.95]	−0.006[†] (0.005)	[−0.016, 0.004]
Implicit theories about SC	−0.22 (0.32)	[−0.86, 0.42]	0.21 (0.24)	[−0.26, 0.68]	−0.005 (0.006)	[−0.015, 0.006]
Openness	−0.44 (0.26)	[−0.95, 0.08]	0.03 (0.20)	[−0.35, 0.42]	0.001 (0.005)	[−0.008, 0.001]
Conscientiousness	−0.01 (0.36)	[−0.71, 0.69]	0.49 (0.26)	[−0.03, 1.01]	−0.001 (0.006)	[−0.013, 0.011]
Extraversion	−0.11 (0.30)	[−0.70, 0.48]	**0.57[†] (0.22)**	[0.14, 1.01]	−0.001 (0.005)	[−0.011, 0.009]
Agreeableness	**0.78 (0.31)**	[0.17, 1.39]	−0.28 (0.23)	[−0.73, 0.18]	−0.010 (0.005)	[−0.021, 0.000]
Neuroticism	0.24 (0.30)	[−0.35, 0.84]	0.03 (0.22)	[−0.40, 0.47]	0.003 (0.005)	[−0.006, 0.013]
Random effects (multi-variable)	*Var* (*SE*)	95% CI	*Var* (*SE*)	95% CI	*Var* (*SE*)	95% CI
Depletion manipulation	1704 (211)	[1338, 2171]	2328 (268)	[1859, 2916]	0.169 (0.60)	[0.085, 0.338]
Participants	6251 (668)	[5070, 7707]	6949 (736)	[5647, 7552]	0.442 (0.073)	[0.319, 0.611]
Depletion manipulation × Participants	−619 (270)	[−1149, −89]	−788 (320)	[−1416, −160]	−0.030 (0.051)	[−0.131, 0.70]
Target	2040 (582)	[1166, 3567]	1838 (497)	[1082, 3122]	—	—
Residual	15 348 (115)	[15 123, 15 575]	17 859 (109)	[17 647, 18 075]	$\pi^2/3$	—

Note: Estimates in bold do not include zero in their 95% CI. MSIT, multi-source interference task; $RT_{incongruent}$ = reaction time on incongruent trials; MSIT effect, the difference between reaction time on interference and control trials; Errors = error in a trial of the MSIT; 95% CI, 95% confidence interval; SC, self-control.

[†]Estimates do not include zero in their 95% CI when crossed effects are removed from the model.

[‡]Estimates still include zero in their 95% CI when adjusting for multiple comparisons via the false discovery rate.

DOI: 10.1002/per

Method

Participants and design

Power calculation was based on a small intervention effect of Cohen's $d = 0.33$, with a power of $1 - \beta = .95$, an α level of .05, and an expected total attrition rate of 10%. One hundred thirty-seven participants were recruited in two waves, of which 11 participants dropped out throughout the study. We excluded all participants who completed less than 33% of the signals, leaving a total of 125 participants (77.6% women; $M = 22.9$ years, $SD = 5.1$). For more information on the inclusion criteria, design, and compensation, please refer to the study protocol (Rowland et al., 2016).

Procedure

The SMASH study combines seven weekly laboratory sessions with a 6-week experience sampling. After signing an informed consent form in the first lab session, participants completed questionnaires regarding self-control and personality traits (for a complete overview of the measures, see Rowland et al., 2016). Starting the next day and for 40 subsequent days, participants received six signals per day between 10:00 and 20:00 hours via the movisensXS experience sampling application (movisens GmbH, Karlsruhe, Germany) that presented them with questionnaires regarding their momentary self-control. The signals were randomly distributed, with the condition that they were always 45 to 200 minutes apart. The average interval was 103.4 minutes ($SD = 34.3$). After 40 days with weekly laboratory sessions, the study concluded with a post-measurement laboratory session.

Participants returned to the lab to weekly sessions to talk about any issues regarding the study and to keep adherence high. Moreover, participants completed state and trait questionnaires on mindfulness. Additionally, participants in the mindfulness training condition performed a computer-based guided breathing meditation for five weekly sessions. Importantly, it was only an ultra-brief meditation training that significantly influenced state mindfulness and subjective depletion but only yielded expectable small effect sizes. Moreover, the mindfulness intervention did not impact the variables relevant to this study. However, to reduce the risk of presenting spurious associations due to this training, we included this intervention as a control variable in our analyses.

Materials

Descriptive statistics and reliability of all measures are shown in Table 5. All continuous variables were transformed to POMP scores. However, we did not multiply the scores by 100 because this would have resulted in coefficients in the mixed models that would have been too small to easily present. The transformed scores thus range between 0 and 1.

State measures

Self-control. To assess self-control, we used the measures from Hofmann et al. (2012). Using a Likert-type scale ranging from 0 (*no desire at all*) to 7 (*very strong desire*), participants indicated whether they were currently or had been experiencing a desire within the last 30 minutes (assessing desire strength). For scores greater than 0, participants also completed items on the strength of conflict ('How much has it conflicted with a personal goal?'), resistance ('To which extent have you tried to resist your desire?'), and enactment ('To which extent have you acted on your desire?') within the last 30 minutes, using a Likert-type scale ranging from 0 (*not at all*) to 4 (*very much*). In addition, participants completed three items of the SSCCS (Bertrams, Unger, & Dickhäuser, 2011), which is a Likert-type scale that ranges from 1 (*not true*) to 7 (*very true*) and which assesses the current capacity to regulate oneself in that particular moment (e.g. 'I feel like my willpower is gone'). The mean value was inverted so that higher values reflected higher subjective levels of depletion (i.e. less momentary self-control).

Table 5. Means, standard deviations, range, and reliabilities of all variables in Study 2

	Min	Max	*M*	*SD*	R_{KRN}/α	R_{CN}	ICC
State measures							
Enactment	0	1	0.39	0.40	.87	—	.14
Subjective depletion (SSCCS)	0	1	0.12	0.18	.89	.54	.41
Temptation	0	1	0.08	0.20	.89	—	.16
Resistance	0	1	0.39	0.39	.88	—	.16
Positive affect	0	1	0.53	0.22	.91	.64	.44
Negative affect	0	1	0.19	0.18	.91	.64	.51
Internal motivation	0	1	0.45	0.34	.97	—	.47
External motivation	0	1	0.19	0.19	.96	—	.40
Trait measures							
Self-control	0.25	1	0.64	0.15	.84	—	—
Openness	0.23	0.98	0.68	0.15	.76	—	—
Conscientiousness	0.19	1	0.64	0.17	.77	—	—
Extraversion	0.19	0.97	0.63	0.17	.78	—	—
Agreeableness	0.38	0.97	0.69	0.12	.73	—	—
Neuroticism	0.14	0.89	0.50	0.16	.77	—	—

Note: R_{KRN}, between-person reliability (for state measures); α, Cronbach's α (for trait measures); R_{CN}, within-person reliability (change from session to session); ICC, intra-class correlation; SSCCS, State Self-Control Capacity Scale. As this reflects the amount of variance explained by clustering, here it explains the proportion of variance due to between-person differences.

