



LJMU Research Online

Eayrs, JO, Tobing, HS, Steendam, ST, Prutean, N, Notebaert, W, Wiersema, JR, Krebs, RM and Boehler, CN

Reward and Efficacy Modulate the Rate of Anticipatory Pupil Dilation

<http://researchonline.ljmu.ac.uk/id/eprint/25832/>

Article

Citation (please note it is advisable to refer to the publisher's version if you intend to cite from this work)

Eayrs, JO, Tobing, HS, Steendam, ST, Prutean, N, Notebaert, W, Wiersema, JR, Krebs, RM and Boehler, CN (2025) Reward and Efficacy Modulate the Rate of Anticipatory Pupil Dilation. *Psychophysiology*, 62 (1). ISSN 0048-5772

LJMU has developed **LJMU Research Online** for users to access the research output of the University more effectively. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Users may download and/or print one copy of any article(s) in LJMU Research Online to facilitate their private study or for non-commercial research. You may not engage in further distribution of the material or use it for any profit-making activities or any commercial gain.

The version presented here may differ from the published version or from the version of the record. Please see the repository URL above for details on accessing the published version and note that access may require a subscription.

For more information please contact researchonline@ljmu.ac.uk

<http://researchonline.ljmu.ac.uk/>

ORIGINAL ARTICLE OPEN ACCESS

Reward and Efficacy Modulate the Rate of Anticipatory Pupil Dilation

Joshua O. Eayrs^{1,2}  | Haya Serena Tobing³  | S. Tabitha Steendam²  | Nicoleta Prutean²  | Wim Notebaert²  | Jan R. Wiersema⁴  | Ruth M. Krebs²  | C. Nico Boehler² 

¹School of Psychology, Liverpool John Moores University, Liverpool, UK | ²Department of Experimental Psychology, Ghent University, Ghent, Belgium | ³Faculty of Psychology and Educational Sciences, KU Leuven, Leuven, Belgium | ⁴Department of Experimental Clinical and Health Psychology, Ghent University, Ghent, Belgium

Correspondence: Joshua O. Eayrs (j.o.eayrs@ljmu.ac.uk)

Received: 16 September 2024 | **Revised:** 17 December 2024 | **Accepted:** 20 December 2024

Funding: This work was supported by Universiteit Gent (BOF20-GOA-004).

Keywords: efficacy | effort | preparatory control | pupillometry | reward

ABSTRACT

Pupil size is a well-established marker of cognitive effort, with greater efforts leading to larger pupils. This is particularly true for pupil size during task performance, whereas findings on anticipatory effort triggered by a cue stimulus are less consistent. For example, a recent report by Frömer et al. found that in a cued-Stroop task, behavioral performance and electrophysiological markers of preparatory effort allocation were modulated by cued reward and ‘efficacy’ (the degree to which rewards depended on good performance), but pupil size did not show a comparable pattern. Here, we conceptually replicated this study, employing an alternative approach to the pupillometry analyses. In line with previous findings, we found no modulation of absolute pupil size in the cue-to-target interval. Instead, we observed a significant difference in the rate of pupil dilation in anticipation of the target: pupils dilated more rapidly for high-reward trials in which rewards depended on good performance. This was followed by a significant difference in absolute pupil size within the first hundreds of milliseconds following Stroop stimulus onset, likely reflecting a lagging effect of anticipatory effort allocation. Finally, the slope of pupil dilation was significantly correlated with behavioral response times, and this association was strongest for the high-reward, high-efficacy trials, further supporting that the rate of anticipatory pupil dilation reflects anticipatory effort. We conclude that pupil size is modulated by anticipatory effort, but in a highly temporally-specific manner, which is best reflected by the rate of dilation in the moments just prior to stimulus onset.

Cognitive control is necessary for the performance of almost any task, both in the lab and in daily life, but exercising control is effortful and therefore typically avoided unless adequate incentive is available (Westbrook and Braver 2015). According to the Expected Value of Control model (Shenhav, Botvinick, and Cohen 2013; Shenhav, Cohen, and Botvinick 2016), the intensity of effortful control allocated to a task depends on a computation of the relative reward associated with successful performance and the effort necessary to obtain it. Functional imaging studies established that computation of the expected value of control is reflected in the activity of the dopaminergic

midbrain and anterior cingulate cortex (ACC), and the resulting allocation of effort is associated with response preparation and activation of attentional control structures (e.g., Krebs et al. 2012; Shenhav, Botvinick, and Cohen 2013; Shenhav, Cohen, and Botvinick 2016). Electrophysiological studies have demonstrated that cues indicating reward prospects evoke significantly larger P3 ERP components, the amplitude of which scales with reward value. Later, approaching the target stimulus onset, reward and anticipated task demands interact to determine neural indices of task preparation such as the contingent negative variation (CNV) and posterior parietal

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2025 The Author(s). *Psychophysiology* published by Wiley Periodicals LLC on behalf of Society for Psychophysiological Research.

alpha power (Goldstein et al. 2006; Kostandyan et al. 2019; Schevernels et al. 2014; Van Den Berg et al. 2014).

Pupil dilation provides a sensitive measure of the level of cognitive effort invested in a task in the sense that more difficult tasks lead to larger dilations, provided participants engage accordingly (Hess and Polt 1964; Kahneman and Beatty 1966; Van Der Wel and Van Steenbergen 2018). In tasks that involve cognitive conflict, such as the Stroop or flanker tasks (Eriksen and Eriksen 1974; Stroop 1935), pupil dilations to conflicting stimuli are reliably larger than those for non-conflicting stimuli, reflecting the increased need for effortful control (Rondeel et al. 2015; Van Der Wel and Van Steenbergen 2018). Furthermore, individual differences in pupil responses to conflict predict behavioral outcomes—individuals with greater stimulus-evoked pupil dilations respond faster, make fewer errors, and have smaller conflict-related performance costs (Rondeel et al. 2015; Unsworth and Miller 2023). Physiologically, pupil size is tightly linked to the release of norepinephrine from the Locus Coeruleus (LC), which receives substantial input from the ACC. The LC has widespread efferent connections throughout the cortex and is thought to modulate neural gain—increasing excitation sensitivity to task-stimulus-relevant neurons (Alnaes et al. 2014; Aston-Jones and Cohen 2005; Ferguson and Cardin 2020; Gilzenrat et al. 2010; Murphy et al. 2014; Reimer et al. 2016; Unsworth and Robison 2017).

