American Psychological Association

Journal of Experimental Psychology:
Learning, Memory, and Cognition

0278-7393/20/$12.00

2020, Vol. 46, No. 2, 280–295
http://dx.doi.org/10.1037/xlm0000720

Measuring Task Set Preparation Versus Mind Wandering Using Pupillometry

Keith A. Hutchison
Montana State University

Chad C. Moffitt
University of Utah

Katie Hart and Audrey V. B. Hood
Montana State University

Jason M. Watson
University of Colorado, Denver

Frank M. Marchak
Montana State University

We investigated participants' task set preparation by measuring changes in pupil diameter during a blank interval as they prepared for an easy (i.e., prosaccade) or difficult (i.e., antisaccade) trial. We used occasional thought probes to gauge "on-task" thoughts versus mind wandering. In both studies, participants' pupil diameters were larger when anticipating an antisaccade, relative to a prosaccade, trial. In contrast, their self-reported mind wandering depended upon whether the thought probes occurred after their target detection response (Experiment 1) or occurred in lieu of target detection (Experiment 2). In the latter case, self-reported mind wandering echoed the pupil diameter changes in demonstrating greater off-task behavior when preparing for a prosaccade trial. More important, trial type effects in pupil diameter emerged only when participants reported being "on-task," but disappeared during periods of mind wandering. These results demonstrate that changes in pupil diameter reflect the degree of preparatory control exerted for an upcoming trial, but only when attention is actively focused on the upcoming task.

Keywords: cognitive control, mind wandering, pupillometry, attention

Attentional control refers to the ability to orchestrate thought and action in accord with internal goals—particularly in situations that have the potential for distraction. These distractions can arise through external environmental stimuli or through internal sources, such as intruding thoughts, daydreaming, or mind wandering. Behaviors that involve a high probability of distraction, therefore, require mechanisms to keep attention focused on a goal that will result in successful task completion (Kane & Engle, 2002; Norman & Shallice, 1986).

Perhaps the greatest form of attentional control is overcoming a habitual response in favor of a less-practiced, task-appropriate response. A classic paradigm for investigating this involves making an eye movement either toward (prosaccade) or away (antisaccade) from an abruptly appearing visual cue (Kane, Bleckley, Conway, & Engle, 2001). Performance on prosaccade and antisaccade trials primarily reflects habitual and controlled processes, respectively, with antisaccade trial success requiring top-down suppression of the automatic response of looking toward the cue (Brown, Vilis, & Everling, 2007; Mazaheri, DiQuattro, Bengson, & Geng, 2011).

Consistent with this assumption, antisaccade performance is related to individual differences in working memory capacity, a measured construct that presumably reflects attentional control abilities in addition to primary and secondary memory (Engle, 2001; Shipstead, Lindsey, Marshall, & Engle, 2014; Unsworth, Fukuda, Awh, & Vogel, 2014). When antisaccade and prosaccade trials are blocked separately, working memory capacity correlates only with antisaccade performance (Kane et al., 2001; Unsworth, Schrock, & Engle, 2004). However, when trial types are intermixed and randomly cued, lower working memory capacity individuals perform worse on prosaccade trials as well, presumably because intermixing increases task difficulty through the need to update and maintain the appropriate task goal for each trial (Unsworth et al., 2014). The important factor may be the ability to fully engage such goals, above and beyond the ability to maintain them, as individuals with greater working memory capacity are

Editor's Note. Nash Unsworth served as Action Editor for this article.

This article was published Online First June 20, 2019.

Keith A. Hutchison, Department of Psychology, Montana State University; Chad C. Moffitt, Department of Psychology, University of Utah; Katie Hart and Audrey V. B. Hood, Department of Psychology, Montana State University; Jason M. Watson, Department of Psychology, University of Colorado, Denver; Frank M. Marchak, Department of Psychology, Montana State University.

Complete data files are available at http://www.montana.edu/attmemlab/.

Correspondence concerning this article should be addressed to Keith A. Hutchison, Department of Psychology, Montana State University, P.O. Box 173440, Bozeman, MT 59717-3440. E-mail: khutch@montana.edu

more likely to use long intervals to activate/engage (or instantiate) the antisaccade goal compared with individuals lower in working memory capacity (Meier, Smeekens, Silvia, Kwapil, & Kane, 2018). Thus, rather than successful performance depending on goal maintenance per se, performance may depend on the ability to fully engage (instantiate) the goal before it can ever be maintained.

In effect, one assumption of attentional control is that it takes time to fully engage, and, therefore, performance benefits arise when there is a preparatory period immediately before an attention-demanding trial (Mueller, Swainson, & Jackson, 2009). Evidence for this preparation has been found in blocks of prosaccade and antisaccade trials (Brown et al., 2007; Mueller et al., 2009). For antisaccade trials, prefrontal areas, including the left dorsolateral prefrontal cortex (LDPFC), exhibit significant preparatory period activity before the onset of the saccade stimulus. This leads to enhanced target-appropriate responding at target onset and, therefore, increased performance. Trials without such activity are associated with more errors. Greater prefrontal activity is also associated with faster reaction times (RTs) and less response-triggered activation in areas responsible for making eye movements (i.e., Frontal and Supplementary Eye Fields), suggesting prefrontal preparation allows such areas to work less at suppressing prosaccades. These findings demonstrate that individuals can engage attentional control preparatory processes in anticipation for cognitively demanding tasks, if given adequate time (Braver, Gray, & Burgess, 2007).

Researchers have more recently moved beyond appeals solely to prefrontal and parietal cortex activation as the determinant of attentional control performance and have begun examining the role of the locus-coeruleus-norepinephrine (LC-NE) system in controlling task engagement through modulating arousal, attention, and alertness (Gilzenrat, Nieuwenhuis, Jepma, & Cohen, 2010; Irons, Jeon, & Leber, 2017; Posner & Petersen, 1990; Unsworth & Robison, 2016, 2017b). The LC, located in the brain stem, has numerous projections throughout the central nervous system and is the main site of NE synthesis. According to the Adaptive Gain Theory (Aston-Jones & Cohen, 2005; Aston-Jones, Iba, Clayton, Rajkowski, & Cohen, 2007), the LC-NE system is sensitive to current task utility, allowing increased performance (i.e., exploitation) when task utility is high (i.e., effortful responding is likely to bring about task-related rewards) and allowing disengagement (i.e., exploration) when task utility is low. As long as baseline arousal is above minimal levels, such exploitation versus exploration are expressed through phasic and tonic modes of function (Usher, Cohen, Servan-Schreiber, Rajkowski, & Aston-Jones, 1999). Under phasic mode, there is lower baseline LC activity, but increased task-evoked NE release throughout the cortex, which increases the gain in processing task-relevant stimuli. This NE release enhances activation in frontal-parietal regions responsible for maintenance and use of task goals and for suppressing the default mode network (Raichle et al., 2001) that is active during rest periods and during internal thought (see Unsworth & Robison, 2017b). In contrast, under tonic mode, baseline LC activity is elevated and there is little-to-no phasic task-evoked response, reflecting disengagement from the current task and enhanced processing of task-unrelated stimuli. This mode allows individuals to seek out other activities and begin to mind wander.

Mind Wandering

Allowing one's mind to wander can constitute a major threat to goal engagement and goal maintenance. According to Smallwood and Schooler (2006), such task unrelated thoughts (TUTs) occur when attention shifts away from a current external task and toward internal processing of information. Moreover, such instances of mind wandering can be either intentional or unintentional, often depending upon task difficulty, such that participants more often report intentional mind wandering during easy tasks, but unintentional mind wandering during difficult tasks (Seli, Risko, Smilek, & Schacter, 2016). However, regardless of intentionality, mind wandering often produces ongoing task performance impairments (Seli, Cheyne, Xu, Purdon, & Smilek, 2015) on a variety of tasks including sustained attention (McVay & Kane, 2009), reading comprehension (Smallwood, McSpadden, & Schooler, 2008), and working memory (Mrazek, Phillips, Franklin, Broadway, & Schooler, 2013). Presumably, such performance decrements occur when mind wandering competes for working memory and attentional control resources otherwise directed toward the task (Smallwood & Schooler, 2006).

These patterns of mind wandering fit well within the Adaptive Gain Theory of LC-NE function described above. Specifically, during the tonic mode of task disengagement, baseline LC activity increases and phasic activity becomes "decoupled" from the perceptual events of the current task (Smallwood et al., 2011), allowing attention to focus on internal thoughts to receive potential benefits of mind wandering. Such benefits could include future goal planning (Baird, Smallwood, & Schooler, 2011), creative problem solving (Baird et al., 2012), and relief from boredom (Mooneyham & Schooler, 2013). Although such task disengagement could sometimes occur intentionally, because of diminished task utility (Gilzenrat et al., 2010), it could also occur because of dysregulation of the LC-NE system that produces occasional attentional lapses, especially among those low in attentional control (Unsworth & Robison, 2017b). Such attentional lapses would allow for the intrusion of task unrelated thoughts and subsequent task impairments (e.g., inaccurate responses or exceptionally long RTs) if a task requires controlled processing.

Validity of Mind Wandering Detection

A common approach to measuring such mind wandering is the probe-caught method, which involves inserting thought probes into a task that require participants to report what they were thinking about immediately before the thought probe appeared (Giambra, 1995; Schooler, Reichle, & Halpern, 2005). Because thought probes may catch instances of mind wandering before they reach awareness (Chin, Mrazek, & Schooler, 2012), they produce a good estimate of mind wandering frequency (Chin et al., 2012; Smallwood & Schooler, 2006). However, thought probes still require a certain level of introspection, of which some individuals may have particular difficulty, especially those low in conscious awareness of experiences and behaviors. An example of such "temporal dissociations" (Schooler, 2002) is when people fail to notice they have "zoned out" while reading. Other times, individuals may withhold admitting certain thoughts and feelings, perhaps believing they are unacceptable, and instead report a more acceptable account (Schooler, 2001). Furthermore, people differ in their confidence in reporting off-task thinking, which might be a result of

the thought probe options available. If someone's attention fluctuates within a trial between being on-task and off-task, or if the thought probe occurs when the participant is transitioning between being on-task or off-task, the participant will be forced into choosing one of the thought probe options available (Seli et al., 2015).

One way to combat this drawback is to combine self-reports of mind wandering with physiological measures, which can potentially validate the extent to which self-reports accurately reflect internal states. Cognitive pupillometry provides such a potentially useful noninvasive physiological measure that accurately indexes cognitive engagement. Moreover, as described below, changes in pupil diameter can serve as an index of ongoing LC activity (Rajkowski, Kubiak, & Aston-Jones, 1993).

Cognitive Pupillometry

Although the exact mechanisms are not completely understood, numerous studies have demonstrated that pupil diameter closely tracks LC activity, serving as an index of phasic versus tonic LC modes of task engagement versus disengagement, respectively (Franklin, Broadway, Mrazek, Smallwood, & Schooler, 2013; Gilzenrat et al., 2010; Rajkowski et al., 1993; Reimer et al., 2016; Unsworth, Robison, & Miller, 2018; Unsworth & Robison, 2016). Moreover, direct stimulation of LC results in rapid pupil dilation within approximately 1 s (Reimer et al., 2016).

