

Running head: PUPIL SIZE AND SELF-CONTROL FAILURE

TITLE

Changes in Pupil Size Track Self-Control Failure

AUTHORS

Sean R. O'Bryan*¹, Mindi M. Price¹, Jessica L. Alquist¹, Tyler Davis¹, Miranda Scolari¹

AFFILIATIONS

¹ Department of Psychological Sciences, Texas Tech University, Lubbock, TX, 79409

CORRESPONDING AUTHOR

Correspondence to: Sean R. O'Bryan (sean_obryan@brown.edu)

Cognitive, Linguistic & Psychological Sciences
Brown University
Box 1821
Metcalf Research Building
Providence, RI 02912

Word count (main text): 6,991

Figures: 4

Abstract

Previous studies have found that people are more likely to perform poorly on a self-control task following a previous task requiring self-control (ego-depletion), but the mechanism for this effect remains unclear. We used pupillometry to test the role of cognitive effort in ego-depletion effects. We hypothesized that an elevated pupil diameter (PD)—a common physiological measure of cognitive effort-- during an initial task requiring self-control should be negatively associated with performance on a subsequent control task. To test this hypothesis, we recorded participants' PD throughout a standard ego-depletion paradigm, where participants were assigned to exert either relatively high or low levels of attentional control. Next, we measured all participants' tonic PD and behavioral performance on a subsequent Stroop task. We show that in the initial task, task-associated PD was larger among participants who were assigned to the demanding attention condition. Furthermore, ego-depletion effects were serially mediated by PD: elevated and task-relevant PD emerged rapidly among the experimental group during the manipulation, persisted as a tonic response during the Stroop task, and predicted worse accuracy on incongruent trials. Notably, cognitive effort-related changes in PD are routinely associated with neural responses from the locus coeruleus-norepinephrine (LC-NE) system. This suggests that self-control lapses following exertion are marked by and may even be the result of an elevated LC-NE response. We conclude by exploring the neuroscientific implications of these results within the context of current theories of self-control.

Keywords: Ego-depletion, Self-control, Eye tracking, Pupillometry, Locus coeruleus, Norepinephrine

1 After an individual effortfully controls their behavior, they may struggle to maintain a
2 consistent level of effort on a subsequent task (Baumeister et al., 1998; Muraven et al., 1998; for a
3 meta-analysis, see Hagger et al., 2010). This effect of one self-control act impairing a second task
4 is commonly referred to as ego-depletion in the social psychology literature. Self-control refers to
5 a variety of instances where a behavioral response must be inhibited (but see Fujita, 2011)—either
6 to a salient but irrelevant stimulus (Garrison et al., 2019) or to an immediately gratifying
7 temptation (Mischel et al., 1989). The former, characterized more specifically as cognitive or
8 attentional control, is the focus of the current project.

9 There is considerable controversy about why, when, and whether ego-depletion occurs
10 (Beedie & Lane, 2012; Carter et al., 2015; Dang, 2016; Dang et al., 2021; Evans et al., 2015;
11 Garrison et al., 2019; Hagger et al., 2016; Inzlicht et al., 2014; Kurzban et al., 2013; Vohs et al.,
12 2021). The controversy is due in part to difficulty identifying a mechanism for the effect (Frieese et
13 al., 2019; Inzlicht et al., 2014). The original conceptualization of ego-depletion was that self-
14 control is a limited resource that is drawn on and reduced by sustained exertion. However, efforts
15 to identify this resource, focusing largely on glucose, have been difficult to replicate (Dang, 2015;
16 Finley et al., 2019; Gailliot et al., 2007; Kurzban, 2010). Instead, recent behavioral models can
17 provide much-needed insight into identifying a physiological marker. Some researchers have
18 proposed that waning self-control is characterized by a motivational and attentional shift among
19 people who have recently endured a demanding task (Inzlicht et al., 2012; 2014; Schmeichel et
20 al., 2010). Lapses in self-control, then, are the byproduct of a corrective measure to regain
21 balance between labor (exploitation) and leisure (exploration). A consequence of this
22 motivational shift is that cues signaling a potential, immediate reward gain attentional priority at
23 the expense of cues signaling a need for control. This can lead to self-control failures

characteristic of ego-depletion when persistent task-focused efforts go unrewarded (Inzlicht et al., 2012; 2014).

This account lends itself well to psychophysiological testing: transitions between attentional states are accompanied by concomitant changes in pupil diameter (PD). PD can be generally classified into two (co-occurring) components: a stimulus-driven (phasic) response and a baseline (tonic) response. Relatively strong pupillary responses time-locked to task-relevant stimuli correlate with focused attention and considerable cognitive effort, while shallow responses indicate an “off focus” attentional state, where behavioral responses tend to be less accurate and highly variable (Arnsten, 2000; Aston-Jones & Cohen, 2005; Gilzenrat et al., 2010; Sara & Bouret, 2012; Unsworth & Robison, 2016), especially in the presence of competing information. Outside of stimulus engagement, attentional focus is also associated with moderate tonic responses-- often indicative of an exploitative mode-- while attentional lapses that involve distraction to task-irrelevant external stimuli are associated with relatively high tonic responses (Unsworth & Robison, 2016)—indicative of a more exploratory mode.

As such, pupillometry is frequently used as a noninvasive proxy for cognitive effort while participants engage in demanding tasks (Eldar et al., 2013; Eldar et al., 2015; Gabay et al., 2011; Geva et al., 2013; Gilzenrat et al., 2010; Hopstaken et al., 2015; O’Bryan & Scolari, 2021; Rondeel et al., 2015; Smallwood et al., 2012; Van der Brink et al., 2016; Wendt et al., 2014). For example, attending to stimuli with low or diminishing subjective utility has been associated with both elevated baseline PD and a shift toward exploratory behavior (Eldar et al., 2013; Gilzenrat et al., 2010; Jepma & Nieuwenhuis, 2011). Similarly, relatively large baseline PD coupled with smaller task-evoked dilations predicts mental fatigue (Hopstaken et al., 2015), mind-wandering (Mittner et al., 2016; Smallwood et al., 2012), and behavioral lapses of attention (Unsworth &

Robison, 2016; Van der Brink et al., 2016). These findings are consistent with recent theories suggesting that tonic PD does not simply reflect task demands, and instead may provide an index of task-independent cognitive effort (da Silva Castanheria et al., 2020; van der Wel & van Steenbergen, 2018). Taken together, converging evidence from pupillometry studies suggests that exercising control (e.g., ignoring salient distracting stimuli in favor of less salient but behaviorally relevant ones) should drive pupillary responses, and such changes may track the performance decrements that characterize ego-depletion.

This increase in tonic PD may reflect increased activity in the locus coeruleus-norepinephrine (LC-NE) system. Single unit recording studies with nonhuman primates revealed a tight coupling between LC-NE system activity and PD (Costa & Rudebeck, 2016; Joshi et al., 2016; Rajkowski, Kubiak, & Aston-Jones, 1994), suggesting that the LC-NE may play a mechanistic role in adjusting pupil size during exploration-exploitation tradeoffs and while deploying attention more generally (Aston-Jones & Cohen, 2005; Sara & Bouret, 2012). Research with both humans and non-human primates has revealed that LC-NE activity increases in response to greater task demands, and in particular, during cognitive effort exertion (Alnæs et al., 2014; Aston-Jones & Cohen; 2005; Rajkowski et al., 1994; Sara & Bouret, 2012; Varazzani et al., 2015).

In light of these findings, we made the following hypotheses. First, we expected that engaging in a demanding attention task would elicit relative increases in PD. To test this prediction, we employed a paradigmatic attention control manipulation in which participants viewed a monotonous video containing salient visual distractors (Hagger et al., 2010; Schmeichel et al., 2003; Wagner & Heatherton, 2013). The design allowed for continuous eye-tracking, enabling us to extract precise measurements of PD, eye position, blinks, and saccades throughout

the manipulation. Participants assigned to the experimental (depletion) group were instructed to ignore the distracting information, while those in the control group viewed the display freely. Thus, we predicted that the depletion group would exhibit a larger task-associated PD over the course of the movie compared to the control group, even when controlling for eye movements and fixation patterns.

We next hypothesized that the sustained, elevated PD elicited during the attention control manipulation would persist into a subsequent cognitive control task. Thus, participants completed a Stroop task (Stroop, 1935) immediately following the attention video (Inzlicht & Gutsell, 2007). In contrast to the manipulation phase, instructions for the Stroop task were matched between groups.