While stimulus-evoked pupil responses reliably index the allocation of cognitive effort to a task, the modulation of the pupil in anticipation of a task is less straightforward. Cognitive control can be allocated to a task both reactively (e.g., to resolve attentional conflicts arising from an unexpected distractor stimulus) but also proactively to minimize the need for reactive adjustments (e.g., by adopting an attentional set which already excludes potential distractors; Braver 2012). Accordingly, pupil size has been shown to increase in anticipation of target stimuli following cues indicating that a reward can be earned for good performance. This effect has been observed in various experimental paradigms, such as visual discrimination (Boehler et al. 2011), pro/antisaccade tasks (Wang, Brien, and Munoz 2015; Wang et al. 2016), and simple reaction time tasks (Schneider et al. 2018). However, in contrast, recent investigations using more complex cognitive paradigms have found little or no evidence for reward-motivated preparatory effort allocation reflected in pupil size. Kostandyan et al. (2019) used both pupillometry and EEG to investigate the effects of reward on allocation of control to a flanker task. Participants completed an arrow flanker task (Eriksen and Eriksen 1974) in which they could earn rewards for fast and accurate responses. Similar to previous experiments, the cue-evoked CNV component in anticipation of the target stimulus was significantly larger, and participants' responses were significantly faster for rewarded than unrewarded trials. Furthermore, pupil dilation evoked by the target stimulus was significantly larger for rewarded than unrewarded trials. However, there was no evidence that pupil size was differentially modulated for reward trials prior to the target onset. Although, intriguingly, the reward effect on the target-evoked pupil response emerged almost instantly after target onset, several hundreds of milliseconds earlier than the effect of flanker congruence. This was taken to suggest that reward did

modulate anticipatory pupil size, but in a way that was tightly linked with the temporal onset of the target stimulus.

In a recent extension of this line of work directly testing the Expected Value of Control model, Frömer et al. (2021) found that participants' proactive effort allocation depended not only on the level of reward but also on the 'performance efficacy' they anticipated in the task. In their experiments, participants were first presented with a cue indicating the monetary reward (relatively large or small), which could be earned on that trial as well as whether that reward would be purely performance-contingent (i.e., a correct and fast response yielded a reward) or if it would be random (i.e., not dependent upon performance). They found that cued reward and efficacy interacted to determine participants' performance as well as replicating the effects of reward on the P3 and CNV components reported previously (with predicted efficacy behaving similarly to predicted task difficulty): CNV amplitude was greatest for large-reward trials with high efficacy, suggesting that participants allocated more effort to the task when they expected this to be more instrumental in earning the reward. However, similar to Kostandyan et al. (2019) pupil dilation following the cue was not significantly modulated by reward and was in fact significantly smaller for high-efficacy trials. The authors speculated that in their paradigm, pupil size effects reflected increased levels of arousal due to uncertainty induced by the low-efficacy condition, rather than preparatory effort. At the same time, Frömer et al. (2021) fitted a canonical pupil response function that peaks in the cue-target interval, which may not be ideal for capturing preparatory processes that gear up toward target presentation.

For example, Unsworth and Miller (2023) recently reported that pupil size 'ramps up' prior to the onset of the stimulus in a Stroop paradigm. That is, pupil size gradually increased over the last few hundreds of milliseconds before the stimulus appeared, and the authors suggest that the *rate* of pupil dilation (i.e., the speed at which the pupil increases in size) may be a better index of the allocation of preparatory effort than absolute pupil size. Furthermore, they found that individual differences in working memory capacity, RT on incongruent Stroop trials and behavioral Stroop congruency effects were all correlated with pupil size slope: Participants with larger working memory capacity, faster RTs and smaller Stroop congruency effects all had relatively steeper increases in pupil size. In a second experiment, they showed that the 'ramping' was also steeper when participants were cued to anticipate a conflicting stimulus than when they expected a congruent stimulus. The authors concluded that individuals with greater capacity for cognitive control (as indexed by working memory capacity; see e.g., Eayrs and Lavie 2018, 2019; Engle and Kane 2003) are better able to prepare for conflict, and this preparation is indexed by the rate of pre-stimulus pupil dilation. While Unsworth and Miller (2023) did not include any incentive or reward manipulation, another study from the same group (Unsworth, Miller, and Aghel 2022) observed similar increases in pre-stimulus pupil dilation rates for a simple psychomotor vigilance task when participants were simply cued to 'try hard' on a subset of trials. Similarly, Tromp, Nieuwenhuis, and Murphy (2022) related the first temporal derivative (i.e., rate of change, or 'ramping') of anticipatory pupil size to the effects of simulated neural gain on Stroop task performance (although their task did not include pre-cues, stimulus onset was still predictable,

allowing preparatory effort allocation). Similar to Unsworth and Miller (2023), they found that steeper anticipatory pupil dilation was associated with faster responding (albeit, in their study, at the cost of increased Stroop interference), globally supporting its use as an index of anticipatory neural gain.

The lack of reward- or efficacy-modulated effects on pupil size during the cue-to-target interval in the results of Frömer et al. (2021) and, to some extent, also in Kostandyan et al. (2019) may be explained by two key factors: First, Frömer et al. (2021) used cues with significantly greater information-density, informing about both the reward and efficacy of each trial, whereas the other studies included only information about either difficulty (Boehler et al. 2011; Wang, Brien, and Munoz 2015; Wang et al. 2016) or reward (Kostandyan et al. 2019; Schneider et al. 2018) or simply provided a warning for the start of the trial (Unsworth and Miller 2023). Prior work has established that relatively long periods can be necessary for proactive allocation to be affected following reward cues in tasks involving cognitive conflict (Bugg and Smallwood 2016), possibly due to the time taken to compute the expected value of control. Relatedly, Unsworth and Miller (2023) used a variable cue-to-target delay, which may have caused participants to allocate cognitive effort across a wider time-range in order to also be ready for particularly early target trials. Second, Frömer et al. (2021) and Kostandyan et al. (2019) did not measure the ‘ramping’ of pupil size in the cue-to-target interval, but rather the absolute pupil size (or of the modeled pupil response function). According to the results of Unsworth and Miller (2023), proactive allocation of effort results in a gradually ramping pupil size rather than a transient dilation (or sustained larger pupil). This pattern is evident in several figures from previous reports, wherein pupil size can be seen to increase over the course of a cue-to-target interval (e.g., Boehler et al. 2011; Kostandyan et al. 2019; Schneider et al. 2018), sometimes only starting to diverge very near or at the onset of the target stimulus (Kostandyan et al. 2019). Indeed, in animal studies it has been shown that activity of cortical noradrenergic neurons is better characterized by the first temporal derivative of pupil size than by its absolute size (Reimer et al. 2016). Furthermore, in humans, the first temporal derivative of pupil size during pre-stimulus baseline has been shown to predict subsequent behavioral performance (Tromp, Nieuwenhuis, and Murphy 2022; Van Den Brink, Murphy, and Nieuwenhuis 2016).