DOI: 10.1002/per

We adopted a functional–cognitive framework (De Houwer, 2011) to capture both the cognitive (subjective feelings of depletion, assessed by the SSCCS) and the functional aspects (behavioural consequences, assessed by enactment of desires) of ego depletion. Momentary self-control was operationalized as the association between momentary resistance and momentary enactment or subjective depletion (via the SSCCS), such that successful momentary self-control was characterized by less enactment of temptations and less subjective depletion. Consecutive self-control was assessed by the interaction between momentary resistance and prior resistance on either current enactment or subjective depletion. Thus, ego depletion would be observed if the negative association between momentary resistance and current enactment or subjective depletion was positively moderated by prior resistance.

Temptations. Based on Milyavskaya and Inzlicht (2017), temptation strength (i.e. conflicted desires) was assessed by computing the product of desire strength (0 to 7) and conflict strength (0 to 4).

Affect. We assessed momentary affect using eight items (Kuppens, Allen, & Sheeber, 2010) selected from the circumplex model of affect (Russell, 1980). According to this model, single affective states can be differentiated by the two universal dimensions of valence and arousal. For this study, we focused on the valence dimension, computing mean scores for positive affect (happy, excited, relaxed, and satisfied) and negative affect (anxious, angry, depressed, and sad). Items were rated on a visual analogue scale ranging from 0 to 100, with higher values reflecting higher levels of positive and negative affect.

Motivation. Internal and external motivation was assessed based on the two respective subscales of the Situational Motivation Scale (Guay, Vallerand, & Blanchard, 2000). On the basis of the Situational Motivation Scale, we assessed internal motivation ('I set this goal because I think it is interesting and fun') and external motivation ('I set this goal because it is something I have to do') in regard to the goal participants were currently pursuing or within the last 30 minutes. The two items were rated from 1 (*not at all*) to 7 (*very much*), with higher values indicating greater motivation.

Trait measures. *Self-control*To assess trait levels of self-control, participants completed the German version of the BSCS at baseline (see section in Study 1).

Big Five personality traits. Participants completed 42 items assessing the Big Five personality traits using the German version (Lang, Lüdtke, & Asendorpf, 2001) of the Big Five Inventory (John, Donahue, & Kentle, 1991). Items were rated on a Likert scale, ranging from 1 (*strongly disagree*) to 5 (*strongly agree*).

Statistical analyses

Again, to increase transparency, we computed univariable and multi-variable mixed models with random intercepts for participants and days and random slopes for resistance and temptation in Stata 15. Again, we indicate whether the results changed without the random slopes and the random intercept for days and when the results were corrected for multiple testing of all moderators pertaining to a specific research question by adjusting the 95% intervals. The number of tests used for the correction was based on the number of tests that were computed for each research question. For example, there were 12 relevant coefficients for the moderation of consecutive self-control: four state moderators and eight trait moderators for each of the cognitive (subjective depletion via the SSCCS) and functional levels (current enactment) of analysis. Given that the subjective feels of depletion and its consequences do not have to overlap (Friese et al., 2018; Wenzel et al., in press), we adjusted for both outcomes separately (as in Study 1). Thus, the fourth lowest p-value is compared with an α level of .05/12 * 4 = 0.016.

Analysis of the ego-depletion effect in daily life

To investigate the depletion effect in daily life, either current enactment of temptations or current subjective depletion via the SSCCS was predicted by the two-way interaction between current and prior resistance and its main effects, controlled for the experimental group (0 = controls; 1 = mindfulness training), study wave (0 = first assessment wave; 1 = second assessment wave), day (1–40), and signal number per day (1–6) as well as prior and current temptation strength. Participants were crossed with days, and both current resistance and temptation were allowed to be random, and thus, the random part consisted of the random intercepts of the participants (P), the random intercepts of the days (D), the random slope of current resistance (R) and of temptation (T), and the residual variance (ε). Equation (4) shows the model for the ith participant of the jth day.

$$\begin{aligned} Y_{ik} = {} & \beta_0 + \mu_i^P + \mu_j^D + \left(\beta_1 + \mu_i^R\right)R_{ij} + \beta_2 priorR_{ij} \\ & + \beta_3 R_{ij} priorR_{ij} + \left(\beta_4 + \mu_i^T\right)T_{ij} + \beta_5 priorT_{ij} \\ & + \beta_6 Group_{ij} + \beta_7 Wave_{ij} + \beta_8 Day_{ij} + \beta_9 Signal_{ij} + \varepsilon_{ij} \end{aligned} \quad (4)$$

Moderation analysis of the ego-depletion effect in daily life

As with Study 1, we separated the models for state and trait moderators and computed both univariable and multi-variable models for both enactment and subjective depletion. Moderation of ego depletion was assessed by adding the three-way interaction between current and prior resistance with the moderator, along with all two-way interactions and main effects.

$$\begin{aligned} Y_{ik} = {} & \beta_0 + \mu_i^P + \mu_j^D + \left(\beta_1 + \mu_i^R\right)R_{ij} + \beta_2 priorR_{ij} \\ & + \beta_3 R_{ij} priorR_{ij} + \left(\beta_4 + \mu_i^T\right)T_{ij} + \beta_5 priorT_{ij} \\ & + \beta_6 Group_{ij} + \beta_7 Wave_{ij} + \beta_8 Day_{ij} + \beta_9 Signal_{ij} \\ & + \beta_{10} Moderator_{ij} + \beta_{11} R_{ij} Moderator_{ij} \\ & + \beta_{12} priorR_{ij} Moderator_{ij} + \beta_{13} R_{ij} priorR_{ij} Moderator_{ij} \\ & + \varepsilon_{ij} \end{aligned} \quad (5)$$

Level 1 variables such as state moderators were person mean centred to capture within-person processes, while level 2 variables such as trait moderators were grand mean centred.

DOI: 10.1002/per

RESULTS

Individual differences in adherence to the study protocol

A total of 75.9% of all ambulatory assessment were completed by the participants, which indicates good adherence to the study protocol. Participants reported experiencing desires at 35.0% of the signals and 59.2% of the time they tried to resist these desires to varying degrees.

We also looked at individual differences of response rates in our sample. To that end, we computed the zero-order correlation of the aggregated moderators and the self-control relevant variables desire, conflict, resist, enact, and SSCCS. Out of 17 correlations, only negative affect ($r = -.20$), conflict ($r = -.19$), and SSCCS ($r = -.36$) were significantly associated with the response rate. When applying the Šidák error correction (Šidák, 1967), the correlation with SSCCS was still significant, which indicates that participants who reported higher levels of subjective depletion over the course of the study recruited less effort to adhere to the study protocol. Interestingly, resistance itself was not associated with adherence ($r = -.01$).