In addition to being a commonly used measure in depletion studies (Inzlicht & Gutsell, 2007; Garrison et al., 2019; Lin et al., 2020), the Stroop task has been used to explore the link between phasic pupillary responses and attentional control. Incongruent trials, where the dimensions of each stimulus are associated with conflicting responses, elicit larger phasic PD relative to congruent trials (Laeng et al., 2011). This likely reflects the magnitude of the demands of each trial type, where a large phasic response on incongruent trials is indicative of successful response conflict resolution. However, when these expected phasic patterns are coupled with relatively large tonic responses, attentional control is expected to diminish, resulting in poorer performance in the face of conflicting input. Thus, we expected that an elevated tonic PD among the depletion group would lead to measurable lapses of control on the Stroop task.

In summary, exerting control should adaptively elicit increased PD among the depletion group during an initial task, but persistent elevated responses that carryover into a second task—reflected by high tonic PD—should maladaptively lead to poorer performance. To account for the

sequential nature of the tasks and correspondingly discrete PD sampling periods, we used a serial mediation model to quantify our results. Importantly, this approach also allows us to account for individual differences in the degree of carryover from one task to the next—a potentially critical element that, if unaccounted for, could lead to unreliable behavioral effects. Finally, given the demonstrated relationship between PD and LC activation, we consider our results in the context of the LC-NE system.

Methods

Participants. We performed a power analysis to estimate sample size prior to data collection. The effect size of task on pupil dilation—our primary measure of interest—is consistently medium to large (Eldar et al., 2013; Gilzenrat et al., 2010), with typical sample sizes for eye tracking studies ranging between 12 and 35 participants per condition. Estimates of depletion effect sizes, however, vary widely (from zero to large; Hagger et al., 2010; Hagger et al., 2016), and recent ego-depletion studies have recruited more than 200 participants per condition (Garrison et al., 2018; Lin et al., 2020). Sample sizes of this magnitude are not realistic in the context of eye tracking, however, where the experimenter continually monitors eye position in real time for each single participant. We therefore chose to use a medium effect size of Cohen’s $f = .30$. With an $\alpha = .05$ and power = .80, the projected sample size needed for the behavioral effect was 90. Data collection was completed prior to analysis and all materials were posted to www.osf.io/m6kax before data collection began.

Ninety-one participants (61 women, 18 men, 12 gender not reported; $M_{age} = 19.57$, $SD_{age} = 1.93$; 31.46% Hispanic or Latino, 62.92% Not Hispanic or Latino, 5.62% not reported; 70.79% White, 7.87% Black or African American, 3.37% Native American, 2.25% Asian, 1.12% Native Hawaiian or Other Pacific Islander, 4.49% more than once race, 10.11% unknown or not

reported) completed the study in exchange for partial course credit through an undergraduate psychology participant pool. Participants reported normal or corrected-to-normal vision and normal color vision as verified by an Ishihara color test (Ishihara, 1913; see Supplementary Material). Two participants were excluded from analyses due to glare on their glasses that prevented accurate eye tracking. All experimental procedures were approved by the presiding Institutional Review Board, and all participants provided written informed consent before beginning the experiment in compliance with the Declaration of Helsinki.

Materials and Procedures. Visual stimuli were rendered in MATLAB 2017b with Psychtoolbox3 (Kleiner et al., 2007) on a desktop PC using Windows 7. Stimuli from all tasks were displayed on a high-resolution (1920 x 1080 pixels) color monitor (BenQ XL2430T) measuring 58 cm wide, with a refresh rate of 100 Hz. Eye tracking measures, including pupil size and eye position, were recorded from the right eye at a sampling rate of 500 Hz using an EyeLink 1000 Plus infrared eye tracker (SR Research, Ontario, Canada), which was positioned in front of the monitor approximately 55 cm from the participant.

Our experimental procedures included a standard manipulation and measure of depletion effects (Hagger et al., 2010; Schmeichel et al., 2003) with minor modifications to accommodate continuous eye tracking (Fig. 1). Participants were randomly assigned to either the depletion (N = 44) or control (N = 45) condition and informed at the outset that they would complete two ostensibly unrelated research projects. Participants were told that the first (attention control task; see below) was designed to assess individual differences in reading body language, and the second (Stroop task; see below) would assess color perception. Unbeknownst to the subjects, however, the attention control task – which was always completed first – was used to manipulate self-control, and the Stroop task would be used to measure potential ego-depletion effects.

Participants' subjective experiences with the tasks were collected at the end to assess the effects of the manipulation and for exploratory analyses (see www.osf.io/m6kax for a complete list). As a manipulation check, participants were asked to report how effortful they found each task to be on a scale from 1 (not at all) to 7 (very much; see Supplementary Material for analyses of self-reported motivation). Participants were informed of the true purposes of the tasks upon debriefing.

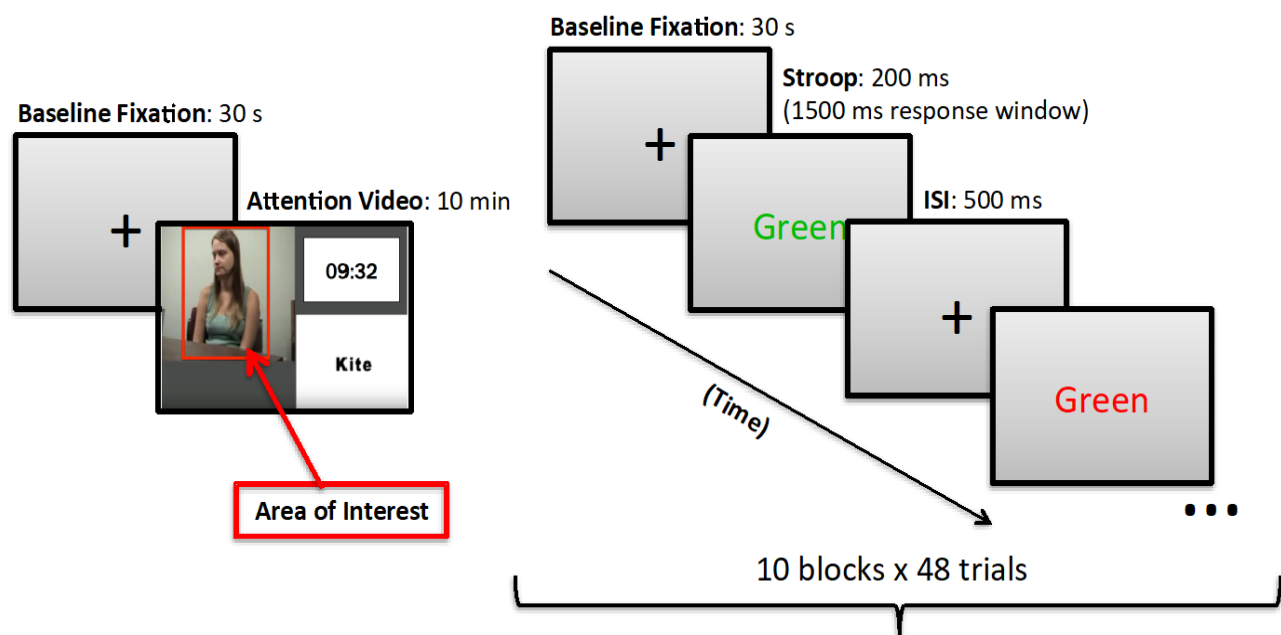


Fig. 1. Schematic illustration of the experimental design, depicting stimuli from both tasks. To control for luminance exposure between task conditions, PD analyses for the attention control video (left) were restricted to the area of interest indicated in red. A 30 s fixation period preceded the video and each block of the Stroop task (right).

Each participant was seated in a lighted room 92 cm from the stimulus monitor, and informed that their eye movements would be recorded via the eye tracker to determine where they were attending. An experimenter dressed in a white lab coat remained in the testing room with the participant and monitored eye position in real-time on a second monitor.

Attention Control Task. All participants watched a silent 10-minute video (available at osf.io/m6kax) of a seated female interviewee. The video was positioned left-of-center (5.3°), measured $10.9^\circ \times 19.7^\circ$ of visual angle, and had an average luminance of 41.7 cd/m^2 . In the lower-right portion of the screen (6.8° from center), a series of common, neutral, one-syllable words were displayed in black text on a white background (78.5 cd/m^2 ; measuring between $1.9^\circ \times 0.9^\circ$ and $4.4^\circ \times 0.9^\circ$; Schmeichel et al., 2003). A clock (86.7 cd/m^2) counting down to the end of the video was also displayed in the upper-right portion of the screen (positioned 6.8° from center and measuring $6.1^\circ \times 2.2^\circ$ (Price & Brewer, 2015). To maximize the likelihood that they would capture attention, the words and clock were made visually salient compared to the monotonous video via use of contrast and abrupt onsets.