Because pupillary responses allow for an indirect measure of LC activity, and because such activity itself reflects task engagement versus disengagement, it is not surprising that task-evoked pupillary responses have been obtained on a variety of tasks. Such tasks include short-term memory (STM; Kahneman & Beatty, 1966), sustained attention (van den Brink, Murphey, & Nieuwenhuis, 2016), working memory (Beatty & Kahneman, 1966), cognitive control (Rondeel, van Steenbergen, Holland, & van Knippenberg, 2015), and complex reasoning (Bradshaw, 1968; Hess & Polt, 1964). Moreover, task-evoked pupil changes are a sensitive measure of task difficulty and load, as well as mental effort (Beatty, 1982; Heitz, Schrock, Payne, & Engle, 2008; Hess & Polt, 1964; Kahneman, 1973; Peavler, 1974). Pupil changes can also provide an online index of the amount of resources allocated to a task (van Der Meer et al., 2010). For instance, changes in pupil diameter are greater when processing difficult, as opposed to simple, sentences (Just & Carpenter, 1993) and reflect use of response preparation following a cue in the AX-CPT task (Chatham, Frank, & Munakata, 2009; Chiew & Braver, 2013). Further, pupillary responses not only distinguish between on and off-task states, but also between different types of off-task states (i.e., mind wandering vs. being distracted), suggesting that pupil diameter can measure distinct types of attentional lapses (Unsworth & Robison, 2016). Therefore, using pupillometry cannot only provide a psychophysiological marker of cognitive effort, but also help eliminate problems associated with self-reported mind wandering. Finally, because of its high temporal resolution, phasic pupil changes can show the time course of cognitive engagement during both task-related and task-unrelated thoughts.

Task Engagement, Pupillometry, and Mind Wandering

Given the relation of LC-NE function to optimal task engagement versus task disengagement and mind wandering, researchers have examined how pupil diameter changes during hard versus easy tasks and during reported "on-task" versus "off-task" mental states. For instance, Gilzenrat et al. (2010, Experiment 3) found evidence for a tonic disengagement mode (i.e., higher baseline pupil diameter and less phasic task-evoked pupil change) when task utility reached a threshold of progressively increasing difficulty and diminished likelihood of reward. Similarly, Franklin et al. (2013) measured pupil diameter during reading and found the tonic disengagement pattern immediately before probes in which participants reported mind wandering. Smallwood et al. (2011) found evidence for tonic LC mode and more self-reported mind wandering during a simple choice RT task, but evidence for phasic LC mode (stimulus-evoked pupil change) in a more difficult working memory version of the paradigm.[1]

In addition to ongoing task difficulty, researchers have recently examined pupil changes during the preparatory period preceding difficult versus easy tasks or trials (Irons et al., 2017; Wang, Brien, & Munoz, 2015). For instance, Irons et al. (2017) gave participants a cue before each trial indicating whether they would have to perform an easy (e.g., red among blue items) or difficult (e.g., square among diamonds) target discrimination. They found greater cue-evoked pupil dilation in preparation for the more difficult target discrimination, indicating they were activating and maintaining task goals in preparation for the upcoming trial (see also Unsworth et al., 2018). Similarly, Wang et al. (2015) examined pupil size during the preparatory period in which a red or green fixation point directed participants to make a prosaccade or antisaccade response, respectively, to a stimulus presented 1.2 s later. Pupil size increased more in the 300 ms immediately preceding antisaccade trials than prosaccade trials. Combined with the Irons et al. (2017) study, this shows cue-evoked phasic changes in pupil diameter can accurately reflect degree of task set preparation.

LC-NE regulation could also vary across trials, which can lead to occasional attentional lapses during preparatory periods that can impair upcoming performance. Indeed, Unsworth and Robison (2015) found that lower working memory capacity individuals have greater trial-to-trial variability in baseline pupil diameters, suggesting more lapses in attention across trials. In a later latent variable analysis, they (Unsworth & Robison, 2017a) found that baseline and phasic pupil variability was associated with more mind wandering, which correlated with both attentional control and working memory capacity. Such observations are consistent with Unsworth and Robison's (2017b) conclusion that low working memory capacity individuals have greater LC-NE dysregulation, leading to more frequent attentional lapses, which produce greater incidents of mind wandering and impaired task performance. According to Unsworth et al. (2018), "pupillary responses provide a consistent means of tracking fluctuations in intrinsic

[1] Although tonic disengagement mode in these studies was associated with larger baseline pupil diameter, Unsworth and Robison (2015) found that errors on a working memory task were associated with much smaller than normal pretrial (tonic) baseline pupil diameters. Similarly, Kristjansson, Stern, Brown, and Rohrbaugh (2009) found that tonic pupil diameter was much smaller on trials preceding very slow RTs (indicative of lapses of attention) compared with trials where RT was close to the mean. Therefore, both larger and smaller than normal baseline pupil diameter (along with fluctuations in pupil diameter) are associated with lapses of attention and poorer performance (van den Brink et al., 2016; Unsworth & Robison, 2017b).

alertness and attention (linked to LC-NE and cortical sustained attention network functioning) during tasks that demand a great deal of sustained attention for optimal performance" (p. 1251).

Current Study

Similar to Wang et al. (2015), we examined cue-evoked phasic pupil changes during the fixation period preceding a pro- or antisaccade. However, within this task, we also examined the relationship between changes in pupil diameter and self-reports of mind wandering. In general, we predicted to replicate Wang et al.'s finding of greater phasic pupil dilation preceding antisaccade trials than prosaccade trials, indicating greater task set preparation. In contrast, we predicted no pupil increase and greater mind wandering in the preparatory period preceding the easier prosaccade task (Smallwood & Schooler, 2006). In addition, following past research (Smallwood et al., 2011; Unsworth & Robison, 2017a) this task-based difference in phasic pupil changes should be absent when people report mind wandering, indicative of decoupling attention from the necessary task set during attentional lapses. Consistent with this, individuals with greater variability in phasic pupil changes should have less saccade accuracy and report more frequent mind wandering.

Participants received a 1,500 ms cue to direct their eye movements either toward or away from a flashed stimulus to detect a target stimulus presented on the same (i.e., prosaccade) or opposite (i.e., antisaccade) side of the screen. A remote-controlled infrared eye camera measured pupil diameter during the postcue fixation period, which lasted between 500 and 8,000 ms and preceded the flashed stimulus. To obtain self-reported mind wandering, we presented thought probes on 25% (Experiment 1) or 17% (Experiment 2) of the trials. Measuring cue-evoked phasic changes in pupil diameter allowed us to both obtain a physiological measure of controlled processing and validate self-reported instances of mind wandering.

Experiment 1

Participants completed randomized prosaccade and antisaccade trials with varied fixation delays before saccade cue onset. Our saccade task is based on a similar antisaccade task by Hutchison (2007) and relies on target-identification accuracy as the index of saccade accuracy, based on the assumption that correct saccades should lead to correct target-identification whereas incorrect saccades should lead to chance responding.[2] Performance on this previous task has been found to positively correlate with other measures of attentional control (Hutchison, 2007; Hutchison, Heap, Neely, & Thomas, 2014; Shipstead et al., 2014). We used an infrared eye tracker to measure pupil diameter during the delay period when participants prepared for either an easy (prosaccade) or difficult (antisaccade) trial. We predicted that participants would mind wander less often during the more difficult antisaccade trials, as reflected by self-report and larger cue-evoked phasic pupil increases, than during prosaccade trials. Following Unsworth and Robison (2017a, 2017b), variability in phasic pupil changes should negatively correlate with saccade accuracy and positively correlate with mind wandering frequency.

Method

Participants and design. In accord with Simmons, Nelson, and Simonsohn (2011), we report how we determined our sample size, all data exclusions, all manipulations, and all measures in the study. This experiment was the final experiment (of three) in the second author's master's thesis, which investigated saccade performance, individual differences in working memory capacity, and mind wandering over fixation delays (Moffitt, 2013). To be consistent with the first two experiments in that thesis, we chose to run at least 100 participants and decided to continue running participants until the end of the semester even if we had already reached our goal to achieve enough participants after data exclusions. [We did not conduct a power analysis.] There were 135 Montana State University undergraduate students who participated for partial credit in an introductory psychology course. Although we did not ask participants to report their age or gender, this population typically features freshman between 18 and 20 years old, of whom approximately 55–60% are female.

We removed data from 17 participants because of technical issues with the eye tracker or participants choosing not to complete the experiment. This resulted in usable data from 118 participants. Each participant was tested individually in a laboratory session lasting approximately 1 hr. Four fixation delays (500, 2,000, 4,000, and 8,000 ms) and two saccade trial types (prosaccade and antisaccade) varied within subjects. We examined pupil diameter and target accuracy as a function of trial type, fixation delay, and self-reported mind wandering state.

Apparatus. We used E-studio E-prime software from Psychology Software Tools (Version 2.0.8.90) to program and present the saccade stimuli and a Panasonic CF-50 ToughBook laptop, with a Mobile Intel Pentium 4-M 2.00 GHz processor, 768 MB of RAM, and an AT Mobility Radeon 7500 Display Adapter to run the experiment. We presented task stimuli on a 17-in. NEC Multisync LCD 1760v monitor, with 1,024 × 768 screen resolution and a 60 Hz refresh rate, attached to the laptop via an RS232 USB serial port.

To measure pupil diameter, we used a contact-free, remote-controlled infrared eye camera (RED) with automatic gaze and head trackers designed by SensoMotoric Instruments (SMI). Thus, participants could freely view the monitor without having to use a chinrest. The tracker had binocular temporal resolution of 120 Hz, with spatial resolution of 0.03° and gaze position accuracy at 0.4°. Participants sat approximately 60 cm from the RED camera positioned directly under the monitor presenting task stimuli. An RS232 USB serial port on the Panasonic ToughBook laptop allowed the SMI RED tracking software to communicate with the E-prime software that ran the saccade task.

Procedure and stimuli. This study received permission from the Institutional Review Board at Montana State University. Figure 1 displays the trial sequences for both antisaccade and prosaccade trials. At the start of each saccade trial, participants saw a light gray background that remained onscreen while the following stimuli were presented sequentially. All stimuli were presented in

[2] Indeed, when we examined saccades, we found that correct antisaccade eye-movements led to 90% correct target identification. In contrast, when participants incorrectly made a prosaccade on an antisaccade trial their target identification was exactly at chance (.50).

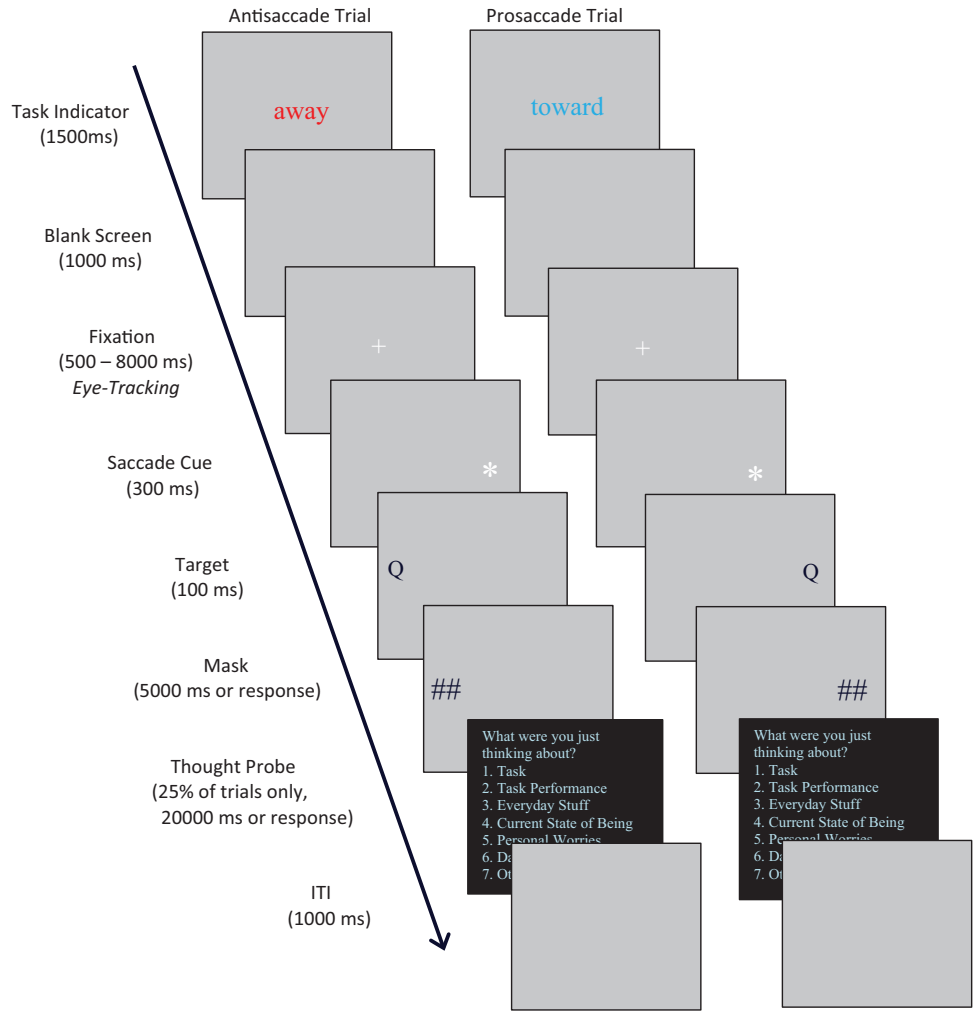

Figure 1. Trial sequence for antisaccade and prosaccade trials in Experiment 1. See the online article for the color version of this figure.