Analysis of the ego-depletion effect in daily life

The mixed models revealed that exerting prior resistance was significantly associated neither with resisting momentary desires, $b = -0.02$, $SE = 0.04$, 95% CI [−0.11, 0.06], $d_{av} = -0.03$, nor with momentary subjective depletion as assessed by the SSCCS, $b = 0.03$, $SE = 0.02$, 95% CI [−0.01, 0.07], $d_{av} = 0.11$. These results mirror the results in Study 1, in that we found very weak evidence that prior acts of self-control negatively impact current self-control efforts. However, the effects were not significant and, more importantly, were only negligibly small, on both the subjective and behavioural level.

Moderation analysis of the ego-depletion effect in daily life

As indicated in Tables 6 and 7, out of 24 coefficients for the univariable analysis of the subjective and behavioural consequences of consecutive self-control, all 95% CIs concluded zero, indicating no signs of substantial moderation. The picture in the multi-variable analysis was similar, except for trait self-control and state motivation. When participants exerted self-control during the last episode, participants with low levels of trait self-control (1 *SD* below the mean) reported higher levels of subjective depletion, $b = 0.11$, $SE = 0.02$, 95% CI [0.06, 0.15], whereas participants with high levels (1 *SD* above the mean) did not, $b = -0.00$, $SE = 0.02$, 95% CI [−0.05, 0.04], respectively. In turn, when participants did not resist earlier, current resistance was not associated with subjective depletion in participants with either low levels of trait self-control, $b = 0.01$, $SE = 0.02$, 95% CI [−0.04, 0.06], or high levels, $b = 0.04$, $SE = 0.03$, 95% CI [−0.01, 0.09].

Moreover, participants high in self-control reported more not less enactment when they previously resisted, $b = -0.42$, $SE = 0.07$, 95% CI [−0.56, −0.29], compared with when they did not, $b = -0.55$, $SE = 0.07$, 95% CI [−0.69, −0.41], with a difference of $b = 0.13$, $SE = 0.06$, 95% CI [0.01, 0.25]. In turn, participants low in self-control reported less enactment when they previously resisted, $b = -0.63$, $SE = 0.06$, 95% CI [−0.75, −0.52], compared with when they did not, $b = -0.49$, $SE = 0.06$, 95% CI [−0.61, −0.36], with a difference of $b = -0.15$, $SE = 0.06$, 95% CI [−0.27, −0.02].

Table 6. Fixed- and random-effects estimates of enactment and subjective depletion as a function of consecutive self-control and state moderators in Study 2

	Enactment		Subjective depletion (SSCCS)	
Fixed effects	*b* (*SE*)	95% CI	*b* (*SE*)	95% CI
Univariable analyses				
Positive affect	0.29 (0.23)	[−0.17, 0.74]	−0.09 (0.10)	[−0.29, 0.11]
Negative affect	−0.36 (0.31)	[−0.96, 0.24]	0.13 (0.13)	[−0.14, 0.39]
Internal motivation	0.19 (0.23)	[−0.25, 0.64]	0.06 (0.14)	[−0.21, 0.33]
External motivation	0.41† (0.21)	[−0.01, 0.83]	0.08 (0.13)	[−0.17, 0.34]
Multi-variable analyses				
Positive affect	−0.34 (0.39)	[−1.11, 0.43]	0.06 (0.20)	[−0.34, 0.46]
Negative affect	−0.43 (0.48)	[−1.36, 0.51]	0.39 (0.25)	[−0.10, 0.88]
Internal motivation	**0.51† (0.25)**	[0.01, 1.01]	0.07 (0.13)	[−0.18, 0.33]
External motivation	**0.63† (0.24)**	[0.16, 1.10]	0.03 (0.13)	[−0.22, 0.28]
Random effects (multi-variable)	*Var* (*SE*)	95% CI	*Var* (*SE*)	95% CI
Resistance	0.077 (0.020)	[0.047, 0.128]	0.001 (0.002)	[0.001, 0.030]
Temptation	0.053 (0.20)	[0.026, 0.111]	0.018 (0.007)	[0.009, 0.039]
Participant	0.010 (0.003)	[0.005, 0.198]	0.018 (0.003)	[0.013, 0.026]
Day	—	—	—	—
Residual	0.065 (0.003)	[0.060, 0.70]	0.018 (0.003)	[0.017, 0.019]

Note: Estimates in bold do not include zero in their 95% CI. SSCCS, State Self-Control Capacity Scale.
†Estimates do not include zero in their 95% CI when crossed effects are removed from the model.
‡Estimates still include zero in their 95% CI when adjusting for multiple comparisons via the false discovery rate.

DOI: 10.1002/per

Table 7. Fixed- and random-effects estimates of enactment and subjective depletion as a function of consecutive self-control and trait moderators in Study 2

	Enactment		Subjective depletion (SSCCS)	
Fixed effects	*b* (*SE*)	95% CI	*b* (*SE*)	95% CI
Univariable analyses				
Age	0.43 (0.36)	[−0.28, 1.14]	−0.15 (0.18)	[−0.50, 0.21]
Gender	0.07 (0.11)	[−0.16, 0.29]	−0.01 (0.06)	[−0.12, 0.11]
Self-control	0.44 (0.29)	[−0.13, 1.02]	−0.23 (0.14)	[−0.51, 0.06]
Openness	0.27 (0.36)	[−0.43, 0.98]	−0.11 (0.18)	[−0.46, 0.24]
Conscientiousness	0.11 (0.24)	[−0.37, 0.58]	−0.03 (0.12)	[−0.27, 0.20]
Extraversion	−0.27 (0.27)	[−0.79, 0.25]	−0.09 (0.13)	[−0.35, 0.17]
Agreeableness	−0.15 (0.36)	[−0.86, 0.57]	−0.28 (0.18)	[−0.63, 0.07]
Neuroticism	0.21 (0.29)	[−0.35, 0.77]	−0.03 (0.14)	[−0.30, 0.25]
Multi-variable analyses				
Age	0.43 (0.42)	[−0.39, 1.26]	−0.25 (0.21)	[−0.65, 0.16]
Gender	0.13 (0.15)	[−0.16, 0.42]	0.07 (0.07)	[−0.07, 0.21]
Self-control	**1.29**† **(0.49)**	[0.34, 2.25]	**−0.65**† **(0.24)**	[−1.12, −0.18]
Openness	0.50 (0.43)	[−0.33, 1.34]	−0.10 (0.21)	[−0.51, 0.32]
Conscientiousness	−0.30 (0.42)	[−1.13, 0.52]	**0.43**† **(0.21)**	[0.03, 0.84]
Extraversion	−0.57 (0.32)	[−1.19, 0.06]	0.06 (0.16)	[−0.25, 0.37]
Agreeableness	−0.27 (0.42)	[−1.09, 0.55]	−0.35 (0.21)	[−0.75, 0.06]
Neuroticism	0.35 (0.32)	[−0.27, 0.98]	−0.15 (0.16)	[−0.46, 0.16]
Random effects (multi-variable)	*Var* (*SE*)	95% CI	*Var* (*SE*)	95% CI
Resistance	0.050 (0.013)	[0.030, 0.085]	0.000 (0.000)	[0.000, 0.001]
Temptation	0.031 (0.015)	[0.012, 0.079]	0.022 (0.008)	[0.011, 0.046]
Person	0.018 (0.004)	[0.012, 0.026]	0.013 (0.002)	[0.010, 0.018]
Day	0.001 (0.001)	[0.000, 0.003]	0.000 (0.000)	[0.000, 0.007]
Residual	0.091 (0.003)	[0.086, 0.096]	0.022 (0.001)	[0.020, 0.025]