Here, we sought to investigate the effects of expected reward and performance efficacy on proactive allocation of cognitive effort as indexed by both absolute pupil size before and after stimulus onset and by the ‘ramping’ of pupil size in the pre-stimulus interval. We used a paradigm modeled on that of Frömer et al. (2021) and with an analysis approach informed by the findings of Kostandyan et al. (2019) and Unsworth and Miller (2023). If, as suggested by the expected value of control theory, effort is allocated in proportion to its expected value, and if, as suggested by Unsworth and Miller (2023), this proactive effort is characterized by a gradually ‘ramping’ pupil size, then we should expect that a larger reward should result in steeper pupil size increases. Further, this effect should be modulated by expected efficacy, wherein effort is only allocated in trials where performance is critical to securing the reward. Following the results of Kostandyan et al. (2019), the effect may also be observed in absolute pupil size in the early portion of the stimulus–response, reflecting a lagging effect of this preparatory allocation.

1 | Method

1.1 | Participants

Based on the average of the three samples tested by Frömer et al. (2021), a sample of 38 participants, recruited from the Ghent University student pool, took part in the experiment in exchange for a cash payment of 12 Euros, with a maximum possible additional bonus payment of 10 Euros. Participants provided informed consent prior to taking part in the study. The research was conducted according to the ethical rules presented in the General Ethics Protocol of the Faculty of Psychology and Educational Sciences of Ghent University. One participant was excluded due to excessive blink-related artifacts, leading to more than 50% of their data being interpolated. The final sample consisted of 6 men and 31 women, with an average age of 22.14 years ($SD = 2.7$). A sensitivity analysis using MorePower6.0 (Campbell and Thompson 2012) suggested that this would provide 82% power to detect an effect size of $\eta_p^2 = 0.2$, comparable to those reported by Frömer et al. (2021) for the interaction of efficacy and reward on CNV amplitudes and RT with a standard 0.05 alpha threshold. While our primary aim was to test these between-condition effects, we also report exploratory individual differences analyses based on those of Unsworth and Miller (2023); a post hoc power analysis suggests that our sample of $n = 37$ would have approximately 46% power to detect a correlation of $r = 0.29$ as reported by Unsworth and Miller (2023).

1.2 | Stimuli and Procedure

The task was based closely on the one used by Frömer et al. (2021). On each trial, participants were first presented with a cue stimulus (Figure 1A) consisting of two parts, an outline image of a hand and of a ‘money sack’; these images could either be gray or else colored in cyan or magenta to indicate the reward and efficacy condition. Isoluminance of the cue colors was confirmed with a Minolta LS-110 luminance meter. Gray always indicated ‘low’ (reward or efficacy), whereas cyan and magenta indicated high (reward or efficacy; mapping counterbalanced between participants). The ‘hand’ and ‘sack’ component of the cue each subtended approximately 1.2×2.5 degrees of visual angle. Each was presented 1° to the left and right of the central fixation cross.

The cue remained onscreen for 2200 ms and was followed immediately by a Stroop stimulus consisting of the Dutch word for “Red” (“Rood”), “Yellow” (“Geel”), “Green” (“Groen”), or “Blue” (“Blauw”). The luminance of these colors was not controlled, but all colors appeared equally frequently in each reward and efficacy condition. The stimulus subtended between 3.4×2.5 and 4.5×2.5 degrees of visual angle and were presented at fixation. The Stroop stimulus remained on-screen for 1000 ms before being replaced by a blank screen (fixation cross only) for a further 1000 ms. Participants were instructed to respond as fast and accurately as possible immediately upon presentation of the Stroop stimulus by pressing the “d”, “f”, “j”, or “k” key, each of which corresponded to a unique ink color (mapping counterbalanced between participants). A response deadline of 1000 ms determined whether their response was considered in-time (in order to earn a reward on high-efficacy trials, if the response

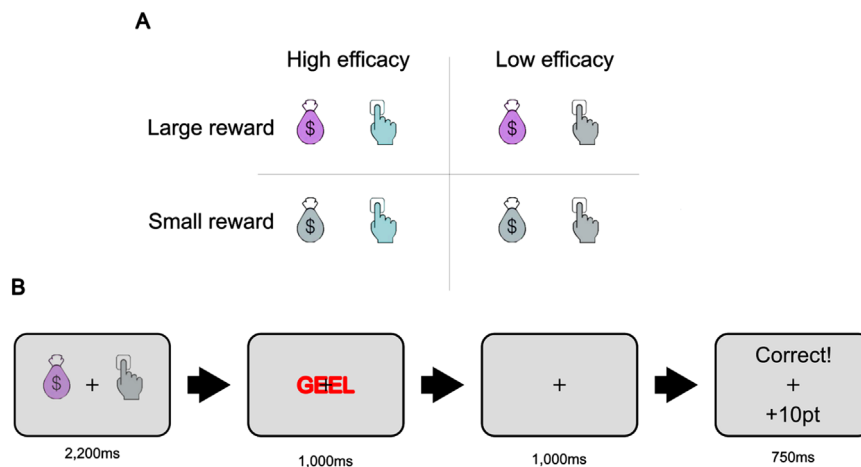


FIGURE 1 | Experiment stimuli (A) and trial structure (B). For cue stimuli (A), gray always indicated ‘low’ reward and/or efficacy, while cyan and magenta ‘hands’ and ‘bags’ indicated ‘high’ reward and/or efficacy. The side of the screen on which the hand or bag was presented was constant for a given participant but counterbalanced between participants, as was the color coding for reward and efficacy. For each trial (B) all Stroop stimuli were incongruent (different word meaning and ink color; “GEEL” is the Dutch word for “YELLOW”); an ITI of 1500–2500ms occurred between trials in which only the fixation cross was presented onscreen.

was also correct). Finally, a feedback screen was presented, consisting of text indicating the reward earned (“0.1 pt” or “10 pt” for a correct and fast response in high or low reward, respectively; “0 pt” for an incorrect response and “Too slow!” for a response after the deadline). A fixation-only ITI screen was presented between trials with a duration drawn randomly on each trial from a uniform distribution between 1500 and 2500 ms.