Courier New bold font. First, either the word "toward" (in blue 18-point font) instructed participants to look toward an upcoming cue to catch a target (prosaccade trial) or the word "away" (in red 18-point font) instructed participants to look away from the cue to catch the target (antisaccade trial). Next, a white 22-point central fixation cross (+) appeared and remained onscreen for 500, 2,000, 4,000 or 8,000 ms, which constituted the fixation delay periods. Then, a 36-point white saccade cue (*) appeared on either the left or right side of the computer screen for 300 ms. Following this, either an 'O' or a 'Q' target in black 20-point font appeared on the opposite side as the cue for 100 ms and was immediately replaced (masked) by two '##' symbols in black 25-point font, which remained on the screen for 5 s, or until target response. The cue, target, and mask appeared approximately 12.5 cm horizontally from the center of the fixation cross, resulting in approximately 11.89 ° visual angle between the location of the fixation cross and the location of the cue, target, and mask. Participants were instructed to identify the target by pressing either the 'O' or 'Q' button on the keyboard. Following a response, there was a 1,000 ms intertrial interval preceding the next trial. Luminance levels were 29 cd/m^2 for both the "away" and "toward" task cue screens and 30 cd/m^2 for all other trial screens.

Thought probes immediately followed 25% of the saccade trials (36 prosaccade and 36 antisaccade). We used thought probes based on McVay and Kane's (2009) instructions and explained these instructions to participants before beginning the saccade task. Thought probe instructions appeared in 14-point Courier New cyan font on a black background. The luminance level for this screen was 2 cd/m^2. Specifically, on thought-probe trials, participants saw the question "What were you just thinking about?" appear on the screen, along with seven response options: (a) *task* (i.e., thinking about the stimuli and the appropriate response); (b) *task performance* (i.e., evaluating one's own performance); (c) *everyday stuff* (i.e., thinking about recent or impending life events or tasks); (d) *current state of being* (i.e., thinking about conditions such as hunger or sleepiness); (e) *personal worries* (i.e., thinking about concerns, troubles, or fears); (f) *daydreams* (i.e., having fantasies disconnected from reality); or (g) *other* (i.e., other thought types). Following McVay and Kane (2009), we defined responses 1 and 2 as "on-task" thoughts and responses 3–7 as

"off-task" thoughts. Participants responded by pressing the corresponding number on the keyboard. After the participant's response, the next saccade trial began.

Participants first completed three practice blocks containing 12 trials each (36 total). The first practice block contained only prosaccade trials, the second block contained only antisaccade trials, and the final block contained six prosaccade trials and six antisaccade trials, presented in random order, designed to mimic the actual experiment. Participants were then instructed about the thought probes. Following the practice blocks, participants completed three experimental blocks, with each block containing 12 prosaccade trials and 12 antisaccade trials at each of the four fixation delays, resulting in 288 total experimental trials (96 trials per block). All trials occurred in random order. The number of fixation delay conditions and percentage of thought probes remained equal across blocks and saccade type, such that each fixation delay occurred 12 times for each saccade type per block, and each fixation delay per block contained three thought probes. Fixation delay and thought probes were presented in random order. The entire experimental session lasted approximately 1 hr.[3]

Phasic pupil diameter measurement. We measured pupillometry during the 4,000 and 8,000 ms conditions to examine the time course of cue-evoked phasic pupil diameter changes during the fixation delay as a function of expected trial type. The eye tracker failed to record pupil data from seven participants and recorded less than 65% of the trials for one other participant. For the remaining participants, blink trials (in which pupil diameter measured zero) were excluded from analysis, as were trials in which the eye tracker failed to capture at least half of the possible observations (sampled approximately every 8 ms). These criteria removed an average of 7.49 trials (3.6%) per participant. For each trial, the first 30 ms of the fixation screen served as a baseline to examine phasic cue-evoked pupil changes. We calculated cue-evoked phasic changes in pupil diameter (averaged across eyes) for each 1-s bin by subtracting the 30 ms baseline from the average pupil diameter during that bin so that positive values reflect dilation and negative values reflect constriction.

Results

In all analyses, we use a two-tailed p value of .05 as our criterion for significance. Because of unequal variance across delays, we corrected all such p values using the Greenhouse-Geisser correction.

Behavioral results.

Saccade target accuracy. There was a main effect of trial type (F[1, 117] = 1130.733, $p < .001$, $\eta_p^2 = .906$), with higher participant accuracy on prosaccade ($M = .918$, $SE = .005$) than on antisaccade trials ($M = .576$, $SE = .010$). There was also a main effect of delay (F[3, 351] = 33.134, $p < .001$, $\eta_p^2 = .221$) and a Trial Type × Delay interaction, $F(3, 351) = 2.965$, $p < .037$, $\eta_p^2 = .025$. To decompose this interaction, we examined the effects of delay separately for antisaccade and prosaccade trials. For both kinds of trials, accuracy improved across delay, $F(3, 351) = 12.835$, $p < .001$, $\eta_p^2 = .099$ and $F(3, 351) = 30.938$, $p < .001$, $\eta_p^2 = .209$ for antisaccade and prosaccade trials, respectively. However, the pattern differed depending upon trial type. For antisaccade trials, this improvement followed a linear trend, $F(1, 117) = 25.491$, $p < .001$, $\eta_p^2 = .179$, such that accuracy increased significantly from 500 ms ($M = .546$, $SE = .011$) to 2,000 ms ($M = .568$, $SE = .010$), numerically, but not significantly, from 2,000 to 4,000 ms ($M = .582$, $SE = .012$), and significantly from 4,000 to 8,000 ms ($M = .606$, $SE = .013$). In contrast, for prosaccade trials, accuracy significantly increased from 500 ms ($M = .883$, $SE = .008$) to 2,000 ms ($M = .927$, $SE = .005$), but then remained stable for the 4,000 ms ($M = .931$, $SE = .006$) and 8,000 ms ($M = .932$, $SE = .006$) delays.

Thought-probe responses. We used a 2 (trial type) × 4 (delay) analysis of variance (ANOVA) to examine thought probe responses. Overall, there was a significant effect of trial type ($F(1, 117) = 7.224$, $p = .008$, $\eta_p^2 = .058$). Surprisingly, participants reported a greater proportion of TUTs on antisaccade trials ($M = .41$, $SE = .02$) than on prosaccade trials ($M = .38$, $SE = .02$). Neither the main effect of delay, nor the Trial Type × Delay interaction were significant, $F(1, 117) = 0.505$, $p = .679$, $\eta_p^2 = .004$ and $F(1, 117) = 2.058$, $p = .106$, $\eta_p^2 = .017$, respectively.

Pupil diameter analyses.

Trial type effects. We used a 2 (trial type) × 4 (delay) ANOVA to examine cue-evoked phasic pupil changes during the 4,000 ms fixation delay and a 2 (trial type) × 8 (delay) ANOVA to examine these changes during the 8,000 ms fixation delay. This initial analysis included both accurate and error trials. The 4,000 and 8,000 ms data are shown in Figure 2. In both analyses, the main effect of delay was significant (F[3, 330] = 4.44, $p = .025$, $\eta_p^2 = .039$; F[1, 110] = 52.89, $p < .001$, $\eta_p^2 = .325$, respectively) indicating reduced pupil diameter across the delay. More important, there was a main effect of trial type (F[1, 110] = 24.35, $p < .001$, $\eta_p^2 = .181$; F[1, 110] = 45.50, $p < .001$, $\eta_p^2 = .293$ for the 4,000 and 8,000 ms analyses, respectively). Specifically, participants' pupil diameters were smaller when preparing for a prosaccade trial, relative to an antisaccade trial. Finally, the interaction was significant in both analyses (F[3, 330] = 16.12, $p < .001$, $\eta_p^2 = .128$; F[7, 770] = 24.68, $p < .001$, $\eta_p^2 = .183$, respectively), indicating that changes in pupil diameter across delay depended upon expected trial type. To examine this interaction, we explored pupil diameter changes separately for each trial type. When expecting an antisaccade trial, participants' pupil diameter remained relatively fixed during the first 4 s, and then reduced during the final 4 s (see Figure 2). This was revealed by a null effect of delay in the 4,000 ms analysis (F[3, 330] = 0.32, $p = .670$, $\eta_p^2 = .003$) combined with a significant effect of delay in the 8,000 ms analysis (F[7, 770] = 17.66, $p < .001$, $\eta_p^2 = .138$). In the 8,000 ms analysis, pupil diameter increased significantly above baseline during the first 1,000 ms (t[110] = 3.14, $p < .002$), returned to baseline levels during seconds 2–5 (ts < 1.1, ps > .31), and then decreased below baseline during seconds 6–8 (all ts > 2.09, all

[3] In addition to these tasks, we also included a shortened OSPAN task (Hutchison, 2007). The behavioral patterns replicated previous work on individual differences, saccade performance, and fixation delay (Kane et al., 2001; Meier et al., 2018; Unsworth et al., 2004) showing high span advantages on antisaccade performance, especially at longer delays. However, we do not report the OSPAN results because (a) many subjects in Experiment 2 were accidentally run on an incorrect version of this task and (b) we only had a single measure of WMC, whereas a composite of multiple measures is recommended (Foster et al., 2015). Thus, the data were not helpful in demonstrating whether or not WMC contributes to the current results. Nonetheless, we included the OSPAN data in the uploaded data file for anyone interested.

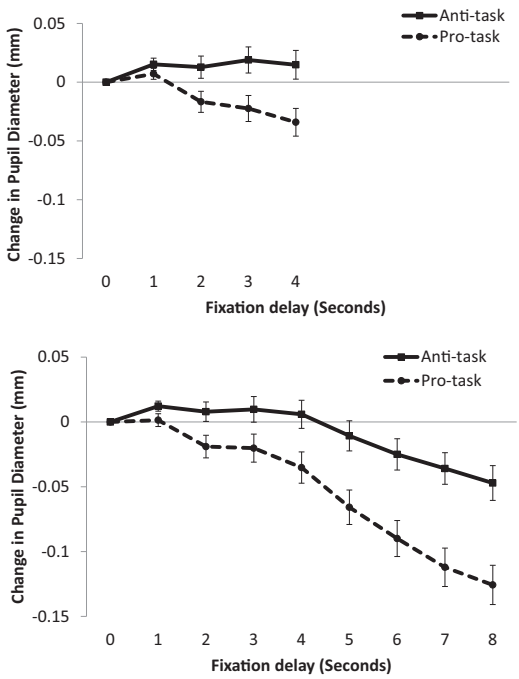

Figure 2. Mean pupil diameter change across delay in Experiment 1 as a function of cued task for trials lasting at least 4,000 ms (top) and trials lasting 8,000 ms (bottom). Error bars reflect *SEM*.