Note: Estimates in bold do not include zero in their 95% CI. SSCCS, State Self-Control Capacity Scale.
†Estimates do not include zero in their 95% CI when crossed effects are removed from the model.
‡Estimates still include zero in their 95% CI when adjusting for multiple comparisons via the false discovery rate.

In regard to motivation, unpacking the interaction reveals that participants high in internal or external motivation (+1 *SD* above the mean) reported less success in resisting current desires that conflicted with a self-set goal when they previously resisted, $b = -0.47$, $SE = 0.06$, 95% CI [−0.58, −0.36], and $b = -0.45$, $SE = 0.05$, 95% CI [−0.56, −0.35], respectively, compared with when they did not, $b = -0.60$, $SE = 0.06$, 95% CI [−0.71, −0.48], and $b = -0.59$, $SE = 0.06$, 95% CI [−0.71, −0.48], respectively.

However, when the 95% CIs were adjusted for multiple testing using the false discovery rate, the confidence of the moderating associations was lower, as reflected in that the CIs of all three-way interactions then included zero.

DISCUSSION

Study 2 investigated ego depletion in daily life, operationalized as the behavioural and cognitive consequences of resisting current desires, depending on prior instances of self-control. We did not find compelling evidence for ego-depletion effects in daily life: for example, prior acts of self-control did not lead to impaired current resistance efforts towards temptations on the cognitive and behavioural level of analysis, as evidenced by very small effect sizes, which equated to $d_{av} = -0.03$ for enactment and $d_{av} = 0.11$ for subjective depletion. Moreover, only a few of the moderators could significantly explain this variability but only when multiple testing was not accounted for. Furthermore, the effect of trait self-control was only significant when the other moderators, such as conscientiousness, were also entered into the model (that is, when the shared variance with the other moderators was removed from the model). We therefore suggest caution when further interpreting this effect and believe that further research is needed before concluding that ego depletion on the cognitive level can only be observed in individuals with low levels of trait self-control. Moreover, we found it difficult to explain the effect why high levels of both internal and external motivation were associated with less success in resisting momentary desires.

GENERAL DISCUSSION

Although we found only marginal ego-depletion effects, both in a laboratory setting (Study 1) and in daily life (Study 2), with only very small, practically meaningless effect sizes around an absolute value of $d_{av} = 0.05$, applying the cross-over design to examine ego depletion revealed a relatively large interindividual variability of the ego-depletion effect, evidencing the existence of moderators that can explain these differences. In both studies, we investigated the most prominent moderators of ego depletion, namely, affect, motivation, and self-control relevant personality traits. Although we

DOI: 10.1002/per

found some significant moderators, these moderating influences mostly disappeared when we controlled for multiple testing using the false discovery rate. Although this was somewhat disappointing for us, and we would have been glad to present rosier results, these findings demonstrate the importance of capturing other sources of variance, such as crossed random factors and adjusting α levels to the number of tests pertaining to a specific research question. Given that more coefficients were significant when only the random intercepts of participants were included in the random part, these findings also illustrate the importance of random effects, demonstrating that the fixed effect is associated with much more uncertainty when the fixed effect is averaged over the slopes for each moderator and for each condition, compared with when the fixed effect is directly estimated without allowing variability in the slopes between the participants. Future research should thus not only employ within-person designs such as the crossover design or ambulatory assessment to increase power but also better reflect the underlying processes and differences.

Implications for future research

Given the large interindividual variability in the ego-depletion effect in Study 1, it may be surprising that even moderators close to the theoretical concept of ego depletion such as motivation and trait self-control did not explain significant variability. One explanation could be that the operationalization of both task motivation and trait self-control might not have captured the self-control demands of the second task. Motivational accounts of ego depletion highlight the role of conservation and reward motivation (e.g. Inzlicht & Schmeichel, 2012), and research has shown that depleted participants engaged less in an effortful task for which they could receive monetary rewards (Giacomantonio, Jordan, Fennis & Panno, 2014). However, none of the motivation measures we employed in Studies 1 and 2 were constructed to assess the specific motivational processes but rather to assess other aspects associated with current motivation, such as challenge and interest in Study 1 or internal and external motivation in Study 2. Because validated measures assessing state reward motivation were not available when planning and conducting this study, future research may develop tailored motivation questionnaires (e.g. by adopting the BIS/BAS questionnaire for momentary changes) that may be better suited to capturing motivational processes underlying consecutive self-control.