Prior to starting the video, participants were randomly assigned to one of two conditions via task instructions from Schmeichel and colleagues (2003). Those in the depletion condition were explicitly instructed to maintain attention on the video while ignoring the salient words and clock. Participants in the control condition were not given any instructions about the words or clock, nor were they explicitly instructed where to attend. Both groups were told they would complete a short questionnaire following the video to probe their impressions of the interviewee, though no such questionnaire was ultimately administered. After task instructions and just prior to video onset, each participant fixated a black central fixation cross presented on a middle gray background (subtending $0.6^\circ \times 0.6^\circ$; luminance: 75.8 cd/m^2) for 30 s, which was later used to determine participants' initial baseline pupil diameter (see *Eye-tracking and Pupillometry* for data cleaning and normalization procedures).

Stroop Task. The attention control task was immediately followed by a Stroop task (Stroop, 1935): a frequently used and effective measure of control in the ego-depletion literature

(Gailliot et al., 2007; Garrison et al., 2019; Inzlicht & Gutsell, 2007; Job et al., 2010; Webb & Sheeran, 2003), as well as in pupillometry studies (Laeng et al., 2011), and thus lends itself well to the goals of this study. First, participants fixated a black cross (identical to the one encountered before the attention control video) for 30 s to acquire a measure of tonic PD. This long fixation period was repeated before each block. Each trial then began with the same fixation cross for 500 ms, which was then replaced by a 200 ms stimulus word (“*RED*” subtending $2.2^\circ \times 0.9^\circ$; “*GREEN*” subtending $3.4^\circ \times 0.9^\circ$) presented in either red (RGB: 255, 0, 0; luminance: 47.4 cd/m^2) or green (RGB: 0, 255, 0; luminance: 64.0 cd/m^2) font. Participants reported the font color as quickly as possible, and within a 1,500 ms response window, using the left and right arrow keys on the computer keyboard. On congruent trials (66.67% of all trials), the color word appeared in a font that matched its meaning (e.g., *RED* presented in red font); on incongruent trials (33.33% of all trials), the font color did not match the word’s meaning (e.g., *RED* presented in green font). Red and green fonts were evenly distributed across conditions. Participants completed 10, 48-trial blocks.

Eye Tracking and Pupillometry. Before both tasks, the eye tracker was calibrated for each participant in chinrest-free mode using a 13-point grid and validated with a $<1^\circ$ average cartesian prediction error. Offline, pupil diameter (collected via default EyeLink settings, in arbitrary units) was extracted in 500 ms bins for 1) the length of the 30 s fixations prior to the attention control task and interleaved with the Stroop task blocks, and 2) the length of the 10-minute video. These bins were then averaged together for the 30 s fixation periods to estimate PD for each subject after removing whole bins that contained blinks, saccades, or abnormally small average PD indicative of blinks and/or eye movements (below the 5th percentile; Mathôt et al., 2018). When this measurement is taken during task engagement, we refer to it as task-associated PD, as it should

reflect cognitive effort driven by ongoing attention to visual stimuli. This is the case during the video task. When taken during fixation periods either outside of or between periods of task engagement, we refer to it as tonic PD, as it should reflect a baseline response.

We used a standard subtraction approach to normalize participant's task-associated PD measurements for the video phase (Mathôt et al., 2018). Specifically, each participant's raw PD averaged over the last 25 s of the initial fixation period (baseline) was subtracted from the values recorded during the video. Thus, analyses and figures depicting PD reflect unit increases and/or decreases for each participant relative to their initial baseline. The first 5 s of fixation were omitted from the baseline computations to ensure participants' pupils had adjusted to the luminance of the monitor (notably, the experiment took place in a lit room so participants did not experience drastic changes in lighting conditions between entering the lab and initiating eye tracking).

During the attention control task, participants from each condition were presented with the same display but given different instructions about making eye movements. Because we were interested in relative differences in task-associated PD between conditions, and such responses can be influenced by changes in eye position across a display of unequal luminance (Brisson et al., 2013), we restricted all PD analyses to periods of stable fixation on the interviewee (5.3° from center, subtending $6.2^\circ \times 8.7^\circ$; luminance: 41.7 cd/m^2 ; see Fig. 1) by omitting all 500 ms sampling bins containing saccades or eyeblinks. This was applied uniformly to all participants, thus matching eye position and the luminance of the overtly attended region between groups. To verify that this procedure appropriately eliminated any undue influences from preceding eye movements across the variable display, we correlated each participant's fixation count to the high-contrast distractors, and the proportion of time spent maintaining fixation within the interest

area, with our restricted PD. Absent correlations would indicate that eye movements had no discernable impact on the PD estimates used in our main analyses.

An analogous normalization procedure was performed on the tonic PD data for the Stroop task, where baseline PD was calculated from the 30 s fixation period encountered just after the attention control task and before the first Stroop block. Three participants (two from the control group and one from the depletion group) failed to maintain fixation during this period and were thus omitted from PD analyses performed on fixation data during the Stroop task.

In addition to tonic PD during the Stroop task, we measured the phasic responses for congruent and incongruent trials separately, allowing us to estimate relative cognitive demands for each trial type. To calculate phasic PD, we first extracted the data in 100 ms bins. For each trial, we defined the baseline response as the average of the five 100 ms bins that contained the trial-level fixation period preceding stimulus presentation. To account for the delay between measurable changes in pupil size and LC-NE activation (Aston-Jones et al., 1994; Aston-Jones & Cohen, 2005), we defined the stimulus response for each trial as the bin containing the maximum pupil size following stimulus offset and within the response window. This is also consistent with the epoch within which Laeng et al. (2011) observed reliable Stroop effects in PD. Finally, we normalized each stimulus-driven response by subtracting from it the trial-level baseline.

Statistical Analyses. Statistical analyses were carried out using R (<https://www.r-project.org/>). For behavioral analyses involving nested effects, we computed linear mixed models using the “lme4” package in R with restricted maximum likelihood estimation and random intercepts for subjects. Mean proportion correct for each Stroop trial type (congruent and incongruent) served as the primary outcome variable for behavioral analyses, with the effect of task manipulation on performance assessed as the interaction between trial type and self-control

condition (depletion and control). The same analyses were secondarily performed on reaction time (RT) data from correct trials. Reported effect sizes for simple main effects reflect Cohen's d ($\frac{\mu_1 - \mu_2}{\sigma_{pooled}}$). Interaction effect sizes in linear mixed models were computed using a near approximation of Cohen's d , $d_{residual}(\frac{(a_1 - b_1) - (a_2 - b_2)}{2\sigma})$, where subscripted letters represent cell means and σ represents the residual standard deviation of a given mixed model.

We used a serial multiple mediation model (Hayes, 2013) to test whether depletion effects were driven by the task-associated PD response more so than the direct effect of the experimental manipulation. The serial mediation straightforwardly extends standard single-mediator models by assuming that correlated variables sampled at different time points might exert unique effects on a dependent variable. Thus, it differs from standard multiple mediation models by assuming a causal relationship between mediators. We predicted that elevated PD during the video (M_1) would directly influence tonic PD during the subsequent Stroop task (M_2), and that tonic PD during the Stroop task would ultimately predict the magnitude of depletion effects in behavior (Y). Because ego-depletion should specifically impair cognitive control, we used incongruent trial accuracy as the behavioral outcome variable.

Three models corresponding to each outcome variable (M_1 , M_2 , and Y) were used to estimate the indirect effects of depletion condition (X) through M_1 and M_2 respectively, the serial indirect effect of condition through M_1 and M_2 , and the direct effect of condition while holding all indirect effects and interactions constant. For the serial mediation results, parameter $a_1 * b_1$ represents the indirect effect of condition on behavior through M_1 (manipulation PD) alone; $a_2 * b_2$ represents the indirect effect of condition on behavior through M_2 (Stroop PD) alone; $a_1 * d_{21} * b_2$ represents the indirect effect of condition on behavior through M_1 and M_2 in series; and c'

represents the direct effect of condition when adjusting for all indirect effects. All models were computed at the mean subject level using simple linear regression.

Finally, we created bootstrapped 95% confidence intervals by randomly sampling subjects with replacement up to the original sample size (10,000 iterations) to relax assumptions of normality with respect to the sampling distributions of the indirect effects. Originally represented in arbitrary units corresponding to PD, both mediators were standardized to facilitate interpretation. Values for the behavioral outcome Y (mean proportion correct on incongruent Stroop trials) were kept in their original, meaningful units.