*p*s $< .04$). In contrast, when expecting a prosaccade trial, participants' pupil diameter continually decreased across delay in both analyses, as revealed by significant effects of delay ($F[3, 330] = 13.55$, $p < .001$, $\eta_p^2 = .11$; $F[7, 770] = 77.08$, $p < .001$, $\eta_p^2 = .41$, respectively). In the 8,000 ms analyses, pupil diameter decreased from baseline within 2 s ($t[110] = -2.18$, $p = .04$, respectively) and continually decreased thereafter. In the 4,000 ms analyses, the pupil decrease was marginal at 2 s ($t[110] = -2.18$, $p = .064$), but significant at the later durations. Finally, phasic pupil diameter changes reliably differed between prosaccade and antisaccade trial types at all delays for both the 4,000 and 8,000 ms analyses, despite no significant difference between trial types during the first 30 ms used as baseline, $t(110) = -1.75$, $p = .08$.

Antisaccade Accuracy × Pupil Diameter effects. Because prosaccade accuracy was near ceiling, we did not have enough error observations to examine prosaccade pupil changes as a function of trial accuracy. Therefore, we focused on antisaccade trials to examine pupil diameter on successful versus nonsuccessful trials. We used a 2 (accuracy) × 4 (delay) ANOVA to examine the 4,000 ms trials and a 2 (accuracy) × 8 (delay) ANOVA to examine the 8,000 ms trials. These data are shown in Figure 3. In the 4,000 ms analysis, pupils significantly dilated above baseline on accurate trials ($M = .024 \pm .020$ mm, $\pm$ = 95% confidence interval [CI]), but not on error trials ($M = 0.005 \pm .020$ mm), revealing greater engagement of cognitive effort (or preparatory control) before accurate responses. This difference resulted in a significant main effect of accuracy, $F(1, 109) = 4.79$, $p = .031$, $\eta_p^2 = .042$. There was also a marginal Accuracy × Delay interaction ($F[1, 109] = 2.55$, $p = .056$, $\eta_p^2 = .023$), indicating greater pupil differences between correct and error trials across delay. On accurate trials, pupil diameter increased above baseline at the 1 s ($t[109] = 3.367$, $p < .001$), 3 s ($t[109] = 2.310$, $p = .023$), and 4 s ($t[109] = 2.093$, $p = .039$) delays, and was marginally above baseline at the 2 s delay ($t[109] = 1.908$, $p = .059$), suggesting that participants maintained cognitive effort across time to improve accuracy performance. In contrast, on error trials, pupil diameter marginally increased above baseline at the 1 s interval ($t[110] = 1.715$, $p = .089$), but did not differ from baseline at any of the other delays (all *p*s $> .510$). This accuracy effect was in the same direction for the 8,000 ms analysis, but did not reach significance $F(1, 110) = 1.304$, $p = .256$, $\eta_p^2 = .012$.

Variability in phasic pupil response. We conducted a final analysis to examine individual differences in saccade accuracy and reported TUTs as a function of variability in phasic pupil responses. To accomplish this, we took the *SD* of participants' phasic pupil response in both the 4,000 and 8,000 ms conditions for each trial type and correlated this with their accuracy and self-reported TUTs. These correlations are shown in Table 1. Within measures, there were stable individual differences in phasic pupil variability, mind wandering rate, and saccade accuracy across trials. Specifically, there were strong correlations (i.e., Pearson *r* above .500) between the four Trial Type × Delay conditions in phasic pupil variability and TUT rate and between both prosaccade accuracy and antisaccade accuracy across delay conditions. Examining phasic pupil variability, as predicted, this measure correlated negatively with saccade accuracy (see rows 5–8, columns 1–4) and positively with TUT rate (see rows 9–12, columns 5–8).

Discussion

The surprising result from Experiment 1 was that self-reported rates of mind wandering as a function of trial type were in conflict

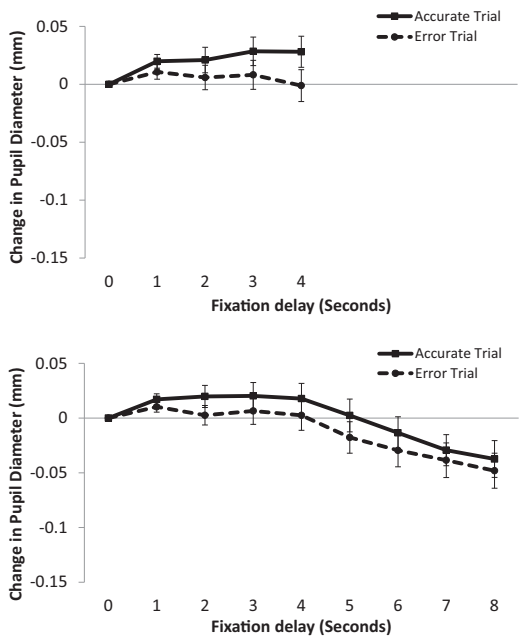

Figure 3. Mean pupil diameter change across delay for antisaccade trials in Experiment 1 as a function of trial accuracy for trials lasting at least 4,000 ms (top) and trials lasting 8,000 ms (bottom). Error bars reflect *SEM*.

Table 1
Correlations Between Saccade Accuracy, Variability (SD) in Phasic Pupil Response, and Task Unrelated Thoughts (TUTs) at the 4,000 and 8,000 ms Conditions in Experiment 1

Experiment 1	1	2	3	4	5	6	7	8	9	10	11	12
1. Pro, 4,000 ms accuracy	—											
2. Pro, 8,000 ms accuracy	.562*	—										
3. Anti, 4,000 ms accuracy	.109	.229*	—									
4. Anti, 8,000 ms accuracy	.187*	.218*	.738*	—								
5. Pro, 4,000 ms pupil *SD*	−.222*	−.151	−.290*	−.361*	—							
6. Pro, 8,000 ms pupil *SD*	−.290*	−.170	−.336*	−.352*	.880*	—						
7. Anti, 4,000 ms pupil *SD*	−.189*	−.124	−.430*	−.420*	.809*	.830*	—					
8. Anti, 8,000 ms pupil *SD*	−.232*	−.222*	−.399*	−.465*	.788*	.785*	.793*	—				
9. Pro, 4,000 ms, TUT	.081	−.122	−.208*	−.156	.240*	.246*	.294*	.353*	—			
10. Pro, 8,000 ms, TUT	−.001	−.208*	−.282*	−.272*	.375*	.333*	.399*	.450*	.722*	—		
11. Anti, 4,000 ms, TUT	.064	−.083	−.149	−.116	.229*	.235*	.247*	.291*	.698*	.649*	—	
12. Anti, 8,000 ms, TUT	−.077	−.229*	−.286*	−.269*	.420*	.419*	.407*	.492*	.719*	.725*	.755*	—

* $p < .05$ (two-tailed).

not only with previous studies investigating effects of task difficulty (see Smallwood & Schooler, 2006, for a review), but also with participants' own pupillometry data. If the thought probes truly captured participants' mind wandering, then we should have observed the reverse pattern, such that anticipation of a difficult antisaccade trial should have reduced mind wandering, relative to performing the more habitual prosaccade task (Smallwood, Beach, Schooler, & Handy, 2008; Smallwood, Nind, & O'Connor, 2009). Indeed, this is what we found in the pupillometry analysis. Specifically, participants' pupils were considerably smaller, relative to baseline, on prosaccade trials than on antisaccade trials, and this effect of trial type increased over time. Such a pattern is consistent with participants remaining fully engaged on the task when expecting difficult antisaccade trials, but allowing themselves to relax when expecting easy prosaccade trials.

Given these findings, it is possible that participants used their task performance to base their thought probe response. With excellent performance on prosaccade trials and poor performance on antisaccade trials, participants could have occasionally rationalized their poor antisaccade performance as the result of mind wandering. We think presenting the probes immediately after prosaccade or antisaccade responses might have caused such potential reactivity. In other words, on thought probe trials, participants would respond to a thought probe immediately after providing their target identity response. This is problematic for several reasons. First, because the probe is presented after the saccade response has already occurred, it might not truly assess what participants were thinking about before their response. Some participants may simply be unable to assess their previous state of mind upon reflection. Second, and related, given a potential difficulty in introspection, participants might have justified unsure responses by reporting mind wandering. Consistent with this second possibility, participants were more likely to report mind wandering following errors than accurate responses on both prosaccade ($t[115] = 2.625$, $p = .010$) and antisaccade trials, $t(114) = 2.887$, $p = .005$. (Note, two participants reported no mind wandering on antisaccade trials and one participant reported no mind wandering on prosaccade trials.)

As predicted, the cue-evoked phasic pupil changes differed as a function of trial type. When participants were expecting prosaccade trials, their pupils monotonically constricted relative to baseline. In contrast, when expecting an antisaccade, their pupils either dilated or remained flat (depending upon accuracy) during the first 4 s and then gradually constricted thereafter, suggesting an early engagement of cognitive effort followed by a steady decrease in effort. This pattern differs from Wang et al.'s (2015) finding of initial constriction for both types of trials combined with greater dilation for antisaccade then prosaccade trials in the 200 ms immediately preceding onset of the saccade stimulus. We believe this difference is because of the variable delays used in the current study, requiring participants to engage preparatory control early in the fixation delay and maintain vigilance throughout the entire delay. In contrast, when the stimulus always occurs at the same time (constant ISI), it requires less focused attention and typically results in better overall performance on sustained attention tasks (Unsworth et al., 2018). Indeed, in a subsequent study using the identical program and a constant 5,000 ms delay, we found the typical constant ISI pattern of gradually increasing pupil dilation that peaks immediately before stimulus onset (Bradshaw, 1968, 1969; Jennings, van der Molen, & Steinhauer, 1998; Richer & Beatty, 1987; Richer, Silverman, & Beatty, 1983; Unsworth et al., 2018; van der Molen, Boomsma, Jennings, & Nieuwboer, 1989).

The current requirement to maintain vigilant preparation across variable delays up to 8 s likely not only influenced the pattern of phasic pupil changes, but also made this task much more difficult. This temporal uncertainly, combined with the requirement to switch tasks based on random cues, likely reduced mind wandering relative to less demanding tasks such as the SART task, in which young adult TUT rates are usually around 50% (McVay & Kane, 2009; McVay, Meier, Touron, & Kane, 2013). This difficulty is consistent with the generally sparse pupil dilation and our finding that overall antisaccade accuracy was not much above chance ($M = 57\%$).

Finally, we examined variability in phasic pupil changes to track individual differences in attentional lapses across trials. Consistent with Unsworth and colleagues' LC-NE account for individual variations in attentional control (Unsworth & Robison, 2017a, 2017b; Unsworth et al., 2018), we found that individuals showing greater fluctuations in attentional preparation, as measured through variability in phasic pupil changes across trials, were less accurate on both prosaccade and antisaccade trials. Consistent with Un-

sworth and Robison (2017b), this measure of attentional lapses also positively correlated with self-reported mind wandering frequency, although we need to be cautious in interpreting any correlations with mind wandering in Experiment 1, because of the potential participant reactivity. Despite this potential limitation, however, the results overall appear consistent with the view that performance problems during attentionally demanding tasks reflect moment-to-moment fluctuations in attentional control and the resulting mind wandering.

Experiment 2

To provide a more valid measure of mind wandering, in Experiment 2 we presented thought-probes in *lieu* of the saccade cue-target-mask stimuli. In this way, probes should provide a more accurate "in-the-moment" assessment of mind wandering, while also eliminating the possibility of using thought probes to justify nonconfident responses. Further, we reduced the frequency of thought probe presentation to 17% to reduce any potential artifact of thought probes serving as reminders to stay on-task. If this presentation change indeed increases the validity of thought-probes, we would expect to see a reversal of the mind wandering results in Experiment 1, such that participants should report more mind wandering on prosaccade than antisaccade trials. Yet, we also predicted that the results should parallel the results of Experiment 1 in terms of cue-evoked phasic pupil changes correlating with saccade performance and rate of mind wandering. By extension, if we can validate the self-reported mind wandering with pupil responses, we will then have more confidence in participants' probe responses and can then use them to examine phasic pupil responses separately for trials in which participants are "on-task" versus "off-task," as well as use fluctuations in pupil dilation to predict mind wandering (or off-task thoughts).