In a similar vein, research on cognitive control had previously assumed cognitive control to be the ability to overcome one's inner responses (e.g. Posner & Snyder, 1975); this has branched out to a more diverse concept, for example, by differentiating between the processes of inhibition, updating, and switching (Miyake, Friedman, Emerson, Witzki, Howerter, & Wager, 2000). The MSIT in the second task requires primarily inhibition, whereas the dual-task design also involves switching, although in a much more elongated form than operationalized in cognitive control research. However, out of the 13 items of the BSCS we used in the present study, only three items mainly involve inhibition (Item 1: 'I am good at resisting temptations', Item 9: 'Pleasure and fun sometimes keep me from getting work done', and Item 12: 'Sometimes I cannot stop myself from doing something, even if I know it is wrong'); the other items assess other aspects of self-control (e.g. Item 3: 'I am lazy' or Item 7: 'I wish I had more self-discipline'). Unsurprisingly, research has thus shown that self-reports such as the BSCS do not correlate with tasks predominantly involving inhibition such as the MSIT or the Stroop task (Duckworth & Kern, 2011; Saunders, Milyavskaya, Etz, Randles, & Inzlicht, 2018). Instead, trait self-control may be more strongly related to not being tempted in the first place, instead of with better performance in inhibiting conflicting desires (Hofmann et al., 2012). To connect to this research, we computed a linear regression to predict temptation strength in daily life based on trait self-control, controlling for wave and group. We found a small effect, $b = -0.018$, $SE = 0.007$, 95% CI [−0.033, −0.004], $\beta = -0.22$, and the 95% CI did not include zero, such that participants with high compared with low levels of trait self-control reported weaker temptations on average. Similarly, Milyavskaya and Inzlicht (2017) reported that trait self-control was not significantly associated with goal attainment and depletion in daily life but rather with the experience of temptations. The role of temptations in self-control is further highlighted in our data, whereby temptation strength was only positively associated with enactment of desires in individuals with low but not high levels of trait self-control. With regard to consecutive control processes, trait self-control was a significant moderator for subjective depletion, such that individuals with high levels of trait self-control who were confronted with prior temptations reported higher subjective depletion when confronted with current temptations but not when they currently resisted. Taken together, we believe that future research would benefit from gaining a deeper understanding of the many aspects of self-control by exploring how different lines of research could be connected with and inform each other, such as aspects of executive function in self-control (inhibition, updating, and switching), state resistance and temptations strength, and facets of trait self-control.

Another explanation for these null effects could be that the ego-depletion effect, both as assessed by the crossover design and in general, might not be reliable. This is a crucial issue because a moderator would only explain variance in the ego-depletion effect if this variance was largely systematic and thus attributable to other factors. In Study 1, we assessed test–retest reliability of the MSIT measures in the second task and found that only the $RT_{incongruent}$ showed good values of test–retest reliability; in contrast, the MSIT effect and errors made had only moderate test–retest reliability at best, with values around $r = .50$. We therefore recommend focusing on $RT_{incongruent}$ when using the MSIT. Although we controlled for the order of the conditions, other unmeasured differences between the two-order conditions could have confounded these findings. However, it is important to note that this analysis does not assess the reliability of ego-depletion effects, and we are not aware of any study that has investigated this crucial issue. The crossover design has promise, in that it offers an opportunity to estimate the

DOI: 10.1002/per

test–retest reliability of the ego-depletion effect, specifically by employing two crossover designs with a long washout period between each crossover session.

Furthermore, even if the MSIT and the ego-depletion effect were robust and reliable, associations on the trait level might be low because of low systematic between-person variability of the MSIT, which some call the *reliability paradox* (Hedge, Powell, & Sumner, 2018). Hedge et al. (2018) argue that the experimental effects of cognitive tasks such as the MSIT or the Stroop task only become well established in the case of low between-person variability. However, this condition weakens correlations with trait moderators, undermining conclusions that could be based on the correlational relationships between cognitive tasks and personality traits. Please note that this could only explain the null effects for trait moderators in Study 1, because state moderators could also explain within-person variability.

Lastly, we found an interesting effect in our data that views ego depletion from a different angle: whereas most variables relevant to Study 2 were not significantly predictive of adherence to the study protocol in Study 2, the best predictor was subjective depletion ($r = -.36$). This was moderate in size and more than three times larger than the twice largest correlation ($r = -.20$). Thus, participant who reported being more subjectively depleted over the course of Study 2 completed less ambulatory assessments. One possible explanation for this effect could be that subjectively depleted participants try to avoid effort (Inzlicht, Shenhav, & Olivola, 2018) such as completing six ambulatory assessments over the course of 40 days. To further unpack this process, we compared the momentary level of subjective depletion before a response and before a missing, which did not reveal a significant association, $b = 0.003$, $SE = 0.003$, 95% CI [−0.003, 0.007], and thus, the associations could not be observed on the within-person level. Given that Study 2 consisted of conditional branching, such that items on resistance and enactment were only presented if the participants currently experienced a desire (or within the last 30 minutes), we examined whether subjectively depleted participants further reduced their effort by reporting fewer desires over the course of Study 2 in order to avoid the additional items. Whereas we, again, did not find an association on the within-person level, $b = 0.002$, $SE = 0.001$, 95% CI [−0.001, 0.004], the association on the between-person level was significant, $b = -0.005$, $SE = 0.001$, 95% CI [−0.007, −0.002], which equates to $d_{av} = -0.29$ based on the estimated marginals and their standard errors, such that participants with higher levels of subjective depletion in general reported fewer desires over the course of the study than participants with lower levels of subjective depletion. Future research might, thus, take a closer look at how individuals perceive and try to avoid effort in daily life.

To sum up, we could not find convincing evidence in either laboratory or ambulatory settings that acts of self-control led to subsequent self-control limitations in the short term. Despite relatively large interindividual variability of the ego-depletion effect, we also did not find evidence of moderation by affect, motivation, or self-control relevant personality traits. However, we still believe that ego depletion is an important research topic that should be studied under conditions that align more closely to the everyday environment of individuals. We suggest that ego-depletion research turn to other fields of research that explore self-control performance over longer periods of time, such as in long tasks (drawing on the mental fatigue literature; e.g. Boksem & Tops, 2008; Gergelyfi, Jacob, Olivier, & Zénon, 2015), studies on effort (Inzlicht et al., 2018), or when individuals engage in real-world problems such as dieting (drawing on the literature on habit formation; Neal, Wood, & Drolet, 2013), or study.

Limitations

We used the same tasks in Study 1 that were used in the multi-lab replication project of ego depletion (Hagger et al., 2016), which uses a computerized version of the letter-e task. This computerized task has been criticized because it may not be depleting enough to yield an ego-depletion effect (Baumeister & Vohs, 2016): instead of establishing a habit first by letting the participants cross out each instance of the letter ‘e’, both the replication project and the original study (Sripada et al., 2014) skip this habit-forming step; as a result, there is no habit that needs to be overridden effortfully. However, research shows that the computerized letter-e task is as effective in invoking performance decrements as the pen-and-paper version (Arber et al., 2017). Moreover, the participants in Study 1 had to complete 20 practice trials, which was not the case in the replication project (Hagger et al., 2016). Although we believe that more practice trials would have been better in establishing a stronger habit, participants in the depletion condition made more errors in the first task (10%) compared with participants in the control condition (1%). Although this does not show that a habit had been established, it does show that the first task was more demanding by additionally requiring to assess whether the ‘e’ was flanked by a vowel.

We also used the same manipulation checks used by the replication project (Hagger et al., 2016), which were criticized given the absence of effects of the depletion manipulation on the fatigue scale in the replication project. However, to combat these problems, we first investigated the factor structure of the manipulation check items and found two factors, which we labelled ‘task difficulty’ and ‘fatigue’. These two factors showed good fit characteristics and strong measurement invariance across both lab sessions and ego-depletion manipulation. Using these two factors, the results demonstrated larger effect sizes of the depletion manipulation for difficulty than for fatigue, reflecting that participants found the depletion more difficult than fatiguing relative to the control condition. However, the effect on the fatigue scale was still medium to large in size, demonstrating the more fatiguing nature of the depletion condition compared with the control condition.