Results

Manipulation checks. Before testing our main pupillometry predictions, we first ensured that (1) our attention control manipulation effectively impacted self-reports of exerted effort; (2) participants adhered to task instructions regarding eye movements; and (3) external stimulus properties do not account for pupil size during the attention control video.

Self-reports of effort. Given the depletion condition was designed to be more cognitively demanding than the control condition, we expected the depletion group would report relatively higher levels of subjective effort during the attention control task. We calculated a composite “effort” score from participants’ responses to two highly correlated ($r = 0.60$, 95% CI [0.44, 0.72], $p < 0.001$), but differently phrased questions. As expected, participants in the depletion condition reported exerting more effort ($M = 4.0$, $SD = 1.6$) than those assigned to the control condition ($M = 3.1$, $SD = 1.3$; $t(85) = 2.97$, $p = 0.004$, $d = 0.62$).

Adherence to task instructions. During the attention control video, individuals assigned to the depletion condition maintained stable fixation within the interest area at a significantly higher rate than those in the control condition (depletion: $M = 63.17\%$, $SD = 14.13\%$; control: $M =$

50.45%, $SD = 18.02\%$; $t(87) = 3.70, p < 0.001, d = 0.79$). Depletion participants also made significantly fewer fixations to the distractors than control participants (depletion: $M = 15.02, SD = 19.64$; control: $M = 138.58, SD = 101.19$; $t(87) = 7.95, p < 0.001, d = 2.05$). Importantly, eye movement patterns only differed between groups during the attention control video. For the subsequent Stroop task where all participants received identical instructions, percent fixation was well matched between groups during the interleaved fixation periods (depletion: $M = 51.97\%, SD = 21.94\%$; control: $M = 49.81\%, SD = 18.43\%$; $t(84) = 0.45, p = 0.65, d = 0.10$). Thus, participants in the depletion group appropriately adhered to task instructions.

Eye movements and pupil size. The distracting stimuli were made difficult to ignore by presenting them at higher contrast than the video (see *Methods: Attention Control Task*). Importantly, all participants were exposed to the same stimulus display. However, eye movements from the low-contrast video to the high-contrast distractors could lead to constricted pupils, such that more fixations to the distractors may produce overall smaller mean PD. We therefore restricted our measure of PD to periods of stable fixation on the video (see *Methods: Eye Tracking and Pupillometry*) to ensure that any group differences can be attributed to attentional effort rather than stimulus properties. Because the groups exhibited different eye movement patterns, we conducted an additional control analysis to check whether this accounted for our PD measure during the attention control task. Subject-level correlations between fixation count to the distractors and PD during stable fixation at the video produced no evidence of a relationship in either the depletion ($r = 0.04, 95\% \text{ CI } [-0.26, 0.34], t(42) = 0.28, p = 0.78$) or the control condition ($r = -0.12, 95\% \text{ CI } [-0.40, 0.18], t(43) = -0.77, p = 0.45$). Similarly, the proportion of time spent maintaining fixation within the interest area was not associated with our PD measure for either group (depletion group: $r = .02, 95\% \text{ CI } [-.28, .31], t(42) = 0.11, p = .92$; control group:

$r = -.15$, 95% CI $[-.43, .15]$, $t(43) = -0.98$, $p = .33$; see Supplementary Fig. 1). Thus, the reported PD was not appreciably driven by luminance differences across the video display for either participant group.

Finally, raw pupil size was well matched between participant groups during baseline (see *Methods: Eye Tracking and Pupillometry*), before the onset of the attentional control task (depletion: $M = 38.41$ units, $SD = 9.24$ units; control: $M = 40.28$ units, $SD = 9.84$ units; $t(87) = 0.93$, $p = 0.36$, $d = 0.20$). In sum, any subsequent differences in task-associated PD can be appropriately attributed to attention demands.

Behavioral results. A sizeable number of studies failed to replicate the depletion effect when only considering participants' mean performance (e.g., Hagger et al., 2016; Inzlicht et al., 2014; Vohs et al., 2021). Consistent with these demonstrations, we failed to detect a significant effect when ignoring the pupillary response. We observed a classic Stroop interference effect in both accuracy, $F(1, 87) = 18.4$, $p < 0.001$, $d = 0.58$, and RT, $F(1, 87) = 81.5$, $p < 0.001$, $d = 0.52$, whereby performance was reliably better on congruent trials (see Supplementary Fig. 2; see Supplementary Material for a speed-accuracy tradeoff analysis). Consistent with the expected direction of an ego-depletion effect, the depletion group performed worse on incongruent trials ($M = 83.9\%$, $SD = 29.3\%$) compared to the control group ($M = 88.03\%$, $SD = 19.9\%$), but this pattern was not significant, $F(1, 87) = 0.91$, $p = 0.34$, $d_r = 0.14$; nor was it observed in RT, $F(1, 87) = 0.13$, $p = 0.72$, $d_r = 0.05$. These results align with our hypothesis that, without accounting for individual differences in carryover effects from the initial task to the subsequent task, behavioral ego-depletion effects may be unreliable. With that in mind, we turn next to the pupillometry data.

339

340 ***Pupillometry results.*** Behavioral performance measures fail to capture individual
341 differences in the degree and duration of exerted cognitive effort during a control task, an
342 important expected precursor to ego-depletion effects. Given that pupil size can serve as a
343 physiological index of ongoing cognitive effort, we next examined (1) how self-control demands
344 modulated PD within task, (2) whether this response persisted into a subsequent task, and (3)
345 whether this pupillary response persistence predicts self-control failure.

346 *Self-control demands in Task 1 modulate PD during Task 1.* We hypothesized that greater
347 cognitive demands for the depletion group would lead to increased PD during the video
348 manipulation. A two-sample t-test confirmed this prediction: participants in the depletion group
349 had larger mean PD relative to baseline ($M = +0.14$ units, $SD = 0.68$ units) compared to the
350 control group ($M = -0.20$ units, $SD = 0.81$ units) for the duration of the attention control video, t
351 $(87) = 2.16$, $p = 0.03$, $d = 0.46$. This effect emerged rapidly after video onset and was sustained
352 throughout the 10-minute run time (see Fig. 3).

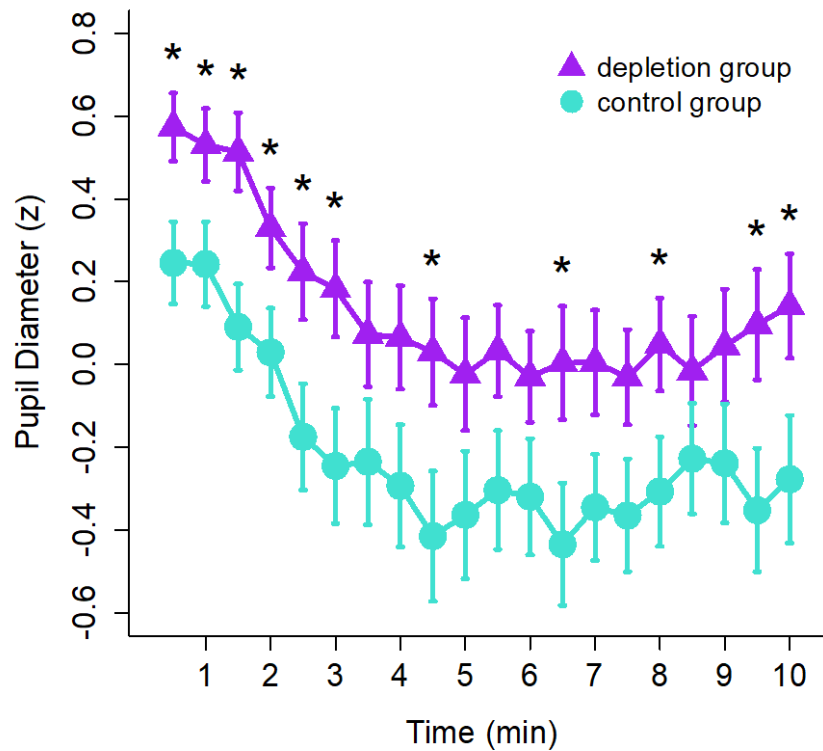


Fig. 2. Task-associated pupil diameter (PD) during the attention control task. Markers (purple triangles: depletion group, $n = 44$; turquoise circles: control group, $n = 45$) depict z-scored change in PD relative to participants' initial baseline PD over a 30 s recording window. Individual time points with significant group differences at $p < 0.05$ (uncorrected) as indicated by pairwise regressions are indicated with asterisks. Error bars depict between-subject SEM.