The reward policy on low-efficacy trials was the same as that used by Frömer et al. (2021): Whereas for high-efficacy trials, a correct response within the response deadline would be rewarded, for low-efficacy trials, reward was not dependent upon participants’ performance, but were instead drawn randomly from a rolling window of the past 10 high-efficacy trials. That is, on each trial, a random sample was selected from the preceding 10 high-efficacy trials and if this trial was rewarded then so was the present low-efficacy trial. Thus, reward rates were approximately matched within a 10-trial window between the high-efficacy and low-efficacy conditions. At the end of the experiment, 10 trials were drawn at random to determine the total bonus reward earned by the participant—which could therefore range between 0 and 10 Euros (mean = 5.17, SD = 1.25).

Before starting the main task, participants were required to complete several practice phases. First, they learned the response mapping by responding to the letters “XXXXX” presented at fixation in the same size, position, and colors as the Stroop stimulus. They were required to press the “d”, “f”, “j”, or “k” keys corresponding to the correct answer for each color as in the main task with no time limit. These trials were presented in blocks of 16 trials (with four repetitions of each ink color) and were required to achieve at least 70% accuracy in one block to proceed.

Next, participants were familiarized with the cue stimuli. The same cues as in the main task were presented on-screen with text below displaying the possible meanings of the cue along with button assignments to indicate those options (e.g., “d: high reward, low efficacy”, “f: low reward, low efficacy” etc.).

Participants pressed a key on the keyboard corresponding to the correct meaning of the cue, again with no time limit for responses. As in the response mapping phase, participants were required to achieve at least 70% accuracy for one block of 16 trials to proceed to the next phase. Next, participants practiced the primary task. This practice was identical to the experimental task, described above (as displayed in Figure 1). Once again, participants had to achieve at least 70% accuracy for one block of 48 trials in order to proceed to the main task.

Our task and procedure were thus identical to those used by Frömer et al. (2021) with the following changes: (1) Our cue-to-target interval was 700 ms longer (2200 ms instead of 1500 ms) to allow more time for preparatory pupil effects to emerge; (2) The luminance of our cue stimuli was matched in order to rule out differential luminance effects on pupil size; (3). We did not include congruent Stroop stimuli to ensure a consistently challenging task and that the cued efficacy and reward were the primary factors affecting performance; (4) The response deadline during the main task was 1000 ms instead of 750 ms to compensate for the increased RTs on incongruent trials.

1.3 | Pupil Size Recording

Pupil size was recorded continuously throughout the main experimental task using an Eyelink 1000 Plus eyetracker (SR Research Ltd., Ottawa, ON, Canada). The camera was positioned below the computer monitor at a distance of 60 cm from the participant. Participants were instructed to remain still with their head in the chinrest throughout experimental blocks and to limit eye movements as much as possible. Data were recorded from both eyes where possible, with a sampling rate of 1000 Hz. However, due to experimenter error, the data for four participants were instead recorded at a sampling rate of 500 Hz, and so all other data were subsequently downsampled to this frequency. For three participants it was not possible to obtain a clear and reliable image of both eyes (e.g., due to the participants’ glasses), and so in these cases only the left eye was recorded.

1.4 | Analysis

Mean accuracy and RT were computed for each participant in each condition of the main task and analyzed with repeated-measures ANOVA (rANOVA). Trials with RTs faster than 200ms or slower than the condition-wise mean plus 2.5 SDs were excluded from further analysis.

Pupil data were analyzed in Python using custom scripts and some functions from the Pypillometry package (Mittner 2020). Data for each participant were preprocessed individually: First the data were smoothed using a Butterworth filter with a low-pass cutoff of 10 Hz. Next, blinks were identified based on velocity thresholds applied to the pupil size data (individual thresholds determined for each participant). Identified blink periods were then interpolated using a cubic spline as recommended by Mathôt et al. (2013). The complete, continuous pupil data for each participant were visually inspected in 5-s epochs to ensure optimal detection and interpolation of blinks. Pupil data were then baseline-corrected by subtracting the mean of the pupil size during the 500ms preceding cue onset from the subsequent signal (for cue-based analyses) or 500ms preceding the Stroop target onset (for Stroop-based analyses). Finally, event-related pupil responses were calculated by taking the mean of all pupil traces

for each condition (Figure 2). Mean, baseline-corrected pupil size was then compared with repeated-measures ANOVA (rANOVA) in sequential 50 ms bins throughout the cue-to-target and post-Stroop intervals (in each case comprising 44×50 ms bins). The results of these analyses were then subject to a false-discovery rate (FDR) correction for multiple comparisons. To analyze preparatory ‘ramping’ of pupil size, a linear function was fit to the mean absolute pupil size over the period from 1500 to 2500 ms after cue onset (with the baseline correction relative to cue onset as described above), thereby overlapping with the stimulus onset by 300 ms.

2 | Results

2.1 | Behavior

Figure 3, left panel displays the mean correct reaction times for each condition. The corresponding rANOVA revealed a significant main effect of reward, $F(1,36)=5.88$, $p=0.021$, $\eta_p^2=0.14$, as participants responded faster to high-reward trials. The main effect of efficacy was also significant, $F(1,36)=4.37$, $p=0.044$, $\eta_p^2=0.11$, with faster responses to high-efficacy trials than low-efficacy trials. Although numerically the fastest responses were