Another limitation of Study 1 pertains to the use of a crossover design, which is susceptible to carry-over effects (i.e. effects of the intervention are carried over from one session to the next one) and practice effects (i.e. effects due to performing the same tasks multiple times; Wellek & Blettner, 2012). Regarding a carry-over effect, the interaction between

DOI: 10.1002/per

depletion manipulation and period was significant for neither RTs nor errors made, which indicates no evidence for a carry-over effect. However, period itself was significant in all models, such that participants generally performed better in the second lab sessions, independent of whether this session was the depletion or control condition. Practice effects can be combated with washout periods. We used a washout period of 2 weeks, which was evidently not long enough to avoid practice effects. However, a carry-over effect was not present for any of the three MSIT outcome. Moreover, although this practice effect is a problem, it also shows a brighter side of consecutive self-control: participants are not fragile individuals who are demotivated or depleted after 7.5 minutes of initial self-control but rather learn and adapt to the self-control demands at hand, the effects of which can be detected even 2 weeks later. Moreover, the practice effects were larger than the depletion effects, d_{av} = 0.38 versus d_{av} = 0.06 for RTs, which demonstrates the ability of participants to adapt. This finding probably also reflects the artificial nature of these tasks (which may be more prone to practice effects than real-world challenges), which may limit the size of the ego-depletion effect.

An additional limitation of Studies 1 and 2 is that our sample is not specifically tailored to some of the moderators such as gender and age given that our sample predominately consisted of young participants who identify themselves as women. Moreover, in order to maximize the statistical power within the scope of our possibilities, we did not manipulate either motivation or affect but instead used measured variables based on self-reports to test for moderation, which is likely not equivalent and might reduce the size of the associations.

One limitation of Study 2 was the burden placed on participants, with up to 50 items administered in the study for 6 weeks; this likely played a strong role in the 25% missed-signal rate. However, we would like to note that this is a general problem associated with most experience sampling conducted over a longer time period and is not specific to our study. Moreover, we still managed to collect an average of approximately 180 episodes per participant, which compares favourably to other ambulatory assessment studies.

Conclusion

In the present study, we investigated how ego depletion in a laboratory setting and in daily life was moderated by affect, motivation, and personality traits. We found no compelling or consistent evidence for ego-depletion effects or moderating influences, in either the laboratory or an ambulatory setting. Our results demonstrate the need for new pathways and research avenues to gain a better understanding of an important phenomenon that is worth studying despite the methodical and conceptual problems of this field.

ACKNOWLEDGEMENTS

This research was supported by a scholarship from the Stipendienstiftung Rheinland-Pfalz awarded to Zarah Rowland. We are thankful to our research assistants Maike Vogel, Julius Welzel, Markus Müssig, and Kimberly Holtz who assisted us in the organization and in the data collection of the study.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

REFERENCES

Agrawal, N., & Wan, E. W. (2009). Regulating risk or risking regulation? Construal levels and depletion effects in the processing of health messages. *Journal of Consumer Research*, *36*, 448–462. https://doi.org/10.1086/597331.

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

Baayen, R. H., & Milin, P. (2010). Analyzing reaction times. *International Journal of Psychological Research*, *3*, 12–28. https://doi.org/10.21500/20112084.807.

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. https://doi.org/10.1037/0022-3514.74.5.1252.

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

Baumeister, R. F., Tice, D. M., & Vohs, K. D. (2018). The strength model of self-regulation: Conclusions from the second decade of willpower research. *Perspectives on Psychological Science*, *13*, 141–145. https://doi.org/10.1177/1745691617716946.

Baumeister, R. F., & Vohs, K. D. (2016). Misguided effort with elusive implications. *Perspectives on Psychological Science*, *11*, 574–575. https://doi.org/10.3389/fpsyg.2016.01639.

Benjamini, Y., & Yekutieli, D. (2005). False discovery rate-adjusted multiple confidence intervals for selected parameters. *Journal of the American Statistical Association*, *100*, 71–81. https://doi.org/10.1198/016214504000001907.

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. https://doi.org/10.1026/0012-1924.55.1.2.

Bertrams, A., Unger, A., & Dickhäuser, O. (2011). Momentan verfügbare Selbstkontrollkraft—Vorstellung eines Messinstruments und erste Befunde aus pädagogisch-psychologischen Kontexten [Momentarily available self-control strength—Introduction of a measure and first findings from educational–psychological contexts]. *Zeitschrift für Pädagogische Psychologie*, *25*, 185–196. https://doi.org/10.1024/1010-0652/a000042.

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

Borkenau, P., & Ostendorf, F. (2008). *NEO-FFI: NEO-Fünf-Faktoren-Inventar nach Costa und McCrae (2., neu normierte und vollständig überarbeitete Auflage) [NEO-FFI: NEO-five-factor-inventory by Costa and McCrae; 2nd ed. including new norms]. Göttingen, Germany: Hogrefe Verlag*.

Boucher, H. C., & Kofos, M. (2012). The idea of money counteracts ego depletion effects. *Journal of Experimental Social Psychology*, *48*, 804–810. https://doi.org/10.1016/j.jesp.2012.02.003.

DOI: 10.1002/per

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. https://doi.org/10.1038/sj.mp.4001217.

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. https://doi.org/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. https://doi.org/10.3389/fpsyg.2014.00823.

Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS scales. *Journal of Personality and Social Psychology*, *67*, 319–333. https://doi.org/10.1037/0022-3514.67.2.319.

Cohen, P., Cohen, J., Aiken, L. S., & West, S. G. (1999). The problems of units and the circumstance for POMP. *Multivariate Behavioral Research*, *34*, 315–346. https://doi.org/10.1207/S15327906MBR3403_2.

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. https://doi.org/10.1037/a0014604.

Cunningham, M. R., & Baumeister, R. F. (2016). How to make nothing out of something: Analyses of the impact of study sampling and statistical interpretation in misleading meta-analytic conclusions. *Frontiers in Psychology*, *7*, 1639. https://doi.org/10.3389/fpsyg.2016.01639.

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

Dai, H., Milkman, K. L., Hofmann, D. A., & Staats, B. R. (2015). The impact of time at work and time off from work on rule compliance: The case of hand hygiene in healthcare. *Journal of Applied Psychology*, *100*, 846–862. https://doi.org/10.1037/a0038067.

De Houwer, J. (2011). Why the cognitive approach in psychology would profit from a functional approach and vice versa. *Perspectives on Psychological Science*, *6*, 202–209. https://doi.org/10.1177/1745691611400238.