Task 1 PD responses persist as tonic PD during Task 2. Next, we predicted the sustained PD elicited during the initial task would positively correlate with tonic PD responses in the subsequent task. A Pearson correlation revealed a strong positive association between changes in subject-level PD during the attention control video and tonic PD during the Stroop task, $r = 0.48$, 95% CI [0.29, 0.63], $t(84) = 4.97$, $p < 0.001$, the magnitude of which did not differ between groups (depletion: $r = 0.48$, 95% CI [0.21, 0.68], $t(41) = 3.48$, $p = 0.001$; control: $r = 0.50$, 95% CI [0.23, 0.69], $t(41) = 3.67$, $p < 0.001$), $F(1, 82) = 0.09$, $p = 0.76$.

Taken together, the pupillometry results thus far suggest that the additional demands of the depletion condition led to increased PD during the video task, and that participants' PD

responses to the video task largely persisted into the subsequent Stroop task. Interestingly, the same relationship was not observed for phasic PD time locked to the Stroop stimuli for either group (depletion: $r = -0.04$, 95% CI [-0.34, 0.26], $t(42) = -0.29$, $p = 0.77$; control: $r = 0.21$, 95% CI [-0.09, 0.47], $t(43) = 1.41$, $p = 0.17$). Thus, the persistent carryover effect from the video task was restricted to the tonic response during the Stroop task, as we predicted.

Tonic PD during Task 2 predicts Task 2 performance. We predicted that a sufficiently strong and sustained LC-NE response carried over from a demanding task should impact performance on a subsequent task. A crucial component for testing this theory using pupillometry is that tonic PD during the Stroop task must track task performance (Gilzenrat et al., 2010; Unsworth & Robison, 2016; Van der Brink et al., 2016). First, incongruent trials should be relatively more cognitively demanding given the presence of salient, conflicting information (i.e., word meaning), indicated by larger phasic PD time locked to stimulus presentation (Laeng et al., 2011). As expected, the phasic response across groups was larger on incongruent trials compared to congruent trials, $t(88) = 7.57$, $p < .001$, $d = .31$. This result, coupled with the significant behavioral Stroop effect, indicates that a relatively larger, persistent tonic response should be especially detrimental on the more demanding incongruent trials, especially as it relates to an (here, maladaptive) exploratory mode.

We therefore employed a multiple regression to further examine how depletion condition, tonic PD, and their interactions related to mean accuracy on incongruent Stroop trials. The model revealed a significant main effect of tonic PD, $\beta = -0.15$, $t(78) = -2.79$, $p = 0.01$, indicating that larger mean response during the interleaved fixation periods was associated with lower accuracy on incongruent trials across groups. Furthermore, the negative effect of tonic PD on accuracy was significantly more pronounced for the depletion group, $\beta = 0.18$, $F(1, 78) = 5.21$, $p = 0.03$ (see

Supplementary Material for a similar pattern in phasic PD). To illustrate this continuous interaction effect, we divided participants within each group according to whether they exhibited high (above median) or low (below median) accuracy on the Stroop task (see Fig. 4).

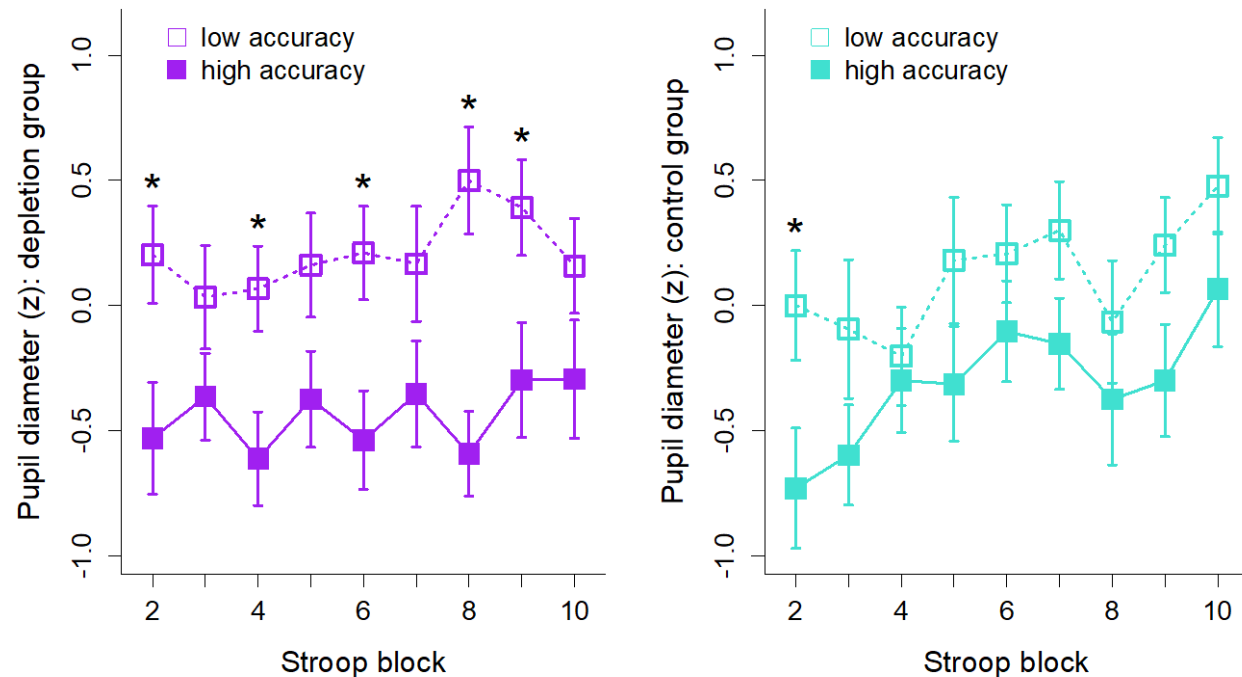


Fig. 3. Tonic pupil diameter (PD) during the Stroop task for both groups as a function of task performance. For illustrative purposes, “high accuracy” (solid markers) corresponds to above-median Stroop performers on incongruent trials within each respective group, and “low accuracy” (open markers) corresponds to below-median performers. Markers depict z-scored change in PD relative to participants’ pre-Stroop baseline PD over a 30 s recording window. The first Stroop fixation period was used for normalization and is thus omitted. Individual time points with significant group differences at $p < 0.05$ (uncorrected) as revealed by pairwise regressions are indicated with asterisks. The interaction between attention control condition and PD in predicting accuracy was significant ($p = 0.03$). Error bars depict between-subject SEM.

Task 1 PD mediates ego-depletion effects in Task 2. Finally, we used a serial multiple mediation framework to consider all effects described above in a single model (Hayes, 2013). Here, elevated PD during the initial task should mediate the relationship between task demands and self-control failures in the sequential-task paradigm. Specifically, we expected greater demands during the manipulation to be associated with temporally sustained increases in PD, that

these initial modulations of PD would be predictive of tonic PD in the Stroop task that followed, and finally, that larger tonic PD during the Stroop task would reliably predict poorer performance on incongruent Stroop trials. Rather than assuming independence among mediators, serial mediations are designed to test for indirect effects of a manipulation through two or more mediators that are thought to be causally related (here, task-associated PD during the attention control task and tonic PD during the Stroop task). This analysis is therefore ideally suited for investigating potential physiological mediators in the sequential-task paradigm used here and more broadly in ego-depletion research. In particular, the serial mediation can account for subject-level variability in the carryover of indirect effects: physiological effects that are measured only during a manipulation might not uniformly persist into a subsequent self-control task across participants, and a physiological effect only measured during a later self-control task could be epiphenomenal with respect to a temporally separated manipulation phase. Accordingly, the serial indirect effect $a_1 * d_{21} * b_2$ served as the primary outcome of interest in our mediation model (see Methods, *Statistical Analysis* for more details). Critically, the presence or absence of a significant total effect (path c) does not bear on the significance nor interpretation of indirect pathways in mediation analysis (Hayes, 2013; Rucker, Preacher, Tormala, & Petty, 2011).

The results of the mediation analysis are depicted in Fig. 4 and Supplementary Table 1. As predicted, Stroop performance was serially mediated by task-associated PD on the initial task. Being assigned to the depletion condition was associated with increased PD during the attention video compared to the control condition (a_1 : $\beta = 0.45$, 95% CI [0.05, 0.87]). This in turn predicted tonic PD in the Stroop task (d_{21} : $\beta = 0.40$, 95% CI [0.16, 0.58]), which itself was associated with poorer accuracy on incongruent trials (b_2 : $\beta = -0.15$, 95% CI [-0.29, -0.01]).