Method

Participants and design. We again sought to run at least 100 participants in the current experiment and, therefore, ran as many as possible until the end of the semester. There were 103 male and female Montana State undergraduate students who participated for partial course credit in an introductory psychology course. We removed data from eight participants because of technical issues with the eye tracker. This resulted in usable data from 95 participants. Each participant was tested individually in a laboratory session lasting approximately 1 hr. Four fixation delays (500, 2,000, 4,000, and 8,000 ms) and two saccade trial types (prosaccade and antisaccade) varied within subjects. We examined pupil diameter and saccade accuracy as a function of trial type, fixation delay, saccade accuracy, and self-reported mind wandering state.

Apparatus, stimuli, and procedure. The apparatus and stimuli are identical to Experiment 1. We altered the thought probe procedure from Experiment 1 in two ways. First, we reduced the number of thought probes from 25% of trials to 17% of trials to reduce reminders to stay on-task, while also collecting enough thought probe trials (48 total) to measure mind wandering. Second, thought probe trials consisted only of the trial type cue and fixation delay, followed immediately by the thought probe (see Figure 4, for an example of a normal and thought probe antisaccade trial in Experiment 2).

Phasic pupil diameter measurement. We again used the 4,000 and 8,000 ms conditions to examine the time course of cue-evoked phasic pupil diameter changes during the fixation delays as a function of expected trial type. The eye tracker failed to record data from five participants and recorded less than 65% of the trials for three other participants. For the remaining participants, blink trials (in which pupil diameter measured zero) were excluded from analysis, as were trials in which the eye tracker failed to capture at least half of the possible observations (at approximately 8 ms each). These criteria removed an average of 12.1 trials (4.5%) per participant. The first 30 ms of each fixation delay again served as a baseline measure to examine phasic cue-evoked pupil changes.

Results

Behavioral results.

Saccade target accuracy. There was a main effect of trial type ($F[1, 94] = 1213.151$, $p < .001$, $\eta_p^2 = .928$), with participant accuracy higher on prosaccade trials ($M = .910$, $SE = .009$) than on antisaccade trials ($M = .560$, $SE = .008$). There was also a main effect of delay ($F[1, 94] = 16.696$, $p < .001$, $\eta_p^2 = .151$), with participant accuracy significantly increasing across the first three delays of 500 ms ($M = .705$, $SE = .008$), 2,000 ms ($M = .736$, $SE = .007$), and 4,000 ms ($M = .750$, $SE = .008$) and remaining stable for the 8,000 ms delay ($M = .750$, $SE = .009$). The Trial Type × Delay interaction was not significant, $F(1, 94) = 0.551$, $p = .639$, $\eta_p^2 = .006$.

Thought-probe responses. We used a 2 (trial type) × 4 (delay) ANOVA to examine thought probe responses. Overall, there was a significant effect of trial type ($F[1, 94] = 7.164$, $p = .009$, $\eta_p^2 = .071$). In contrast to Experiment 1, yet as predicted, participants reported a greater proportion of TUTs on prosaccade trials ($M = .48$, $SE = .03$) than on antisaccade trials ($M = .43$, $SE = .03$). No other effects were significant, Fs < 1, η_p^2s $< .01$.

Pupil diameter analyses.

Trial type effects. We again used a 2 (trial type) × 4 (delay) ANOVA to examine cue-evoked phasic pupil changes during the 4,000 ms fixation delay and a 2 (trial type) × 8 (delay) ANOVA to examine these changes during the 8,000 ms delay. Again, this initial analysis included both accurate and error trials. These data are shown in Figure 5 for the 4,000 ms (top) and 8,000 ms (bottom) data sets. Consistent with Experiment 1, cue-evoked phasic changes in pupil diameter depended upon trial type, with a reduction in pupil diameter when participants prepared for a prosaccade trial relative to an antisaccade trial, $F(1, 94) = 16.915$, $p < .001$, $\eta_p^2 = .153$; $F(1, 94) = 19.980$, $p < .001$, $\eta_p^2 = .175$ for the 4,000 and 8,000 ms data, respectively. Also, pupil diameter again decreased across delays ($F[3, 282] = 16.212$, $p < .001$, $\eta_p^2 = .147$; $F[7, 658] = 55.299$, $p < .001$, $\eta_p^2 = .370$, respectively). The Trial Type × Delay interaction was again significant ($F[3, 282] = 14.206$, $p < .001$, $\eta_p^2 = .131$; $F[7, 658] = 15.046$, $p < .001$, $\eta_p^2 = .138$, respectively), indicating trial type differences in the pattern of phasic pupil changes across delay. Although participants' pupil diameter decreased regardless of upcoming trial type, the decrease was greater when expecting a prosaccade trial than when expecting an antisaccade trial (see Figure 5). When expecting an antisaccade trial, in the 4,000 ms analysis, there was significant pupil dilation during the first 1,000 ms ($t[94] = 2.39$, $p = .038$), followed by a

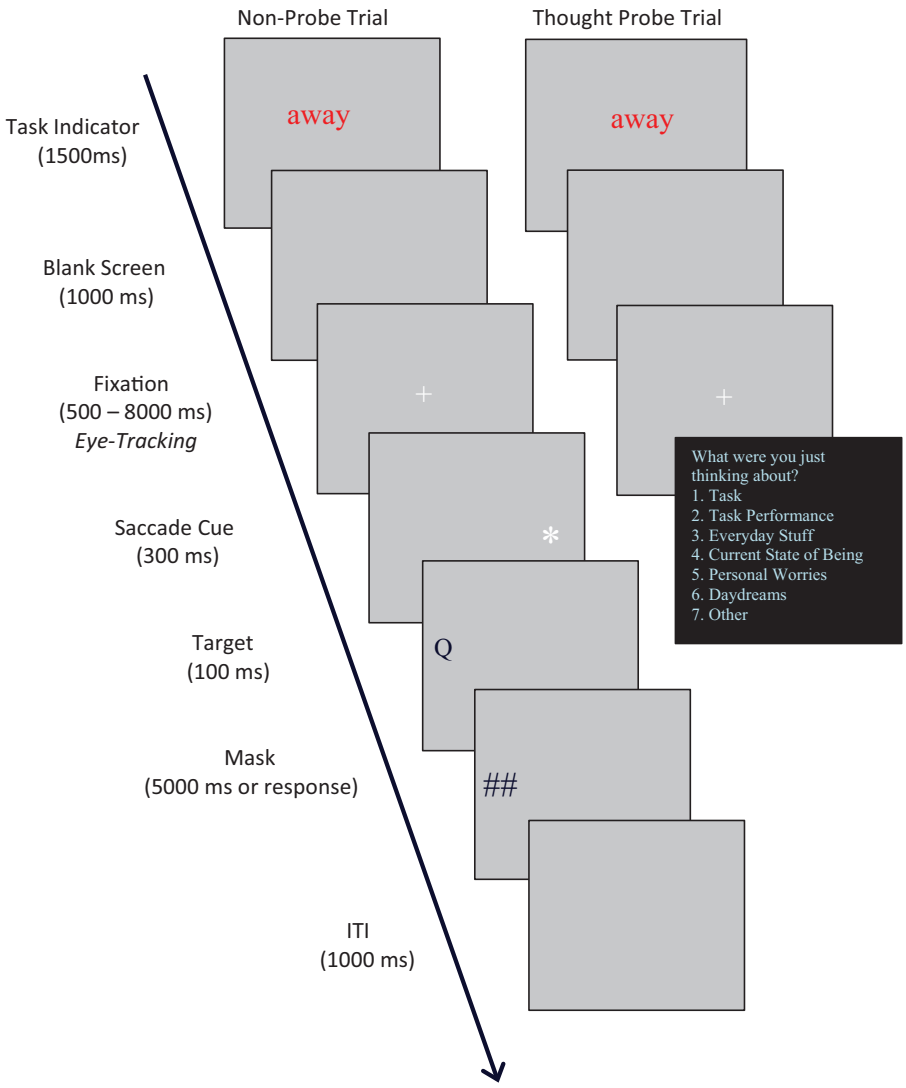

Figure 4. Sample trial sequence for normal and thought probe trials in Experiment 2. See the online article for the color version of this figure.

return to baseline during seconds 2–4 (all ts < 1.4, all ps $> .17$), suggesting early engagement of cognitive effort. In the 8,000 ms analyses, this pattern for the first 4 s remained, along with decreasing pupil diameter below baseline during seconds 6–8 (all ts > 2.48, all ps $< .015$). In contrast to the antisaccade pattern, pupil diameter decreased monotonically when participants expected a prosaccade trial. In both the 4,000 and 8,000 ms analyses, pupil diameter decreased from baseline within 2 s ($t[94] = -2.72$, $p = .008$; $t[94] = -2.77$, $p = .007$, respectively) and continually decreased thereafter (all ts > 2.93, all ps $< .005$), suggesting a continual decrease in cognitive effort (or controlled processing).

In Experiment 2, pupil diameter at baseline was .016 ± .014 mm larger for prosaccade trials ($M = 4.247$ mm) than antisaccade trials ($M = 4.231$ mm). As with Experiment 1, phasic pupil changes reliably differed between prosaccade and antisaccade tasks at all delays for both the 4,000 and 8,000 ms trials.

Antisaccade Accuracy × Trial Type effects. We again used two 2 (accuracy) × 4 (delay) ANOVAs to examine cue-evoked phasic pupil changes as a function of antisaccade response accuracy and delay on 4,000 and 8,000 ms trials. These data are shown in Figure 6. The main effect of accuracy was not significant in the 4,000 ms analysis ($F[1, 94] = 0.050$, $p = .824$, $\eta_p^2 = .001$), but was marginally significant in the 8,000 ms analysis ($F[1, 94] = 3.303$, $p = .072$, $\eta_p^2 = .034$). The main effect of delay was significant in both analyses ($F[1, 94] = 5.163$, $p = .009$, $\eta_p^2 = .052$ and $F[1, 94] = 16.501$, $p < .001$, $\eta_p^2 = .149$ for the 4,000 and 8,000 ms analyses, respectively). There was a significant Accuracy × Delay interaction on 8,000 ms trials $F(1, 94) = 2.675$, $p = .049$, $\eta_p^2 = .028$. This interaction reflected a steeper decrease in pupil diameter from baseline during error trials than accurate trials, suggesting a greater decrease in cognitive effort across time before error trials compared with accurate trials. Specifically, on error trials, pupil diameter dilated above baseline at 1 s ($t[94] = 3.177$, $p = .002$), did not differ from baseline during seconds 2–5 (all ts < 1.93, ps $> .055$), and decreased below baseline on seconds 6–8 (all ts > 2.63, ps $< .01$). In contrast, on accurate trials, pupil diameter

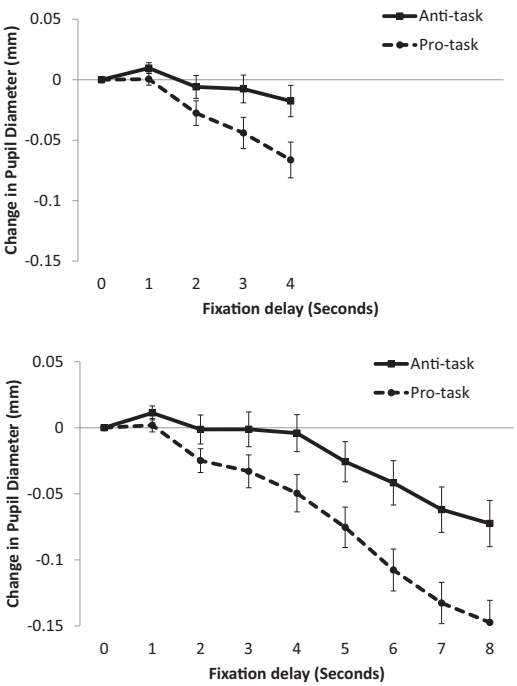

Figure 5. Mean change pupil diameter change across delay in Experiment 2 as a function of cued task for trials lasting at least 4,000 ms (top) and trials lasting 8,000 ms (bottom). Error bars reflect *SEM*.

marginally dilated above baseline at 1 s (t[94] = 1.872, p = .064), did not differ from baseline during seconds 2–7 (all ts < 1.92, ps > .057), and only decreased below baseline during the last second, t(94) = 2.192, p = .031. This pattern suggests fast engagement of control preceding both error and accurate trials, but a continual disengagement on error trials that happens at a faster rate than on accurate trials.