DeWall, C. N., Baumeister, R. F., Stillman, T. F., & Gailliot, M. T. (2007). Violence restrained: Effects of self-regulation and its depletion on aggression. *Journal of Experimental Social Psychology*, *43*, 62–76. https://doi.org/10.1016/j.jesp.2005.12.005.

Duckworth, A. L., & Kern, M. L. (2011). A meta-analysis of the convergent validity of self-control measures. *Journal of Research in Personality*, *45*, 259–268. https://doi.org/10.1016/j.jrp.2011.02.004.

De Vet, E., Nelissen, R. M. A., Zeelenberg, M., & De Ridder, D. T. D. (2013). Ain't no mountain high enough? Setting high weight loss goals predicts effort and short-term weight loss. *Journal of Health Psychology*, *18*, 638–647. https://doi.org/10.1177/1359105312454038.

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. https://doi.org/10.1177/1088868313507533.

Eisenberger, R. (1992). Learned industriousness. *Psychological Review*, *99*, 248–267. https://doi.org/10.1037/0033-295X.99.2.248.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*, 175–191.

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. https://doi.org/10.1016/j.jesp.2009.09.008.

Francis, Z., Milyavskaya, M., Lin, H., & Inzlicht, M. (2018). Development of a within-subject, repeated-measures ego-depletion paradigm. *Social Psychology*, *49*, 271–286. https://doi.org/10.1027/1864-9335/a000348.

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. https://doi.org/10.1016/j.paid.2011.05.033.

Friese, M., Loschelder, D. D., Gieseler, K., Frankenbach, J., & Inzlicht, M. (2018). Is ego depletion real? An analysis of arguments. *Personality and Social Psychology Review*. Advance online publication. https://doi.org/10.1177/1088868318762183.

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*, 176. https://doi.org/10.3389/fnbeh.2015.00176.

Giacomantonio, M., Jordan, J., Fennis, B. M., & Panno, A. (2014). When the motivational consequences of ego depletion collide: Conservation dominates over reward-seeking. *Journal of Experimental Social Psychology*, *55*, 217–220. https://doi.org/10.1016/j.jesp.2014.07.009.

Giner-Sorolla, R. (2018, January 2). *Powering your interaction*. Retrieved from https://approachingblog.wordpress.com/2018/01/24/powering-your-interaction-2/

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

Gröpel, P., Baumeister, R. F., & Beckmann, J. (2014). Action versus state orientation and self-control performance after depletion. *Personality and Social Psychology Bulletin*, *40*, 476–487. https://doi.org/10.1177/0146167213516636.

Guay, F., Vallerand, R. J., & Blanchard, C. (2000). On the assessment of situational intrinsic and extrinsic motivation: The Situational Motivation Scale (SIMS). *Motivation and Emotion*, *24*, 175–213. https://doi.org/10.1023/A:1005614228250.

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. https://doi.org/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. https://doi.org/10.1037/a0019486.

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

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. https://doi.org/10.1037/a0026545.

Hughes, J. R., Keely, J., & Naud, S. (2004). Shape of the relapse curve and long-term abstinence among untreated smokers. *Addiction*, *99*, 29–38. https://doi.org/10.1111/j.1360-0443.2004.00540.x.

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. https://doi.org/10.1002/per.1899.

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. https://doi.org/10.1177/1745691612454134.

Inzlicht, M., Shenhav, A., & Olivola, C. (2018). The effort paradox: Effort is both costly and valued. *Trends in Cognitive Science*, *22*, 337–349. https://doi.org/10.1016/j.tics.2018.01.007.