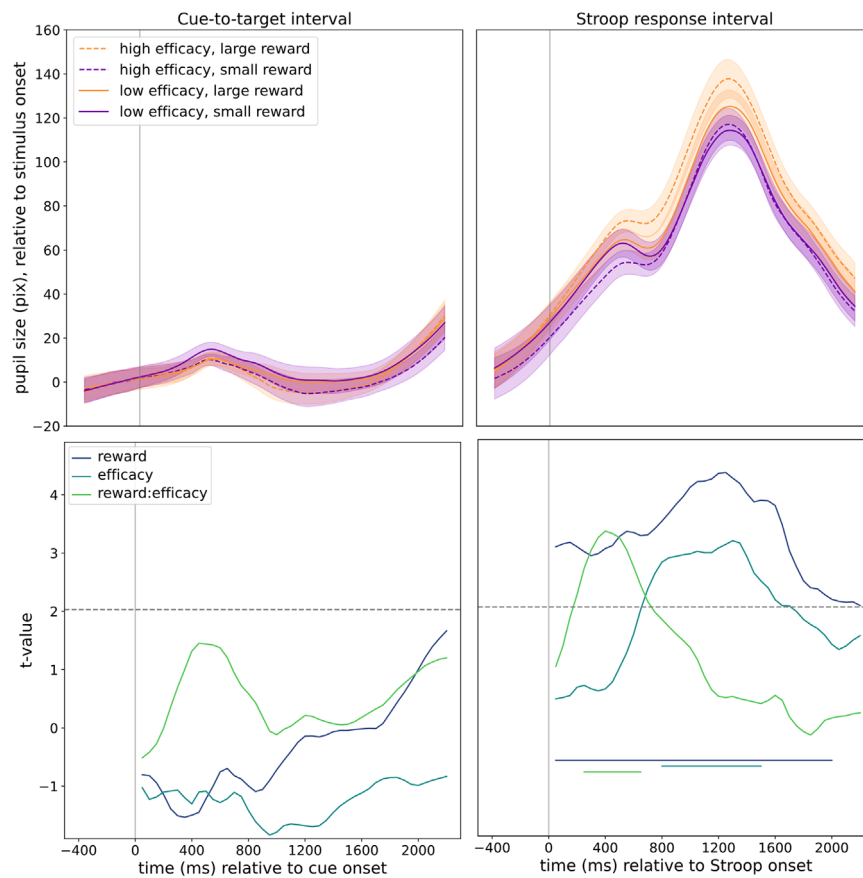


FIGURE 2 | Mean, baseline-corrected pupil size and t -values for the post-cue and post-Stroop time-windows of interest. Top left panel represents the mean, baseline-corrected pupil size (pixels) during the cue-to-target interval; top right panel represents mean, baseline-corrected pupil size following the Stroop stimulus; lower panels represent t -value time-course for comparisons of mean pupil size in sequential 50 ms bins for the cue-to-target interval (bottom left) and following the Stroop stimulus (bottom right). Solid vertical lines represent stimulus onset times; dashed, black horizontal lines indicate the $p=0.05$ significance threshold; solid horizontal lines in the lower portion of the bottom right panel indicate t -values which survive FDR correction for multiple comparisons.

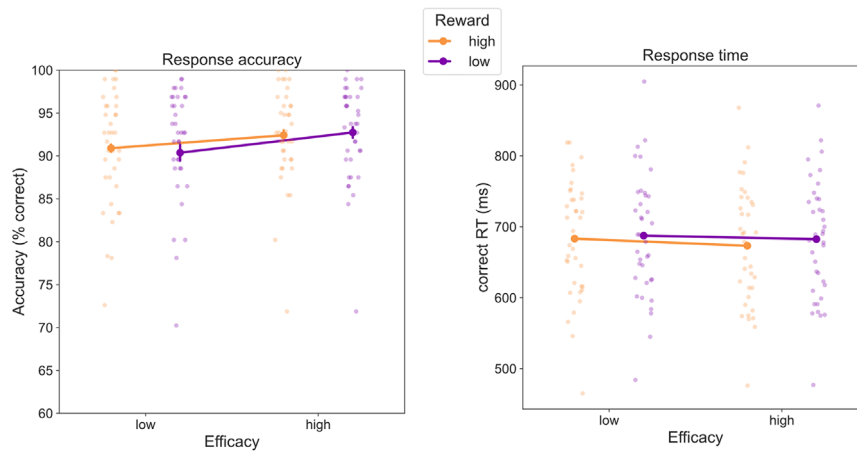


FIGURE 3 | Mean response accuracy (left) and response times (right) in the Stroop task, error bars represent ± 1 SE corrected for within subjects differences (Cousineau 2005).

for high-reward and high-efficacy trials, the interaction was not significant, $F(1,36) = 0.60$, $p = 0.444$, $\eta_p^2 = 0.02$.

A similar rANOVA applied to accuracy (Figure 3, right panel) did not reveal any fully significant effects for reward, $F(1,36) = 0.03$, $p = 0.865$, $\eta_p^2 < 0.01$, efficacy, $F(1,36) = 3.32$, $p = 0.077$, $\eta_p^2 = 0.08$, or their interaction $F(1,36) = 0.59$, $p = 0.447$, $\eta_p^2 = 0.02$. Concerning the trend-level main effect of efficacy, it is important to note that accuracy was higher for high efficacy, ruling out any speed-accuracy trade-offs (performance was authentically higher for high-efficacy trials). Similarly, the absence of any numerical effect for reward also indicates that the RT benefit for these trials can be considered a real behavioral improvement.

2.2 | Pupil Size

2.2.1 | Cue Response

Average pupil size was cut into sequential 50 ms bins, and the mean pupil size in each bin was analyzed in a repeated-measures analysis of variance (rANOVA) with factors for reward (large, small) and efficacy (high, low) (Figure 2, left panels). There were no significant effects of reward, efficacy or their interaction at any point during the cue-to-target interval (all $F < 2.49$, all $p > 0.123$, uncorrected). These results therefore replicate the absence of cue-evoked reward effects reported by Frömer et al. (2021) and others (Kostandyan et al. 2019); however, we also do not observe any effects of efficacy, contrasting with the larger pupils observed by Frömer et al. (2021) for low-efficacy trials.

2.2.2 | Pre-Stroop Slope

Our primary interest was in the effects of reward and efficacy on the rate of anticipatory pupil dilation (e.g., Unsworth and Miller 2023). To this end, a linear function was fit to the mean pupil size for each condition from 700 ms before Stroop onset until 300 ms after Stroop onset (Figure 4). We selected this time window to best capture the ‘ramping up’ effect in anticipation of the Stroop stimulus because, for all conditions, pupils began

to dilate around 700 ms prior to target onset. We reasoned that Stroop-stimulus-evoked responses are unlikely to be observed until at least 400 ms or so after Stroop onset. The resulting linear slopes (Figure 5) were then compared via a rANOVA with factors for reward and efficacy.

This analysis revealed a significant effect of reward, $F(1,36) = 13.8$, $p < 0.001$, $\eta_p^2 = 0.28$, as slopes were significantly steeper for large than for small rewards. The main effect for efficacy was not significant, $F(1,36) = 0.20$, $p = 0.656$, $\eta_p^2 < 0.01$, but there was a significant interaction between reward and efficacy, $F(1,36) = 6.18$, $p = 0.018$, $\eta_p^2 = 0.15$. Follow-up paired-samples t -tests revealed significantly steeper slopes (more rapid pupil dilation) for large rewards than for small rewards within the high-efficacy condition, $t(36) = 4.09$, $p < 0.001$, $d = 0.67$ but not within the low-efficacy condition, $t(36) = 1.1$, $p = 0.279$, $d = 0.18$.