Trial Type × Thought Probe effects. Because self-reported mind wandering in Experiment 2 matched predictions based on previous studies, and was in line with the cue-evoked phasic pupil change data, we felt confident in the validity of this self-report. Therefore, we examined pupil diameter separately for trials in which participants reported being on-task versus off-task. These data are shown in Figure 7 for the 4,000 ms (top) and 8,000 ms (bottom) conditions. This analysis only included the 17% of trials in which participants received a thought probe and only included participants who reported both off-task and on-task responses for each trial type. Because of this, data were missing from 33 participants in the 4,000 ms dataset (remaining N = 62) and 41 participants in the 8,000 ms dataset (remaining N = 54).

We used a 2 (trial type) × 2 (probe response) × 4 (delay) ANOVA to examine the 4,000 ms trials and a 2 (task) × 2 (probe response) × 8 (delay) ANOVA to examine the 8,000 ms trials. On the 4,000 ms trials (see top of Figure 7), none of the effects reached significance (all Fs < 2.08, ps > .155, η_p^2s < .033). On the 8,000 ms trials, there was a significant three-way Trial Type × Probe Response × Delay interaction (F[7, 371] = 3.46, p = .015, η_p^2 = .061), indicating that thought probe responses moderated the trial type differences in pupil changes across delay described above (see bottom of Figure 7). To decompose this interaction, we analyzed the pupil data separately depending upon whether people reported being on-task or mind wandering. When participants reported being on-task (dark lines in Figure 7), cue-evoked phasic pupil changes across delay were strongly dependent upon expected trial type, as revealed by a significant Trial Type × Delay interaction (F[7, 371] = 5.449, p = .001, η_p^2 = .093). Significant trial type differences in pupil change emerged starting at 6 s (t[53] = 2.40, p = .02) and continued until the thought probe. Further, when on-task, participants' pupil diameter remained constant across delay when expecting to perform an antisaccade (all ts < 1.70, ps > .09), but significantly decreased from baseline starting at 5 s (all ts > 2.31, ps < .025) when expecting to perform a prosaccade. In contrast, when off-task, there was only a main effect of delay, such that pupils continually constricted regardless of which task participants expected to perform, F(7, 371) = 10.587, p < .001, η_p^2 = .166).

As a further illustration, we also examined this three-way interaction in terms of trial type. When expecting an antisaccade trial (solid lines in Figure 7), the Thought Probe × Delay interaction was significant (F[7, 371] = 2.999, p = .029, η_p^2 = .054), such that pupil diameter was constricted during off-task thoughts, relative to on-task thoughts, and this difference was significant starting at 5 s. In contrast, when expecting a prosaccade trial (dotted lines in Figure 7), pupil diameter decreased equally over delay regardless of whether participants reported being on-task or off-task, F(7, 371) = 1.357, p = .259, η_p^2 = .025.

Variability in phasic pupil response. As with Experiment 1, we examined individual differences in saccade accuracy, self-reported mind wandering, and variability in phasic pupil responses. The correlations are shown in Table 2. There were again stable individual differences in phasic pupil variability, mind wandering rate, and saccade accuracy across trials, as demonstrated by strong

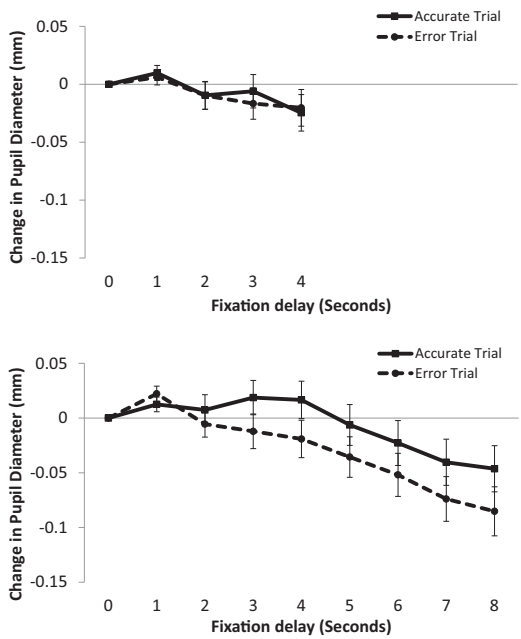

Figure 6. Mean pupil diameter change across delay for antisaccade trials in Experiment 2 as a function of trial accuracy for trials lasting at least 4,000 ms (top) and trials lasting 8,000 ms (bottom). Error bars reflect *SEM*.

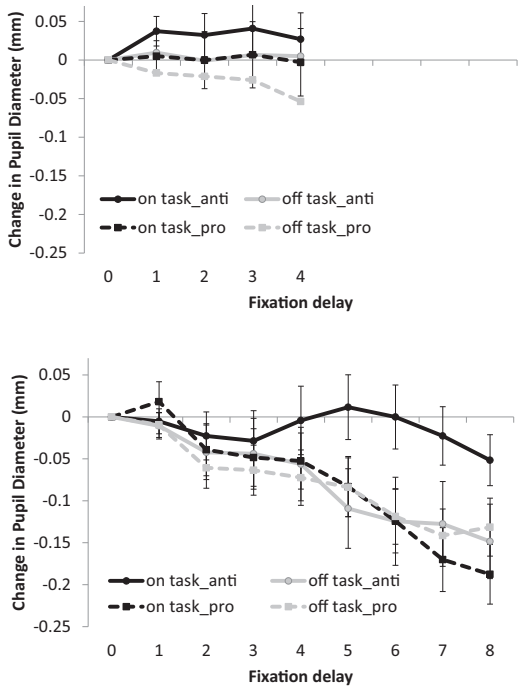

Figure 7. Mean change pupil diameter change across delay in Experiment 2 as a function of cued task and thought probe response for trials lasting at least 4,000 ms and trials lasting 8,000 ms. Error bars reflect *SEM*.

correlations between the four Trial Type × Delay conditions in phasic pupil variability and TUT rate and within the two prosaccade and antisaccade accuracy conditions. Also replicating Experiment 1, phasic pupil variability again correlated negatively with saccade accuracy (see rows 5–8, columns 1–4) and positively with TUT rate (see rows 9–12, columns 5–8).

Discussion

The cue-evoked phasic pupil change data from Experiment 2 replicated the pattern found in Experiment 1. Specifically, participants' pupil diameter decreased to a greater degree when expecting a prosaccade trial relative to an antisaccade trial. This suggests that, when preparing for a prosaccade trial, participants relax control during the delay. In contrast, when expecting an antisaccade trial, participants' pupil diameter initially dilates and then remains at baseline levels until 5,000 ms after cue onset, suggesting early preparatory control engagement. As we discussed following Experiment 1, we suspect this pattern reflects greater vigilance in preparing to make the difficult antisaccade response, combined with the difficulty of maintaining such vigilance throughout the variable delay conditions (Unsworth et al., 2018). In terms of accuracy during antisaccade trials at longer delays, the pattern suggests early engagement of control that wanes earlier preceding error trials than accurate trials. Such a pattern is similar to research using the AX-CPT task, in which there is early pupil dilation immediately after the cue that returns to or below baseline before the probe onset (Chatham et al., 2009; Chiew & Braver, 2013, nonincentive condition).

Of particular interest, the switch in thought probe procedure from Experiment 1 to Experiment 2 successfully reduced reactive responding, such that participants now reported more TUTs on prosaccade than antisaccade trials. This pattern not only replicates task difficulty patterns found in previous studies (see Smallwood & Schooler, 2006, for discussion), but is also consistent with participants' cue-evoked phasic pupil change results from both experiments. In addition, overall TUT rates marginally increased in Experiment 2 (M = 46%) relative to Experiment 1 (M = 39%, $t[211] = 1.88$, $p = .061$ for the difference in TUTs across experiments), which could reflect the reduced frequency of thought probes (reducing reminders to stay on-task), the immediate presentation of thought probes (improving introspection and reducing reactivity), or both.

Given this physiological validation of self-reported mind wandering, we felt comfortable using participants' thought probe responses to separate phasic cue-evoked pupil responses when on-task versus off-task. We found phasic pupil changes as a function of expected trial type when people were on-task, but not when mind wandering (see Figure 7). This finding is consistent with Smallwood et al.'s (2011) decoupling hypothesis and Unsworth and Robison's (2017a) finding that off-task thoughts are associated with reduced phasic pupil responses. Thus, when on-task, the trial

Table 2
Correlations Between Saccade Accuracy, Variability (SD) in Phasic Pupil Response, and Task Unrelated Thoughts (TUTs) at the 4,000 and 8,000 ms Conditions in Experiment 2

Experiment 1	1	2	3	4	5	6	7	8	9	10	11	12
1. Pro, 4,000 ms accuracy	—											
2. Pro, 8,000 ms accuracy	.742*	—										
3. Anti, 4,000 ms accuracy	.164	.172	—									
4. Anti, 8,000 ms accuracy	.150	.191	.330*	—								
5. Pro, 4,000 ms pupil *SD*	−.242*	−.117	−.103	−.162	—							
6. Pro, 8,000 ms pupil *SD*	−.187	−.054	−.057	−.213*	.811*	—						
7. Anti, 4,000 ms pupil *SD*	−.300*	−.140	−.135	−.212*	.834*	.806*	—					
8. Anti, 8,000 ms pupil *SD*	−.178	−.089	−.057	−.151	.821*	.784*	.793*	—				
9. Pro, 4,000 ms, TUT	−.114	−.020	.049	.130	.098	.113	.171	.170	—			
10. Pro, 8,000 ms, TUT	−.109	−.032	.126	.161	.271*	.243*	.265*	.331*	.651*	—		
11. Anti, 4,000 ms, TUT	−.214*	−.136	−.060	.056	.160	.034	.158	.151	.631*	.626*	—	
12. Anti, 8,000 ms, TUT	−.148	−.090	.100	.078	.142	.075	.266*	.223*	.548*	.640*	.686*	—

* $p < .05$ (two-tailed).

type differences in phasic pupil changes reflect differences in the degree of preparatory control exerted for an upcoming trial. However, when attention is focused on internal thoughts, LC-NE activity, as measured by phasic pupil changes, is decoupled from the current task. Further, the finding that pupil changes when preparing for a prosaccade stimulus mimicked pupil changes during mind wandering suggests that preparing for the prosaccade task involves purposeful relaxation of control in allowing oneself to be "captured" by the exogenous saccade stimulus, thereby transitioning control in favor of habitual responses. Thus, in the current paradigm, different cue-evoked phasic pupil change patterns across trial types do not simply reflect task-engagement versus disengagement per se. Rather; they represent the degree of preparatory effort required if engaged in the task, such that on-task preparation for prosaccade trials involves purposeful relaxation of control.