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. https://doi.org/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. https://doi.org/10.1177/0956797610384745.
John, O. P., Donahue, E. M., & Kentle, R. L. (1991). *The Big Five Inventory—Versions 4a and 54*. Berkeley: University of California, Berkeley, Institute of Personality and Social Research.
Lang, F. R., Lüdtke, O., & Asendorpf, J. B. (2001). Testgüte und psychometrische Äquivalenz der deutschen version des big five inventory (BFI) bei jungen, mittelalten und alten Erwachsenen [Validity and psychometric equivalence of the German version of the big five inventory in young, middle-aged and old adults]. *Diagnostica*, *47*, 111–121. https://doi.org/10.1026//0012-1924.47.3.111.
Johnston, D. W., Allan, J. L., Powell, D. J. H., Jones, M. C., Farqharson, B., … Johnston, M. (2018). Why does work cause fatigue? A real-time investigation of fatigue, and determinants of fatigue in nurses working 12-hour shifts. *Annals of Behavioral Medicine*, *53*, kay065–kay562. https://doi.org/10.1093/abm/kay065.
Judd, C. M., Westfall, J., & Kenny, D. A. (2017). Experiments with more than one random factor: Designs, analytic models, and statistical power. *Annual Review of Psychology*, *68*, 601–625. https://doi.org/10.1146/annurev-psych-122414-033702.
Kool, W., & Botvinick, M. (2014). A labor/leisure trade-off in cognitive control. *Journal of Experimental Psychology: General*, *143*, 131–141. https://doi.org/10.1037/a0031048.
Kuppens, P., Allen, N. B., & Sheeber, L. B. (2010). Emotional inertia and psychological maladjustment. *Psychological Science*, *21*, 984–991. https://doi.org/10.1177/0956797610372634.
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. https://doi.org/10.1017/S0140525X12003196.
Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, *4*, 863. https://doi.org/10.3389/fpsyg.2013.00863.
Lo, S., & Andrews, S. (2015). To transform or not to transform: Using generalized linear mixed models to analyse reaction time data. *Frontiers in Psychology*, *6*, 1171. https://doi.org/10.3389/fpsyg.2015.01171.
Lorist, M. M., & Jolij, J. (2012). Trial history effects in Stroop task performance are independent of top-down control. *PLoS ONE*, *7*, e39802. https://doi.org/10.1371/journal.pone.0039802.
Loschelder, D. D., & Friese, M. (2016). Moderators of the ego depletion effect. In E. R. Hirt, J. J. Clarkson, & L. Jia (Eds.), *Self-regulation and ego control* (pp. 21–42). San Diego, CA: Elsevier, https://doi.org/10.1016/B978-0-12-801850-7.00002-0.
Mccrae, R. R., & Costa, P. T. (2004). A contemplated revision of the NEO five-factor inventory. *Personality and Individual Differences*, *36*, 587–596. https://doi.org/10.1016/S0191-8869(03)00118-1.
Meece, J. L., & Painter, J. (2008). Gender, self-regulation, and motivation. In D. H. Schunk, & B. J. Zimmerman (Eds.), *Motivation and self-regulated learning: Theory, research, and applications* (pp. 339–367). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.
Meule, A., Vögele, C., & Kübler, A. (2011). Psychometric evaluation of the German Barratt Impulsiveness Scale—short version (BIS-15). *Diagnostica*, *57*, 126–133. https://doi.org/10.1026/0012-1924/a000042.
Milyavskaya, M., & Inzlicht, M. (2017). What's so great about self-control? Examining the importance of effortful self-control and temptation in predicting real-life depletion and goal attainment. *Social Psychological and Personality Science*, *8*, 603–611. https://doi.org/10.1177/1948550616679237.
Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*, 49–100. https://doi.org/10.1006/cogp.1999.0734.
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. https://doi.org/10.1177/0146167206288008.
Muraven, M., Gagné, M., & Rosman, H. (2008). Helpful self-control: Autonomy support, vitality, and depletion. *Journal of Experimental Social Psychology*, *44*, 573–585. https://doi.org/10.1016/j.jesp.2007.10.008.
Muraven, M., & Slessareva, E. (2003). Mechanisms of self-control failure: Motivation and limited resources. *Personality and Social Psychology Bulletin*, *29*, 894–906. https://doi.org/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. https://doi.org/10.1037/0022-3514.74.3.774.
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. https://doi.org/10.1037/a0032626.
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. https://doi.org/10.1371/journal.pone.0147275.
Nolen-Hoeksema, S., & Corte, C. (2004). Gender and self-regulation. In R. F. Baumeister, & K. D. Vohs (Eds.), *Handbook of self-regulation: Research, theory, and applications* (pp. 411–421). New York, NY, US: The Guilford Press.
O'Connell, K. A., Schwartz, J. E., & Shiffman, S. (2008). Do resisted temptations during smoking cessation deplete or augment self-control resources? *Psychology of Addictive Behaviors*, *2*, 486–495. https://doi.org/10.1037/0893-164X.22.4.486.
Robinson, G. K. (1991). That BLUP is a good thing: The estimation of random effects. *Statistical Science*, *6*, 15–32. https://doi.org/10.1214/ss/1177011926.
Peugh, J. L. (2010). A practical guide to multilevel modeling. *Journal of School Psychology*, *48*, 85–112. https://doi.org/10.1016/j.jsp.2009.09.002.
Posner, M. I., & Snyder, C. R. R. (1975). Attention and cognitive control. In R. L. Solso (Ed.), *Information processing and cognition* (pp. 55–85). Hillsdale, NJ: Erlbaum.
Randles, D., Harlow, I., & Inzlicht, M. (2017). A pre-registered naturalistic observation of within domain mental fatigue and domain-general depletion of self-control. *PLoS ONE*, *12*, e0182980. https://doi.org/10.1371/journal.pone.0182980.
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. https://doi.org/10.1026//0012-1924.47.2.57.
Rowland, Z., Wenzel, M., & Kubiak, T. (2016). The effects of computer-based mindfulness training on self-control and mindfulness within ambulatorily assessed network systems across health-related domains in a healthy student population (SMASH): Study protocol for a randomized controlled trial. *Trials*, *17*, 570. https://doi.org/10.1186/s13063-016-1707-4.
Rowland, Z.*, Wenzel, M.*, & Kubiak, T. (in press). A mind full of happiness: How mindfulness shapes affect dynamics in daily life. *Emotion*.
Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology*, *39*, 1161–1178.
Saunders, B., Milyavskaya, M., Etz, A., Randles, D., & Inzlicht, M. (2018, January 19). *Reported self-control is not meaningfully associated with inhibition-related executive function: A*

DOI: 10.1002/per

Bayesian analysis. Retrieved from https://doi.org/10.31234/osf.io/bxfsu

Savani, K., & Job, V. (2017). Reverse ego-depletion: Acts of self-control can improve subsequent performance in Indian cultural contexts. *Journal of Personality and Social Psychology*, *113*, 589–607. https://doi.org/10.1037/pspi0000099.

Schielzeth, H., & Forstmeier, W. (2009). Conclusions beyond support: Overconfident estimates in mixed models. *Behavioral Ecology*, *20*, 416–420. https://doi.org/10.1093/beheco/arn145.

Šidák, Z. (1967). Rectangular confidence region for the means of multivariate normal distributions. *Journal of the American Statistical Association*, *62*, 626–633.

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

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, NY: Guilford Press.

Sievertsen, H. H., Gino, F., & Piovesan, M. (2016). Cognitive fatigue influences students' performance on standardized tests. *Proceedings of the National Academy of Sciences of the United States of America*, *113*, 2621–2624. https://doi.org/10.1073/pnas.1516947113.

Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2014). P-curve: A key to the file-drawer. *Journal of Experimental Psychology: General*, *143*, 534–547. https://doi.org/10.1037/a0033242.

Spinella, M. (2007). Normative data and a short form of the Barratt Impulsiveness Scale. *International Journal of Neuroscience*, *117*, 359–368. https://doi.org/10.1080/00207450600588881.

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

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

Strobel, A., Beauducel, A., Debener, S., & Brocke, B. (2001). Eine deutschsprachige Version des BIS/BAS-Fragebogens von Carver und White [A German version of Carver and White's BIS/BAS scales]. *Zeitschrift für Differentielle und Diagnostische Psychologie*, *22*, 216–227. https://doi.org/10.1024//0170-1789.22.3.216.

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. https://doi.org/10.1111/j.0022-3506.2004.00263.x.

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. https://doi.org/10.1016/j.jesp.2006.05.007.

Twenge, J. M., Muraven, M., & Tice, D. M. (2004). *Measuring state self-control: Reliability, validity, and correlations with physical and psychological stress*. San Diego, CA: San Diego State University.

Vohs, K. D., Baumeister, R. F., & Schmeichel, B. J. (2013). Erratum to "motivation, personal beliefs, and limited resources all contribute to self-control". *Journal of Experimental Social Psychology*, *49*, 184–188. https://doi.org/10.1016/j.jesp.2012.08.007.

Wellek, S. (2010). *Testing statistical hypotheses of equivalence and noninferiority* (2nd ed.). Boca Raton, FL: CRC Press, https://doi.org/10.1201/EBK1439808184.

Wellek, S., & Blettner, M. (2012). On the proper use of the crossover design in clinical trials: Part 18 of a series on evaluation of scientific publications. *Deutsches Ärzteblatt International*, *109*, 276–281. https://doi.org/10.3238/arztebl.2012.0276.

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 Social Psychology*, *43*, 175–184. https://doi.org/10.1002/ejsp.1936.

Wenzel, M., Lind, M., Rowland, Z., Zahn, D., & Kubiak, T. (in press). The limits of ego depletion: A crossover study on the robustness of performance deterioration in consecutive tasks. *Social Psychology*.

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. https://doi.org/10.1080/02699931.2013.851069.

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. https://doi.org/10.1371/journal.pone.0157009.

DOI: 10.1002/per