2.2.3 | Stroop Response

As in the cue-target interval, average pupil size was once again cut into 50 ms bins, and the average pupil size in each bin analyzed via rANOVA with factors for reward (large, small) and efficacy (high, low) (Figure 2, right panels), a FDR correction was applied to the resulting p -values (surviving effects are represented by the solid horizontal lines at the top of the lower panel of Figure 2). As can be seen from Figure 2, a large and highly significant main effect of reward can be observed almost throughout the entire time window, as pupils were larger for large-reward trials than for small-reward trials. This effect is significant already within the first 50 ms time-bin, $F(1, 36) = 8.56$, $p = 0.006$ (uncorrected, FDR-corrected $p = 0.019$), $\eta_p^2 = 0.19$, suggesting that this reflects a late-preparatory effect as opposed to a Stroop-evoked response, which would take several hundreds of milliseconds to emerge (e.g., Kostandyan et al. 2019). The main effect of efficacy was not significant during this early time-window, but became significant in later time bins, surviving FDR correction from 800 ms until 1550 ms (all $F > 5.65$, all FDR-corrected $p < 0.047$). As can be seen from Figure 2, pupils were larger on high-efficacy trials than on low-efficacy trials. This main effect of efficacy was preceded by a significant interaction effect, which survived FDR correction in sequential bins from

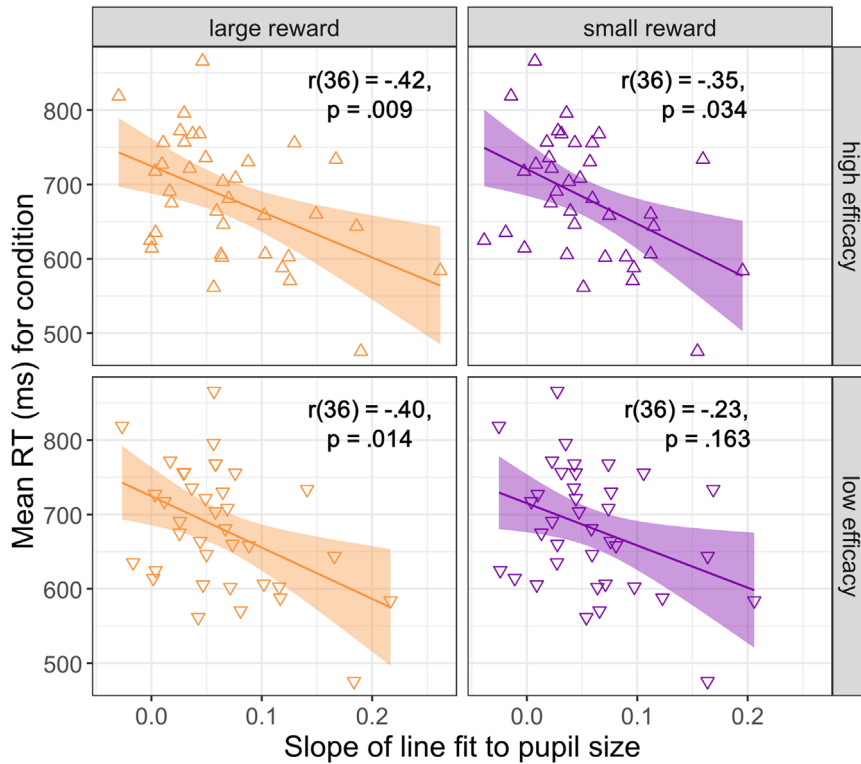


FIGURE 4 | Scatterplots of the correlation between average RT and pupil dilation slopes for each experimental condition.

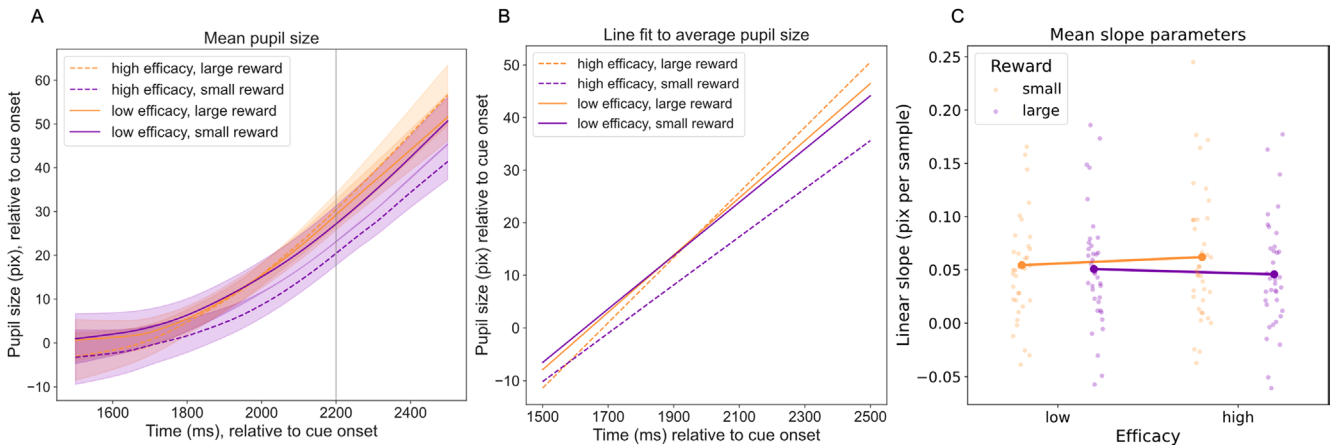


FIGURE 5 | Average pupil size (A), linear fit to pupil size (B), and slope of the linear fits (C) within the time window from 700 ms before until 300 ms after the Stroop stimulus onset.

250 ms until 650 ms (all $F > 5.76$, all FDR-corrected $p < 0.047$). To better understand this interaction, we performed pairwise comparisons in each of the time-windows where the interaction was significant. In each of these windows, pupils were significantly larger for high-reward trials than for low-reward trials, but only when efficacy was high (all t 's > 3.53 , all p 's < 0.002) and not when efficacy was low (all t 's < 1.25 , all p 's > 0.23).