As with Experiment 1, and in line with current research using other tasks (e.g., Unsworth & Robison, 2017a), individuals with greater fluctuations in attentional preparation, as measured through variability in phasic pupil changes, had more errors and reported more mind wandering. These patterns are consistent with Unsworth and Robison's (2017b) LC-NE account of individual differences in attention control. Specifically, dysregulation of LC-NE functioning can cause attentional lapses of control state in which frontal-parietal regions receive less activation and default mode regions are not inhibited. This, in turn, causes increased mind wandering and decreased task performance. By extension, the patterns of the current Experiments further support the view that fluctuations in pupil dilation, as opposed to overall pupil size, relates to moment-by-moment lapses in attention (Unsworth & Robison, 2017a).

An unexpected finding from the current study was that pupil diameters during the first 30 ms of fixation (our baseline measure) were slightly higher (0.016 mm) when participants were cued with a prosaccade than an antisaccade. This same pattern also occurred in Experiment 1, but did not reach significance. Given the equal luminance in the 1,500 ms prosaccade and antisaccade cue screens, and the equal subsequent 1,000 ms ISI screen, it is not clear why this would be. One possibility is that participants use the prosaccade cue to disengage from the current trial by shifting to a tonic control state of exploration characterized by higher baseline pupils and reduced phasic pupil responses (Gilzenrat et al., 2010). However, in a post hoc analysis, we found that participants' pupil diameter (absolute or standardized) during the first 30 ms of fixation did not predict their saccade accuracy and did not correlate with their degree of phasic pupil change, variability in pupil change, or mind wandering (all $ps > .10$), so this initial baseline difference might be spurious.

General Discussion

In the current studies, we investigated participants' task set preparation by measuring changes in pupil diameter as they prepared for an easy (prosaccade) or difficult (antisaccade) trial. We also included thought probes to measure whether participants were on-task or mind wandering. Consistent with Wang et al. (2015), cue-evoked phasic changes in pupil diameter were more positive when anticipating an antisaccade versus a prosaccade trial. We predicted this pattern because antisaccade trials require more cognitive effort (or controlled processing) to execute, as task-evoked pupil changes are a sensitive measure of task difficulty and load, as well as mental effort (Beatty, 1982; Heitz et al., 2008; Hess & Polt, 1964; Kahneman, 1973; Peavler, 1974) and cognitive control (Rondeel et al., 2015). These findings are also in line with previous studies showing that pupil diameter is greater when participants engage preparatory control or are processing difficult stimuli (Chatham et al., 2009; Chiew & Braver, 2013; Just & Carpenter, 1993). Participants prepare for the more difficult antisaccade trials by increasing their level of cognitive engagement, while conversely preparing for the easy prosaccade by allowing their engagement to lapse.

These findings are generally consistent with the Adaptive Gain Theory (Aston-Jones & Cohen, 2005; Aston-Jones et al., 2007) of LC-NE function. The Adaptive Gain theory posits that perceived task utility governs the LC-NE system. When effortful responding is likely to bring reward (i.e., high utility), the system favors task exploitation through lower baseline LC activity and increased task-evoked phasic NE release to increase the gain in processing task-relevant stimuli. In contrast, when effortful responding is not needed and/or unlikely to bring reward (i.e., low utility), the system favors exploration of other activities through higher baseline LC activity and little-to-no phasic task-evoked response, reflecting disengagement from the current task and enhanced processing of task-unrelated stimuli. Because the LC releases NE throughout the neocortex, such activity mobilizes the frontal-parietal executive network to enhance task-relevant processing while reducing external and internal distractions (Unsworth & Robison, 2017b).

We believe the current study supports and ties together several existing theories regarding LC-NE function, task performance, and mind wandering. Although we did not have a tonic baseline measure of pupil diameter before each trial, our phasic cue-evoked pupil changes showed initial dilation before antisaccade responses (especially accurate responses) and constriction preceding prosaccade responses. Moreover, consistent with Smallwood et al.'s (2011) decoupling hypothesis, this expected trial difference in phasic pupil changes only occurred when participants reported being on-task. Finally, consistent with Unsworth and Robison's (2017a) latent variable analysis and attentional lapse account of individual differences in sustained attention, we found that individuals with greater variability in phasic pupil responses were more likely to produce executive errors in the saccade task and more likely to report mind wandering. We believe our study is the first to test all these accounts simultaneously and is the first to demonstrate that (a) cue-evoked phasic pupil changes accurately reflect task set preparation only when participants are on-task and (b) on-task phasic changes during purposeful preparation for a prosaccade trial resemble phasic changes that occur during mind wandering.

Recent findings indicate that greater delays between the fixation and antisaccade stimulus increase the relation between working memory capacity and antisaccade accuracy (Meier et al., 2018; Moffitt, 2013). Specifically, whereas higher working memory capacity individuals use the delay period to prepare target-appropriate responding (Brown et al., 2007), thereby using the increment of time to reach full engagement (or instantiation) of the task goal, lower working memory capacity individuals do not activate the goal as strongly. This finding is consistent with studies showing impairments among low working memory capacity individuals in engaging preparatory control (Braver et al., 2007; Rich-

mond, Redick, & Braver, 2015), as well as in LC-NE regulation (Unsworth & Robison, 2017b) and frontal function (Kane & Engle, 2002). Our current Trial Type × Delay interaction on saccade accuracy indicates that accuracy indeed improved across delay in our studies. However, although we have suggestive evidence that higher working memory capacity individuals within our sample drove this increase in accuracy across delay, our results are not conclusive (see Footnote 2).

An interesting finding in our study is that self-reported mind wandering depended upon when the thought-probe occurred in the task. When the thought probe occurred immediately after the target detection response (Experiment 1), participants reported more mind wandering on antisaccade trials. We believe participants reported being off-task to justify their errors in performance on these harder antisaccade trials. Consistent with this, when the thought probe instead occurred in lieu of the stimulus (Experiment 2), we found the opposite, yet predicted pattern, in that participants reported more mind wandering on prosaccade trials. This finding not only replicates task difficulty patterns found in previous studies (see Smallwood & Schooler, 2006, for discussion), but is also consistent with participants' cue-evoked phasic pupil response from both experiments. Although thought probes are a common method of measuring mind wandering, as mentioned previously, accurately responding to thought probes requires a certain degree of introspection, confidence in that introspection, and honesty in reporting. The current results demonstrate that pupillometry cannot only provide a psychophysiological marker of cognitive effort, but also help validate self-reported mind wandering.

One setback to the current study is that, because of technical issues with the eye tracker, we had to remove data from 8 and 17 participants from Experiments 1 and 2, respectively. Furthermore, in examining pupil diameter separately for trials in which participants reported being on-task versus off-task, we only included data from participants who reported both TUT and TRT responses. This resulted in the removal of a large amount of data. Future studies could include more participants to gain more power and determine whether any marginal results from the current study would become significant. However, we note that our power was sufficient to obtain the three-way interaction between task, probe response, and delay in the 8,000 ms dataset.

In conclusion, the results of this experiment demonstrate that cue-evoked phasic pupil changes accurately reflect the predicted levels of cognitive engagement because of variations in expected task difficulty and delay. Additionally, when participants report being on-task, they selectively regulate effortful control based on expected task demands. Together, these results bridge current theories regarding LC-NE function, task performance, and mind wandering.

References

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. http://dx.doi.org/10.1146/annurev.neuro.28.061604.135709

Aston-Jones, G. S., Iba, M., Clayton, E., Rajkowski, J., & Cohen, J. (2007). The locus coeruleus and regulation of behavioral flexibility and attention: Clinical implications. In G. A. Ordway, M. A. Schwartz, & A. Frazer (Eds.), *Brain norepinephrine: Neurobiology and therapeutics* (pp. 196–235). Cambridge, UK: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511544156.008

Baird, B., Smallwood, J., Mrazek, M. D., Kam, J. W. Y., Franklin, M. S., & Schooler, J. W. (2012). Inspired by distraction: Mind wandering facilitates creative incubation. *Psychological Science, 23,* 1117–1122. http://dx.doi.org/10.1177/0956797612446024

Baird, B., Smallwood, J., & Schooler, J. W. (2011). Back to the future: Autobiographical planning and the functionality of mind-wandering. *Consciousness and Cognition, 20,* 1604–1611. http://dx.doi.org/10.1016/j.concog.2011.08.007

Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological Bulletin, 91,* 276–292. http://dx.doi.org/10.1037/0033-2909.91.2.276

Beatty, J., & Kahneman, D. (1966). Pupillary changes in two memory tasks. *Psychonomic Science, 5,* 371–372.

Bradshaw, J. L. (1968). Pupil size and problem solving. *The Quarterly Journal of Experimental Psychology, 20,* 116–122. http://dx.doi.org/10.1080/14640746808400139

Bradshaw, J. L. (1969). Background light intensity and the pupillary response in a reaction time task. *Psychonomic Science, 14,* 271–272.

Braver, T. S., Gray, J. R., & Burgess, G. C. (2007). Explaining the many varieties of working memory variation: Dual mechanisms of cognitive control. In A. Conway, C. Jarrold, M. J. Kane, A. Miyake, & J. N. Towse (Eds.), *Variation in working memory* (pp. 76–106). New York, NY: Oxford University Press.

Brown, M. R. G., Vilis, T., & Everling, S. (2007). Frontoparietal activation with preparation for antisaccades. *Journal of Neurophysiology, 98,* 1751–1762. http://dx.doi.org/10.1152/jn.00460.2007

Chatham, C. H., Frank, M. J., & Munakata, Y. (2009). Pupillometric and behavioral markers of a developmental shift in the temporal dynamics of cognitive control. *Proceedings of the National Academy of Sciences of the United States of America, 106,* 5529–5533. http://dx.doi.org/10.1073/pnas.0810002106

Chiew, K. S., & Braver, T. S. (2013). Temporal dynamics of motivation-cognitive control interactions revealed by high-resolution pupillometry. *Frontiers in Psychology, 4,* 15. http://dx.doi.org/10.3389/fpsyg.2013.00015

Chin, J., Mrazek, M., & Schooler, J. (2012). Blind spots to the self. In S. Vazire & T. D. Wilson (Eds.), *Handbook of self-knowledge* (p. 79). New York, NY: Guilford Press.

Engle, R. W. (2001). What is working memory capacity? In H. L. Roediger, III, J. S. Nairne, I. Neath, & A. M. Surprenant (Eds.), *Science conference series. The nature of remembering: Essays in honor of Robert G. Crowder* (pp. 297–314). Washington, DC: American Psychological Association. http://dx.doi.org/10.1037/10394-016

Foster, J. L., Shipstead, Z., Harrison, T. L., Hicks, K. L., Redick, T. S., & Engle, R. W. (2015). Shortened complex span tasks can reliably measure working memory capacity. *Memory & Cognition, 43,* 226–236. http://dx.doi.org/10.3758/s13421-014-0461-7

Franklin, M. S., Broadway, J. M., Mrazek, M. D., Smallwood, J., & Schooler, J. W. (2013). Window to the wandering mind: Pupillometry of spontaneous thought while reading. *Quarterly Journal of Experimental Psychology, 66,* 2289–2294. http://dx.doi.org/10.1080/17470218.2013.858170

Giambra, L. M. (1995). A laboratory method for investigating influences on switching attention to task-unrelated imagery and thought. *Consciousness and Cognition, 4,* 1–21. http://dx.doi.org/10.1006/ccog.1995.1001

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. http://dx.doi.org/10.3758/CABN.10.2.252

Heitz, R. P., Schrock, J. C., Payne, T. W., & Engle, R. W. (2008). Effects of incentive on working memory capacity: Behavioral and pupillometric data. *Psychophysiology, 45,* 119–129.