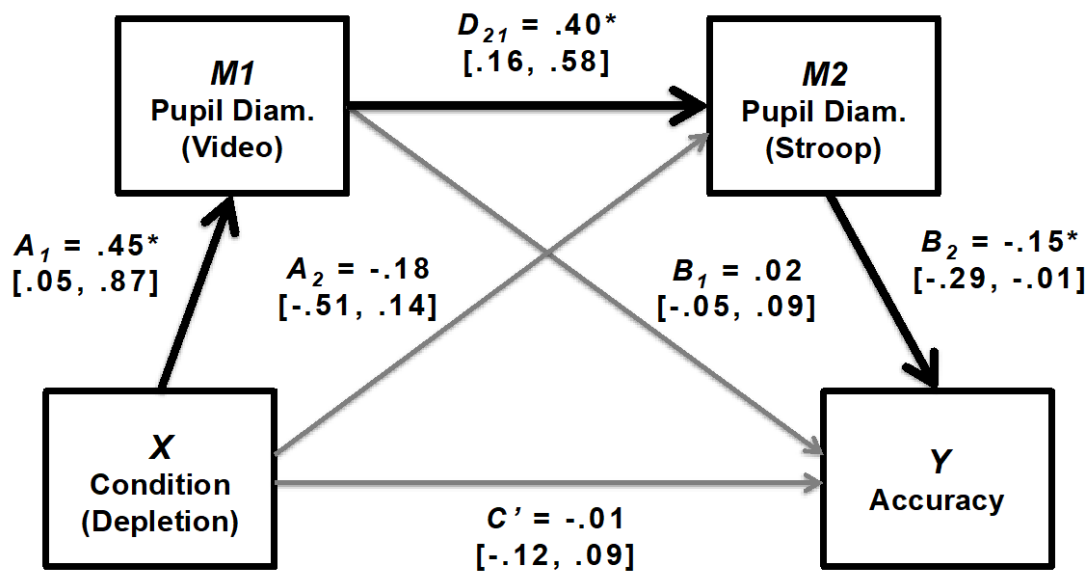


Fig. 4. Serial mediation results. Boxes correspond to the variables used in the serial mediation. Arrows correspond to each path tested in the model, with black arrows reflecting significant paths and gray arrows reflecting non-significant paths. The parameter estimates for each path correspond to the median of 10,000 bootstrapped estimates, and the bracketed values below them correspond to the 2.5th and 97.5th percentiles, respectively. Asterisks indicate that the 95% bootstrapped confidence intervals for the path did not cross zero.

These results highlight the importance of accounting for temporal carryover of physiological responses in sequential task paradigms. Considering only the effect of depletion condition on Stroop tonic PD (a_2 : $\beta = -0.18$, 95% CI [-0.51, 0.14]) or the effect of task-associated PD during the video task on subsequent Stroop task accuracy (b_1 : $\beta = 0.02$, 95% CI [-0.05, 0.09]) fails to capture individual variability in the correlation between task-associated and tonic PD across the two tasks. In contrast, the serial pathway accounting for such variability ($a_1*d_{21}*b_2$: $\beta = -0.02$, 95% CI [-0.09, -0.001]) revealed a clear indirect effect.

Discussion

By measuring fluctuations in PD during a typical ego-depletion paradigm, we provide evidence that pupillometry can track lapses of self-control. We demonstrated via serial mediation that greater self-control demands led to larger task-associated PD during the attention control

video, which in turn predicted tonic PD during the Stroop task. Larger tonic PD during the Stroop task predicted poorer performance, and this effect was more pronounced among participants who were instructed to control their attention during the manipulation. Accordingly, exerting self-control impairs subsequent performance indirectly. Given the demonstrably strong coupling between changes in pupil size and LC-NE activation (Aston-Jones & Cohen, 2005; Aston-Jones et al., 1994; Murphy et al., 2014), we argue this may occur through the modulation of LC-NE activity.

This study lends support to models suggesting that reductions in task-oriented attention plays a role in ego-depletion effects (Inzlicht & Schmeichel, 2012; Inzlicht et al., 2014). According to the process model, depletion is thought to be a byproduct of an attentional shift: after sustained focus (exploitation) on a demanding task, participants engage an exploratory mode in search of more immediate rewards. This weakens their attentional capacity for conflict monitoring and increases the likelihood of self-control failure (Inzlicht & Schmeichel, 2012; Inzlicht et al., 2014). The link between exploitation-exploration tradeoffs and LC-NE is well established in the reinforcement learning literature (Aston-Jones & Cohen, 2005; Eldar et al., 2013; 2016; Gilzenrat et al., 2010; Jepma & Nieuwenhuis, 2011), and we provide evidence that a similar neurobiological mechanism, as tracked by pupil size, underlies ego-depletion effects.

We showed that lapses in self-control follow from an elevated PD that emerges early and is sustained over the course of a long, challenging task. Although the modulatory effects of our manipulation on inhibitory control in the subsequent Stroop task were apparent when modeling the indirect effects of PD, the direct effect of condition on task performance was ambiguous. As such, behavioral depletion effects may only be observed when a preceding task elicits requisite changes in underlying cognitive and neurobiological states. Likewise, accounting for the

mediating effects via PD may reveal significant group differences related to self-control manipulations that would not be apparent from traditional behavioral measures alone.

Because group differences in task-associated PD emerged rapidly during the manipulation, it remains unclear whether lapses in self-control would coincide with the onset of an elevated response or would only emerge following a prolonged response in connection to a task. For example, previous depletion findings have shown just perceiving a task as draining or even imagining exerting self-control can result in depletion effects (Ackerman et al., 2009; Clarkson et al., 2010). Embarking on or imagining a subjectively draining task may induce rapid increases in cognitive arousal, mirrored by increases in NE (Arnsten, 2000; Hermans et al., 2011; Sara & Bouret, 2012). Given that LC-NE activity is regularly tied to cognitive processing of external stimuli (Einhauser et al., 2008; Eldar et al., 2013; Gabay et al., 2013; Gilzenrat et al., 2010; Van der Brink et al., 2016), it is alternatively possible that modulation of NE in self-control is specific to manipulations that require effortful response suppression to task-irrelevant input (Baumeister et al., 1998; Hagger et al., 2010; Schmeichel et al., 2003, Wagner & Heatherton, 2013), as was true of the current study.

Although the precise mechanisms linking LC-NE activity and PD are not fully understood, converging electrophysiological evidence suggests that both spontaneous and temporally sustained fluctuations in PD precisely track NE release (Aston-Jones & Cohen, 2005; Aston-Jones et al., 1994; Costa & Rudebeck, 2016; Joshi et al., 2016; Mittner et al., 2016; Rajkowski et al., 1994; Reimer et al., 2016), allowing researchers to reliably infer LC-NE activity from PD in behaving human participants (Einhäuser et al., 2008; Eldar et al., 2013; Eldar et al., 2016; Gabay et al., 2013; Gilzenrat et al., 2013; Hauser et al., 2019; Mittner et al., 2016; O'Bryan & Scolari, 2021). Critically, our study design and analytical approach enabled such inferences: We restricted

all PD analyses to periods of stable fixation on portions of the screen that were matched for luminance and gaze offset between groups, thus carefully controlling for aspects of the visual display that could differentially affect PD measurements (Brisson et al., 2013; Mathôt et al., 2018).

However, PD is not exclusively driven by LC-NE activity; dopamine and acetylcholine have also been shown to correlate with pupil size in animal models (Hauser et al., 2019; Reimer et al., 2016; Varrazzani et al., 2015). Nonetheless, even more human behavioral studies have linked PD changes to fluctuations in attentional or cognitive control (Einhäuser et al., 2008; Eldar et al., 2013; Eldar et al., 2016; Gabay et al., 2011; Geva et al., 2013; Gilzenrat et al., 2013; Hopstaken et al., 2012; Laeng et al., 2011; Rondeel et al., 2015; Wendt et al., 2014), regardless of the underlying mechanism(s). Given the existing literature that demonstrates (1) similar relationships between exerted control and LC-NE activation as well as PD and (2) a tight temporal coupling between LC-NE and PD, we believe that modulation of LC-NE is the most likely explanation for our pattern of results. Importantly, however, we acknowledge that other neurobiological mechanisms may contribute.