2.2.4 | Individual Differences

Unsworth and Miller (2023) found that the slope of anticipatory pupil dilation was predictive of individual differences in working memory capacity and Stroop task performance (individuals

with steeper pupil size increases had smaller Stroop effects and faster average RTs). Our study was not designed to measure individual differences and thus lacks the sample size or additional measures (working memory capacity) used by Unsworth and Miller. Nevertheless, for completeness, we assessed the correlation between individual's pupil dilation slopes and RT in each condition of our experiment (Figure 4). As can be seen from Figure 4, faster RTs within each condition were associated with steeper increases in pupil size. This was confirmed by the presence of a significant negative correlation between RT and the slope of the linear fit to pupil size in all conditions except for low-reward and low-efficacy (Figure 4, bottom right panel). The strongest correlation was between slope and RT for large-reward and high-efficacy trials, a comparison of correlation strengths

conducted with the COCOR package in R (Diedenhofen and Musch 2015) indicated that this association was significantly stronger than the weakest, non-significant correlation between pupil slopes and RT for small-reward, low-efficacy trials, all $z=2.2$, $p=0.024$ (see Table S1 for full correlation matrix and table of comparisons). The relationship between pupil dilation rate and behavioral performance in this task thus appears to depend upon the motivational factors relating to the expected value of control. As a control analysis, we also examined the correlations between absolute pupil size in the post-Stroop interval and behavioral RT. These analyses (see Supporting Information) revealed the same pattern of results as for preparatory pupil size slopes, suggesting that, to a large extent, the differences in pupil size following the Stroop stimulus can be attributed to changes in preparatory effort allocation.

3 | Discussion

In the present study, we investigated how anticipated reward and efficacy influence preparatory effort using pupillometry. Our work thereby strongly builds on the work by Frömer et al. (2021), who spearheaded work in this domain in an EEG-pupillometry set-up but observed unexpected pupillometry results. Using a slightly adjusted version of the cued-Stroop task developed by Frömer et al. (2021), we cued participants to expect either large or small rewards, which could either be dependent on their performance or on a weighted coin-toss. We found no modulation of absolute pupil size in response to the cues or in the cue-to-target interval; however, we did observe significant differences in pupil size following Stroop target onset. This effect began already within the first few hundreds of milliseconds, too early to have been triggered by the targets themselves, as pupil size changes with a delay of several hundred milliseconds for cognitive effects (Denison, Parker, and Carrasco 2020; Hoeks and Levelt 1993; Kostandyan et al. 2019). The effect therefore likely represents a late (or even well-timed) engagement of preparatory effort in response to the cue. We also observed that pupil size began to ‘ramp up’ rapidly in the final moments before Stroop onset, replicating recent demonstrations by Unsworth and Miller (2023) and Tromp, Nieuwenhuis, and Murphy (2022), who have emphasized the rate of change to be an important characteristic of the pupil response as opposed to its absolute size. Crucially, we found that the slope of this pre-target pupil dilation was modulated by reward and efficacy, such that it was steepest for large-reward trials in which the reward was performance-contingent. Finally, the rate of pupil dilation was significantly correlated with behavioral response times across participants, especially in performance-contingent, large-reward trials.

Our study was closely based on that of Frömer et al. (2021), who used the same cued reward and efficacy manipulation and measured pupil size and EEG responses. Our results differ from those of Frömer et al. (2021) in several key ways, however. First, they observed a significant interaction between reward and efficacy in response times, whereas we observed significant main effects of both reward and efficacy but no interaction (although our participants’ response times did numerically replicate the pattern of results observed by Frömer et al.). It is noteworthy, though, that this interaction was only significant in two of the three experiments reported by Frömer et al. (2021). It is

possible that slight differences in our task drove this difference: to emphasize the cues in our study and to avoid any possible confounds relating to congruency and success rate, we did not include congruent Stroop trials (i.e., trials where the ink color matches the word), whereas Frömer et al. did; we also used a less strict response time cutoff of 1000 ms instead of 750 ms. We did so in order to account for the fact that responses on incongruent trials are slower than on congruent ones, but probably increased the deadline to an extent that made it easier for participants to respond in-time. More crucially, although Frömer et al. (2021) observed the predicted interaction between efficacy and reward in their anticipatory ERP (CNV) results, the pattern of cue-related pupil dilation was markedly divergent from both the results we report here and their own predictions. Specifically, they observed no modulation of pupil size whatsoever by reward but significantly smaller pupils for high-efficacy trials. Similarly, we observed no effect of reward on absolute pupil size in the cue-to-target interval, however, we also failed to observe any effect of efficacy. We did, however, observe significant differences in pupil size very soon after the Stroop stimulus onset, including an interacting effect of reward and efficacy. We also observed significantly steeper pupil dilation in the final hundreds of milliseconds prior to large-reward trials, which interacted with efficacy. Both effects were generally in line with the expectations and with the preparatory state implied by the CNV findings by Frömer et al. (2021). The pupil analyses reported by Frömer et al. (2021) employed a modeling approach in which they compared the peak of a canonical pupil response following stimulus onset. Based on our results, this peak was likely too early to optimally capture the very late-preparatory response to the cue, thereby obscuring any potential effects that may have been present. Similarly, this approach would not be suited to capture the differences in dilation slopes that we observed in the final moments before Stroop onset.

These results confirm and extend recent findings suggesting that the ‘ramping up’ of pupil size in anticipation of a Stroop stimulus reflects anticipatory effort (Tromp, Nieuwenhuis, and Murphy 2022; Unsworth and Miller 2023). Specifically, Unsworth and Miller (2023) found that pupil size increased linearly over the second prior to stimulus onset and was steeper when participants were cued to expect an incongruent than a congruent stimulus. Unsworth and Miller did not include any reward manipulation in their study, and we did not include a congruence or difficulty manipulation, but the convergence of results nevertheless suggests that the rate of pupil dilation reflects anticipatory effort allocation. Whether motivated by greater necessity (in the case of anticipated incongruent stimuli) or by incentive (in the case of reward), pupils dilate more rapidly when more effort is allocated to the task. In addition to the steeper slopes on trials where greater effort was required, Unsworth and Miller also observed significant correlations between individuals’ pupil dilation rates and their behavioral response times (as well as congruency difference scores and working memory capacity; see also Tromp, Nieuwenhuis, and Murphy 2022). While not intended as an individual differences study, our results further replicate these effects and extend them to the context of reward motivation: Not only were steeper slopes associated with faster response times in our data, but this association also depended on motivational factors, such that the relationship was strongest for large-reward trials in which the reward depended on good performance, whereas the association

was significantly weaker and non-significant in small-reward, low-efficacy trials.