Hess, E. H., & Polt, J. M. (1964). Pupil size in relation to mental activity during simple problem-solving. *Science, 143,* 1190–1192. http://dx.doi.org/10.1126/science.143.3611.1190

Hutchison, K. A. (2007). Attentional control and the relatedness proportion effect in semantic priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33,* 645–662. http://dx.doi.org/10.1037/0278-7393.33.4.645

Hutchison, K. A., Heap, S. J., Neely, J. H., & Thomas, M. A. (2014). Attentional control and asymmetric associative priming. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40,* 844–856. http://dx.doi.org/10.1037/a0035781

Irons, J. L., Jeon, M., & Leber, A. B. (2017). Pre-stimulus pupil dilation and the preparatory control of attention. *PLoS ONE, 12,* e0188787. http://dx.doi.org/10.1371/journal.pone.0188787

Jennings, J. R., van der Molen, M. W., & Steinhauer, S. R. (1998). Preparing the heart, eye, and brain: Foreperiod length effects in a nonaging paradigm. *Psychophysiology, 35,* 90–98. http://dx.doi.org/10.1111/1469-8986.3510090

Just, M. A., & Carpenter, P. A. (1993). The intensity dimension of thought: Pupillometric indices of sentence processing. *Canadian Journal of Experimental Psychology, 47,* 310–339. http://dx.doi.org/10.1037/h0078820

Kahneman, D. (1973). *Attention and effort*. Upper Saddle River, NJ: Prentice Hall.

Kahneman, D., & Beatty, J. (1966). Pupil diameter and load on memory. *Science, 154,* 1583–1585. http://dx.doi.org/10.1126/science.154.3756.1583

Kane, M. J., Bleckley, M. K., Conway, A. R. A., & Engle, R. W. (2001). A controlled-attention view of working memory capacity. *Journal of Experimental Psychology: General, 130,* 169–183.

Kane, M. J., & Engle, R. W. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual-differences perspective. *Psychonomic Bulletin & Review, 9,* 637–671. http://dx.doi.org/10.3758/BF03196323

Kristjansson, S. D., Stern, J. A., Brown, T. B., & Rohrbaugh, J. W. (2009). Detecting phasic lapses in alterness using pupillometric measures. *Applied Ergonomics, 40,* 978–986.

Mazaheri, A., DiQuattro, N. E., Bengson, J., & Geng, J. J. (2011). Pre-stimulus activity predicts the winner of top-down vs. bottom-up attentional selection. *PLoS ONE, 6,* e16243. http://dx.doi.org/10.1371/journal.pone.0016243

McVay, J. C., & Kane, M. J. (2009). Conducting the train of thought: Working memory capacity, goal neglect, and mind wandering in an executive-control task. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35,* 196–204. http://dx.doi.org/10.1037/a0014104

McVay, J. C., Meier, M. E., Touron, D. R., & Kane, M. J. (2013). Aging ebbs the flow of thought: Adult age differences in mind wandering, executive control, and self-evaluation. *Acta Psychologica, 142,* 136–147. http://dx.doi.org/10.1016/j.actpsy.2012.11.006

Meier, M. E., Smeekens, B. A., Silvia, P. J., Kwapil, T. R., & Kane, M. J. (2018). Working memory capacity and the antisaccade task: A microanalytic-macroanalytic investigation of individual differences in goal activation and maintenance. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 44,* 68–84. http://dx.doi.org/10.1037/xlm0000431

Moffitt, C. C. (2013). *Working memory capacity and saccade performance across fixation delay: Attentional preparation or goal neglect?* (Master's Thesis). Retrieved from http://scholarworks.montana.edu/xmlui/handle/1/2685

Mooneyham, B. W., & Schooler, J. W. (2013). The costs and benefits of mind-wandering: A review. *Canadian Journal of Experimental Psychology, 67,* 11–18. http://dx.doi.org/10.1037/a0031569

Mrazek, M. D., Phillips, D. T., Franklin, M. S., Broadway, J. M., & Schooler, J. W. (2013). Young and restless: Validation of the Mind-Wandering Questionnaire (MWQ) reveals disruptive impact of mind-wandering for youth. *Frontiers in Psychology, 4,* 560. http://dx.doi.org/10.3389/fpsyg.2013.00560

Mueller, S. C., Swainson, R., & Jackson, G. M. (2009). ERP indices of persisting and current inhibitory control: A study of saccadic task switching. *NeuroImage, 45,* 191–197. http://dx.doi.org/10.1016/j.neuroimage.2008.11.019

Norman, D. A., & Shallice, T. (1986). Attention to action: Willed and automatic control of behavior. In R. J. Davidson, G. E. Schwartz, & D. Shapiro (Eds.), *Consciousness and self-regulation: Advances in research and theory* (Vol. 4, pp. 1–18). New York, NY: Plenum.

Peavler, W. S. (1974). Pupil size, information overload, and performance differences. *Psychophysiology, 11,* 559–566. http://dx.doi.org/10.1111/j.1469-8986.1974.tb01114.x

Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience, 13,* 25–42. http://dx.doi.org/10.1146/annurev.ne.13.030190.000325

Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., & Shulman, G. L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences of the United States of America, 98,* 676–682. http://dx.doi.org/10.1073/pnas.98.2.676

Rajkowski, J., Kubiak, P., & Aston-Jones, G. (1993). Correlations between locus coeruleus (LC) neural activity, pupil diameter and behavior in monkey support a role of LC in attention. *Society for Neuroscience Abstracts, 19,* 974.

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. *Nature Communications, 7,* 13289. http://dx.doi.org/10.1038/ncomms13289

Richer, F., & Beatty, J. (1987). Contrasting effects of response uncertainty on the task-evoked pupillary response and reaction time. *Psychophysiology, 24,* 258–262.

Richer, F., Silverman, C., & Beatty, J. (1983). Response selection, and initiation in speeded reactions: A pupillometric analysis. *Journal of Experimental Psychology: Human Perception and Performance, 9,* 360–370.

Richmond, L. L., Redick, T. S., & Braver, T. S. (2015). Remembering to prepare: The benefits (and costs) of high working memory capacity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41,* 1764–1777. http://dx.doi.org/10.1037/xlm0000122

Rondeel, E. W., van Steenbergen, H., Holland, R. W., & van Knippenberg, A. (2015). A closer look at cognitive control: Differences in resource allocation during updating, inhibition and switching as revealed by pupillometry. *Frontiers in Human Neuroscience, 9,* 494. http://dx.doi.org/10.3389/fnhum.2015.00494

Schooler, J. W. (2001). Discovering memories of abuse in the light of meta-awareness. *Journal of Aggression, Maltreatment & Trauma, 4,* 105–136. http://dx.doi.org/10.1300/J146v04n02_06

Schooler, J. W. (2002). Re-representing consciousness: Dissociations between experience and meta-consciousness. *Trends in Cognitive Sciences, 6,* 339–344. http://dx.doi.org/10.1016/S1364-6613(02)01949-6

Schooler, J. W., Reichle, E. D., & Halpern, D. V. (2005). Zoning out while reading: Evidence for dissociations between experience and metaconsciousness. In D. Levin (Ed.), *Thinking and seeing: Visual metacognition in adults and children* (pp. 203–226). Cambridge, MA: MIT Press.

Seli, P., Cheyne, J. A., Xu, M., Purdon, C., & Smilek, D. (2015). Motivation, intentionality, and mind wandering: Implications for assessments of task-unrelated thought. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 41,* 1417–1425. http://dx.doi.org/10.1037/xlm0000116

Seli, P., Risko, E. F., Smilek, D., & Schacter, D. L. (2016). Mind-wandering with and without intention. *Trends in Cognitive Sciences, 20,* 605–617. http://dx.doi.org/10.1016/j.tics.2016.05.010

Shipstead, Z., Lindsey, D. R. B., Marshall, R. L., & Engle, R. W. (2014). The mechanisms of working memory capacity: Primary memory, sec-

ondary memory, and attention control. *Journal of Memory and Language, 72,* 116–141. http://dx.doi.org/10.1016/j.jml.2014.01.004

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

Smallwood, J., Beach, E., Schooler, J. W., & Handy, T. C. (2008). Going AWOL in the brain: Mind wandering reduces cortical analysis of external events. *Journal of Cognitive Neuroscience, 20,* 458–469. http://dx.doi.org/10.1162/jocn.2008.20037

Smallwood, J., Brown, K. S., Tipper, C., Giesbrecht, B., Franklin, M. S., Mrazek, M. D., . . . Schooler, J. W. (2011). Pupillometric evidence for the decoupling of attention from perceptual input during offline thought. *PLoS ONE, 6,* e18298. http://dx.doi.org/10.1371/journal.pone.0018298

Smallwood, J., McSpadden, M., & Schooler, J. W. (2008). When attention matters: The curious incident of the wandering mind. *Memory & Cognition, 36,* 1144–1150. http://dx.doi.org/10.3758/MC.36.6.1144

Smallwood, J., Nind, L., & O'Connor, R. C. (2009). When is your head at? An exploration of the factors associated with the temporal focus of the wandering mind. *Consciousness and Cognition, 18,* 118–125. http://dx.doi.org/10.1016/j.concog.2008.11.004

Smallwood, J., & Schooler, J. W. (2006). The restless mind. *Psychological Bulletin, 132,* 946–958. http://dx.doi.org/10.1037/0033-2909.132.6.946

Unsworth, N., Fukuda, K., Awh, E., & Vogel, E. K. (2014). Working memory and fluid intelligence: Capacity, attention control, and secondary memory retrieval. *Cognitive Psychology, 71,* 1–26.

Unsworth, N., & Robison, M. K. (2015). Individual differences in the allocation of attention to items in working memory: Evidence from pupillometry. *Psychonomic Bulletin & Review, 22,* 757–765. http://dx.doi.org/10.3758/s13423-014-0747-6

Unsworth, N., & Robison, M. K. (2016). Pupillary correlates of lapses of sustained attention. *Cognitive, Affective & Behavioral Neuroscience, 16,* 601–615. http://dx.doi.org/10.3758/s13415-016-0417-4

Unsworth, N., & Robison, M. K. (2017a). The importance of arousal for variation in working memory capacity and attention control: A latent variable pupillometry study. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 43,* 1962–1987. http://dx.doi.org/10.1037/xlm0000421

Unsworth, N., & Robison, M. K. (2017b). A locus coeruleus-norepinephrine account of individual differences in working memory capacity and attention control. *Psychonomic Bulletin & Review, 24,* 1282–1311. http://dx.doi.org/10.3758/s13423-016-1220-5

Unsworth, N., Robison, M. K., & Miller, A. L. (2018). Pupillary correlates of fluctuations in sustained attention. *Journal of Cognitive Neuroscience, 30,* 1241–1253. http://dx.doi.org/10.1162/jocn_a_01251

Unsworth, N., Schrock, J., & Engle, R. (2004). Working memory capacity and the antisaccade task: Individual differences in voluntary saccade control. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 1302–1321.

Usher, M., Cohen, J. D., Servan-Schreiber, D., Rajkowski, J., & Aston-Jones, G. (1999). The role of locus coeruleus in the regulation of cognitive performance. *Science, 283,* 549–554.

van den Brink, R. L., Murphy, P. R., & Nieuwenhuis, S. (2016). Pupil diameter tracks lapses of attention. *PLoS ONE, 11*(10), e0165274. http://dx.doi.org/10.1371/journal.pone.0165274

van der Meer, E., Beyer, R., Horn, J., Foth, M., Bornemann, B., Ries, J., . . . Wartenburger, I. (2010). Resource allocation and fluid intelligence: Insights from pupillometry. *Psychophysiology, 47,* 158–169. http://dx.doi.org/10.1111/j.1469-8986.2009.00884.x

van der Molen, M. W., Boomsma, D. I., Jennings, J. R., & Nieuwboer, R. T. (1989). Does the heart know what the eye sees? A cardiac/pupillometric analysis of motor preparation and response execution. *Psychophysiology, 26,* 70–80.

Wang, C. A., Brien, D. C., & Munoz, D. P. (2015). Pupil size reveals preparatory processes in the generation of pro-saccades and anti-saccades. *European Journal of Neuroscience, 41,* 1102–1110. http://dx.doi.org/10.1111/ejn.12883

Received October 1, 2018
Revision received February 11, 2019
Accepted March 22, 2019 ■