Narrowing in on NE as a likely mechanism for our PD effects allows us to generate plausible and testable explanations for ego-depletion from the broader LC-NE literature. For example, performing challenging and/or monotonous self-control tasks may produce a mild stress response in some participants (Arnsten, 2000; 2009), which could simultaneously account for both an elevated LC-NE response suggested by our study and performance decrements on tasks requiring cognitive control more broadly observed across the literature. Stress-induced NE release has been mechanistically linked to poor performance on tasks requiring executive function via inhibition of prefrontal cortex (PFC; Arnsten, 2009; Sara & Bouret, 2012). The LC shares

extensive connections with the neocortex, and NE levels in the PFC are essential for tasks that require focused attention, working memory, or cognitive flexibility (Arnsten, 2000; Arnsten, 2009; Arnsten et al., 2012; Ramos & Arnsten, 2007; Sara & Bouret, 2012), all of which are called upon during self-control tasks. Crucially, when NE release exceeds a critical threshold, specific receptors in PFC become relatively inhibited resulting in reduced executive function and an increased tendency to behave in automatic or conditioned ways (Arnsten, 2000; 2009).

The downstream effects of stress on executive function via LC-NE modulation may fittingly characterize the physiological mechanisms underlying ego-depletion, an explanation that is also compatible with the process model described above (Inzlicht et al., 2012; 2014). An acute stress response elicited from a challenging self-control task may predict *when* (and whether) participants will shift from active task exploitation to a more disengaged, exploratory state in pursuit of immediately rewarding stimulation. This tentative explanation is consistent with well-established research on the effects of acute stress on PFC function and executive function (Arnsten, 2000; Arnsten, 2009; Arnsten, Wang, & Paspalas, 2012; Ramos & Arnsten, 2007; Hermans et al., 2011; Sara & Bouret, 2012), and may help inform future research via testable hypotheses about the conditions that will lead to self-control failure.

The original ego-depletion finding has been spectacularly generative, while also yielding pressing questions regarding the equivocal occurrence and mechanism of the effect. By identifying tonic PD as a physiological marker of depletion, which fluctuates with cognitive effort and predicts performance, we offer critical evidence toward reconciling the debate regarding the validity of the effect (Dang, 2016; Hagger et al., 2016). Although we cannot draw causal conclusions from a physiological mediator that was not directly manipulated, this study identifies

544 a candidate mechanism for depletion effects. A sustained LC-NE response may be a reliable
545 precursor to behavioral lapses of self-control.

546

547

548

Statements and Declarations

549

Conflict of Interest Statement

551

552 On behalf of all authors, the corresponding author states there is no conflict of interest.

553

554

Open Practices Statement

556

557 Data, experimental code, and analysis code used to complete the study will be available upon
558 publication at the Open Science Framework: www.osf.io/m6kax

559

560

Funding

562

563 This research did not receive any specific grant from funding agencies in the public, commercial,
564 or not-for-profit sectors.

References

- Ackerman, J. M., Goldstein, N. J., Shapiro, J. R., & Bargh, J. A. (2009). You Wear Me Out: The Vicarious Depletion of Self-Control. *Psychological Science*, 20, 326 – 332.
<https://doi.org/10.1111/j.1467-9280.2009.02290.x>
- Alnæs, D., Sneve, M.H., Espeseth, T., Endestad, T., van de Pavert, S.H., & Laeng, B. (2014). Pupil size signals mental effort deployed during multiple object tracking and predicts brain activity in the dorsal attention network and the locus coeruleus. *Journal of Vision*, 14, 1–20.
- Arnsten, A. F. (2000). Stress impairs prefrontal cortical function in rats and monkeys: Role of dopamine D1 and norepinephrine alpha-1 receptor mechanisms. *Progress in Brain Research*, 126, 183–192. doi: 10.1016/S0079-6123(00)26014-7
- Arnsten, A.F. (2009). Stress signalling pathways that impair prefrontal cortex structure and function. *Nat. Rev. Neurosci.* 10, 410–422.
- Arnsten, A.F.T., Wang, M.J., & Paspalas, C.D. (2012). Neuromodulation of thought: flexibilities and vulnerabilities in prefrontal cortical network synapses. *Neuron*, 76, 223–239.
- Aston-Jones, G., & Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: Adaptive gain and optimal performance. *Annual Review of Neuroscience*, 28, 403–450. doi: 10.1146/annurev.neuro.28.061604.135709
- Aston-Jones, G., Rajkowski, J., Kubiak, P., & Alexinsky, T. (1994). Locus coeruleus neurons in monkey are selectively activated by attended cues in a vigilance task. *J Neurosci* 14, 4467–4480.

- Baumeister, R. F., Bratslavsky, E., Muraven, M., & Tice, D. M. (1998). Ego-depletion: Is the active self a limited resource? *Journal of Personality and Social Psychology*, 74, 1252–1265. doi: 10.1037/0022-3514.74.5.1252
- Beedie, C. J., & Lane, A. M. (2012). The role of glucose in self-control: Another look at the evidence and an alternative conceptualization. *Personality and Social Psychology Review*, 16, 143–153.
- Brisson, J., Mainville, M., Mailloux, D., Beaulieu, C., Serres, J., & Sirois, S. (2013). Pupil diameter measurement errors as a function of gaze direction in corneal reflection eyetrackers. *Behavior Research Methods*, 45, 1322–1331. doi: 10.3758/s13428-013-0327-0
- Carter, E.C., Kofler, L.M., Forster, D.E., & McCollough, M.E. (2015). A series of meta-analytic tests of the depletion effect: Self-control does not seem to rely on a limited resource. *Journal of Experimental Psychology: General*, 144, 796 – 815.
- Clarkson, J., Hirt, E., Jia, L., & Alexander, M. (2010). When perception is more than reality: The effects of perceived versus actual resource depletion on self-regulatory behavior. *Journal of Personality and Social Psychology*, 98, 29–46.
- Costa, V.D. & Rudebeck, P.H. (2016). More than meets the eye: the relationship between pupil size and locus coeruleus activity. *Neuron* 89, 1–10.
- Dang, J. (2016). Commentary: A multilab preregistered replication of the ego-depletion effect. *Frontiers in Psychology*, 7. doi: 10.3389/fpsyg.2016.01155
- da Silva Castanheria, K, LoParco, M., & Otto, R.A. (2020). Task-evoked pupillary responses track effort exertion: Evidence from task-switching. *Cognitive, Affective, & Behavioral Neuroscience*. doi: 10.3758/s13415-020-00843-z

- Einhäuser, W., Stout, J., Koch, C., & Carter, O. (2008). Pupil dilation reflects perceptual selection and predicts subsequent stability in perceptual rivalry. *Proc Nat Acad Sci USA* 105, 1704–1709.
- Eldar, E., Cohen, J. D., & Niv, Y. (2013). The effects of neural gain on attention and learning. *Nature Neuroscience*, 16, 1146–1153. doi: 10.1038/nn.3428
- Eldar, E., Niv, Y., & Cohen, J. (2016). Do you see the forest or the tree? Neural gain and breadth versus focus in perceptual processing. *Psychol Sci* 27, 1632–1643.
- Evans, D. R., Boggero, I. A., & Segerstrom, S. C. (2016). The Nature of Self-Regulatory Fatigue and “Ego-depletion”: Lessons From Physical Fatigue. *Personality and Social Psychology Review*, 20, 291–310. <https://doi.org/10.1177/1088868315597841>
- Finley, A. J., Tang, D., & Schmeichel, B. (2019). Sweet nothings: No effect of self-control exertion on blood glucose levels. *PsyArXiv*. doi: 10.1027/1864-9335/a000376
- Friese, M., Loschelder, D.D., Gieseler, K., Frankenbach, J., & Inzlicht, M. (2019). Is ego-depletion real? An analysis of arguments. *Personality and Social Psychology Review*, 23, 107-131.
- Fujita, K. (2011). On conceptualizing self-control as more than the effortful inhibition of impulses. *Personality and Social Psychology Review*, 15, 352–366. doi: 10.1177/1088868311411165
- Gabay, S., Pertzov, Y., & Henik, A. (2013). Orienting of attention, pupil size, and the norepinephrine system. *Atten Percept Psychophys* 73, 123–129.
- Gailliot, M. T., Baumeister, R. F., DeWall, C. N., Maner, J. K., Plant, E. A., Tice, D. M., ... Schmeichel, B. J. (2007). Self-control relies on glucose as a limited energy source:

- Willpower is more than a metaphor. *Journal of Personality and Social Psychology*, 92, 325–336. doi: 10.1037/0022-3514.92.2.325
- Gilzenrat, M. S., Nieuwenhuis, S., Jepma, M., & Cohen, J. D. (2010). Pupil diameter tracks changes in control state predicted by the adaptive gain theory of locus coeruleus function. *Cognitive, Affective, & Behavioral Neuroscience*, 10, 252–269. doi: 10.3758/CABN.10.2.252
- Hagger, M. S., Chatzisarantis, N. L. D., Alberts, H., Anggono, C. O., Batailler, C., Birt, A. R., ... Zwieneberg, M. (2016). A multilab preregistered replication of the ego-depletion effect. *Perspectives on Psychological Science*, 11, 546–573. doi: 10.1177/1745691616652873
- Hagger, M. S., Wood, C., Stiff, C., & Chatzisarantis, N. L. D. (2010). Ego-depletion and the strength model of self-control: A meta-analysis. *Psychological Bulletin*, 136, 495–525. doi: 10.1037/a0019486
- Hauser, T.U., Eldar, E., Purg, N., Moutoussis, M., & Dolan, R.J. (2019). Distinct roles of dopamine and noradrenaline in incidental memory. *J Neurosci* **39**, 7715–7721.
- Hayes, A. F. (2013). *Introduction to mediation, moderation, and conditional process analysis: A regression-based approach*. New York: The Guilford Press.
- Heatherton, T.F., Wagner, D.D. (2011). Cognitive neuroscience of self-regulation failure. *Trends in Cognitive Sciences*, 15, 132–9.
- Hermans, E.J. et al. (2011). Stress-related noradrenic activity prompts large-scale neural network reconfiguration. *Science*, 334, 1151–1153.
- Hopstaken, J.F., van der Linden, D., Bakker, A.B., & Kompier, M.A. (2015). A multifaceted investigation of the link between mental fatigue and task disengagement. *Psychophysiology* **52**, 305–315.

- Inzlicht, M., & Gutsell, J. N. (2007). Running on empty: Neural signals for self-control failure. *Psychological Science*, *18*, 933–937. doi: 10.1111/j.1467-9280.2007.02004.x
- Inzlicht, M., & Schmeichel, B. J. (2012). What is ego-depletion? Toward a mechanistic revision of the resource model of self-control. *Perspectives on Psychological Science*, *7*, 450–463. doi: 10.1177/1745691612454134
- Inzlicht, M., Schmeichel, B. J., & Macrae, C. N. (2014). Why self-control seems (but may not be) limited. *Trends in Cognitive Sciences*, *18*, 127–133. doi: 10.1016/j.tics.2013.12.009
- Ishihara, S. (1917). *Tests for Color Blindness*. Handaya, Tokyo: Hongo Harukicho.
- Jepma, M. & Nieuwenhuis, S. (2011). Pupil diameter predicts changes in the exploration-exploitation trade-off: evidence for the adaptive gain theory. *J Cogn Neurosci* **23**, 1587–1596.
- Job, V., Dweck, C., & Walton, G. (2010). Ego-depletion: Is it all in your head? Implicit theories about willpower affect self-regulation. *Psychological Science*, *21*, 1686–1693.
- Joshi, S., Li, Y., Kalwani, R.M., and Gold, J.I. (2016). Relationships between pupil diameter and neuronal activity in the locus coeruleus, colliculi, and cingulate cortex. *Neuron* **89**, 221–234.
- Kleiner, M., Brainard, D., Pelli, D., Ingling, A., Murray, R., & Broussard, C. (2007). What's new in Psychtoolbox-3. *Perception* **36**.
- Kurzban, R. (2010). Does the brain consume additional glucose during self-control tasks? *Evolutionary Psychology*, *8*, 244–259.^{[L][SEP]}
- Kurzban, R., Duckworth, A., Kable, J. W., & Myers, J. (2013). An opportunity cost model of subjective effort and task performance. *Behavioral and Brain Sciences*, *36*, 661 - 679. doi: 10.1017/S0140525X12003196

- Laeng, B., Ørbo, M., Holmlund, T., & Miozzo, M. (2011). Pupillary Stroop effects. *Cognitive Processing* **12**, 13–21.
- Lin, H., Saunders, B., Friese, M., Evans, N.J., & Inzlicht, M. (2020). Strong effort manipulations reduce response caution: A preregistered reinvention of the ego-depletion paradigm. *Psychol Sci.*, *31*(5): 531-547. doi:10.1177/0956797620904990
- Mathôt, S., Fabius, J., Van Heusden, E., & Van der Stigchel, S. (2018). Safe and sensible preprocessing and baseline correction of pupil-size data. *Behav. Res. Methods* **50**, 94–106.
- Mittner, M., Hawkins, G.E., Boekel, W., & Forstmann, B.U. (2016). A neural model of mind wandering. *Trends in Cognitive Sciences* **20**, 570–578.
- Muraven, M., & Slessareva, E. (2003). Mechanisms of self-control failure: Motivation and limited resources. *Personality and Social Psychology Bulletin*, *29*, 894–906.
- Muraven, M., Tice, D. M., & Baumeister, R. F. (1998). Self-control as a limited resource: Regulatory depletion patterns. *Journal of Personality and Social Psychology*, *74*, 774–789.
- Murphy, P. R., O’Connell, R. G., O’Sullivan, M., Robertson, I. H., & Balsters, J. H. (2014). Pupil Diameter Covaries with BOLD Activity in Human Locus Coeruleus. *Human Brain Mapping*, *35*, 4140-4154. <https://doi.org/10.1002/hbm.22466>
- O’Bryan, S. R. & Scolari, M. (2021). Phasic pupillary responses modulate object-based attentional prioritization. *Attention, Perception & Psychophysics*, *83*, 1491-1507.
- Price, M. M., & Brewer, L. E. (2015, February). *Social media use and self-control*. Poster presented at the annual meeting of the Society of Personality and Social Psychology, Long Beach, CA.
- Rajkowski, J., Kubiak, P., & Aston-Jones, G. (1994). Locus coeruleus activity in monkey: Phasic

- and tonic changes are associated with altered vigilance. *Brain Research Bulletin*, 35, 607–616.
- Ramos, B.P., & Arnsten, A.F.T. (2007). Adrenergic pharmacology and cognition: focus on the prefrontal cortex. *Pharmacol. Ther.*, 113, 523–536.
- Reimer, J., McGinley, M.J., Liu, Y., Rodenkirch, C., Wang, Q., McCormick, D.A., & Tolias, A.S. (2016). Pupil fluctuations track rapid changes in adrenergic and cholinergic activity in cortex. *Nat Commun* 7, 13289.
- Rucker, D. D., Preacher, K. J., Tormala, Z. L., & Petty, R. E. (2011). Mediation analysis in social psychology: Current practices and new recommendations. *Social and Personality Psychology Compass*, 5(6), 359–371.
- Sara, S. J., & Bouret, S. (2012). Orienting and Reorienting: The locus coeruleus mediates cognition through arousal. *Neuron*, 76, 130–141. doi: 10.1016/j.neuron.2012.09.011
- Schmeichel, B. J., Harmon-Jones, C., & Harmon-Jones, E. (2010). Exercising self-control increases approach motivation. *Journal of Personality and Social Psychology*, 99, 162–173.
- Schmeichel, B. J., Vohs, K. D., & Baumeister, R. F. (2003). Intellectual performance and ego-depletion: Role of the self in logical reasoning and other information processing. *Journal of Personality and Social Psychology*, 85, 33–46. doi: 10.1037/0022-3514.85.1.33
- Smallwood, J., Brown, K.S., Baird, B., Mrazek, M.D., Franklin, M.S., & Schooler, J.W. (2012). Insulation for daydreams: A role for tonic norepinephrine in the facilitation of internally guided thought. *PLoS One*, 7, e33706.
- Stroop, J. R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18, 643–662. doi: 10.1037/h0054651

- Van der Brink, R.L., Murphy, P.R., & Nieuwenhuis, S. (2016). Pupil diameter tracks lapses of attention. *PLOS One*, *11*, e0165274.
- van der Wel, P., & van Steenbergen, H. (2018). Pupil dilation as an index of effort in cognitive control tasks: A review. *Psychonomic Bulletin & Review*, *25*, 2005–2015.
- Varazzani, C., San-Galli, A., Gilardeau, S., & Bouret, S. (2015). Noradrenaline and dopamine neurons in the reward/effort trade-off: A direct electrophysiological comparison in behaving monkeys. *J Neurosci* **35**, 7866–7877.
- Vohs, K. D., Schmeichel, B. J. , Lohmann, S., Gronau, Q., Finley, A. J.... Wagenmakers, E.-J., & Albarracín, D. (in press). A multi-site preregistered paradigmatic test of the ego-depletion effect. *Psychological Science*.
- Wagner, D.D., & Heatherton, D.F. (2013). Self-regulatory depletion increases emotional reactivity in the amygdala. *Social Cognitive & Affective Neuroscience*, *8*, 410 – 417.