These results provide a resolution to the contradictory results in previous literature regarding pupil size as a marker of preparatory effort. While several studies have provided evidence that pre-stimulus pupil size can index reward-induced preparatory effort (Boehler et al. 2011; Wang, Brien, and Munoz 2015; Wang et al. 2016), others have found no effect of reward cues in the cue-to-target interval, but only after the target stimulus has appeared (Frömer et al. 2021; Kostandyan et al. 2019). As noted, however, these studies do observe differences in absolute or baseline-corrected pupil size very shortly after the target appears (e.g., Kostandyan et al. 2019), which we also observed here. Our results provide a possible explanation for these differences: anticipatory effort is reflected in pupil size, but only shortly before the onset of the stimulus, and the rate of pupil dilation is a better marker of this effort than absolute pupil size. This makes sense in the context of neurophysiological studies, which suggest that the first temporal derivative (rate of change) of pupil size is more tightly linked to cortical norepinephrine levels and with less temporal lag (Pfeffer et al. 2022; Reimer et al. 2016; Tromp, Nieuwenhuis, and Murphy 2022). Interestingly, this is then very closely followed by significant effects of reward and, slightly later, an interaction between reward and efficacy in the data that is time-locked to the target onset. This closely mirrors earlier findings by Kostandyan et al. (2019), who observed no reward effects after a cue but immediate effects after target presentation, which were also interpreted as most likely still reflecting the influence of the cues. In fact, given that pupil size is known to be a lagged indicator of the underlying neural processes, it may indicate that participants were able to time their effort investment very well, getting ready just before target presentation. This may then be visible earlier in the ramping response toward the end of the cue-target interval and a little later in pupil size itself.

Absolute pupil size and the slope of pupil dilation are clearly closely related constructs, but are comparatively weakly correlated (van den Brink et al. 2016), and have been shown to some extent to reflect distinct neural correlates (e.g., Reimer et al. 2016). Given that we observe both changes in pre-stimulus pupil slopes and very early post-stimulus changes in absolute pupil size, it is impossible to know the degree to which these effects reflect distinct underlying processes. However, both observations are in line with a proactive account, as both emerge before reactive control can have emerged (Braver 2012).

It may be the case that in our experiment, a very strong proactive mode of control was not required for participants to perform well. Indeed, the task used a relatively relaxed response deadline, and our participants performed very well (i.e., accuracy was generally around 90%). It is therefore possible that a more demanding task, in particular in the sense of tighter time–pressure, would encourage participants to recruit proactive control to a greater extent, which may then manifest in greater differences in pupil dilation or even absolute size prior to stimulus onset. However, the original findings of Frömer et al. (2021) suggest that this still may not be the case. Their response deadline was 25% shorter than ours, yet they still found no reward-related effects on absolute pupil size. Nevertheless, the degree to which proactive control is the

optimal strategic choice in a task is an important consideration which may also play into discrepancies in the previous literature, some tasks may benefit more than others from proactive control, and indeed some participants may have greater capacity to allocate either proactive or reactive control to a given task.

The pattern we observe in pupil dilation is strikingly similar to the pattern observed previously in the CNV component in similar studies (e.g., Frömer et al. 2021; Schevernels et al. 2014), thereby providing converging evidence that they may relate to a similar psychological process, arguably preparatory effort. Specifically, this component is reliably modulated by both reward and other task factors, including difficulty and efficacy, with the main effect of reward emerging first, shortly followed by an interaction in the last moments before target onset. In contrast, these studies typically also report a significant effect of reward on earlier (P3) ERP responses to cue stimuli, which is not reflected in a relatively early, cue-evoked pupil response at any time in the 2-s interval following the cue (i.e., of the type originally modeled by Frömer et al. 2021). This lends support to the suggestion that evaluation of the cued reward and efficacy (or difficulty) represents a distinct neurocognitive process to the later allocation of cognitive effort in anticipation of the task. That is, participants first compute the expected value of control for performing the task, then recruit arousal systems in accordance with that evaluation. The selective modulation of pupil dilation by the latter process would suggest a norepinephrinergetic mechanism of this resource allocation (Reimer et al. 2016; Unsworth and Robison 2017). Future research could elucidate this further by directly assessing the association between pupil dilation and other neural/electrophysiological markers of effort allocation, such as the CNV component.

In our results, as in Frömer et al. (2021), participants still allocated significant effort to the task even on low-efficacy (and low-reward) trials, as reflected by the relatively high average accuracy and fast RTs achieved on these trials. This could in part reflect switch-costs associated with changing levels of control from one trial to the next, as reconfiguring control settings may itself be effortful (Eayrs et al. 2024; Kool et al. 2010). Participants may therefore be slow to adjust their control settings, but with longer time horizons these adjustments may become more pronounced (see e.g. Grahek et al. 2023; Kukkonen et al. 2023). Another possibility is that participants invest cognitive effort into the task for reasons other than the incentives indicated by the cue stimuli, such as a learned association between effort and reward (e.g., Clay et al. 2022; Frömer et al. 2021) or social desirability of good performance.

In our results, it is not possible to completely distinguish between changes in pupil size related to reward anticipation as opposed to effort or resource allocation. Participants in the high-reward condition and high-efficacy conditions may have greater confidence in anticipating a positive outcome from the trial, which may manifest in greater levels of arousal. This explanation seems unlikely, however, as in our task reward probability on low-efficacy trials was continuously matched to the reward rate on high-efficacy trials (as in Frömer et al. 2021). Furthermore, there was no change in pupil size during the early cue-to-target interval, when participants were informed about the reward condition (whereas reward-related changes

in EEG indices of cue processing have been reported in the related previous literature; e.g., Frömer et al. 2021; Kostandyan et al. 2019; Schevernels et al. 2014). From Figure 2, it is also evident that the reward- and efficacy-related changes in pupil size have already dissipated by the time reward feedback appeared onscreen. The fact that all effects of both reward and efficacy on pupil size are tightly linked to the onset of the Stroop stimulus, and the absence of cue-evoked and feedback-preceding differences in pupil size, suggest that these effects reflect preparatory recruitment of cognitive effort to perform the task. Much in line with such an interpretation, Unsworth, Miller, and Aghel (2022) found analogous results to ours (steeper pupil dilation in anticipation of the target stimulus) with a cue simply prompting participants to ‘try hard’ without any reward incentive.

In conclusion, our results suggest that pupil dilation is a sensitive marker of effort allocation, particularly in the moments immediately preceding a task. While absolute pupil size in response to cues did not vary significantly with reward or efficacy before the end of the cue-target interval, the rate of pupil dilation before target onset was clearly modulated by these factors, especially when large rewards were contingent on performance. This finding aligns with and extends previous research, highlighting the importance of considering the temporal dynamics of pupil responses, rather than just static measures of pupil size. The observed correlations between pupil dilation rates and behavioral response times further underscore the link between anticipatory effort and task performance, particularly under conditions of high motivational stakes. Future research should continue to explore the neurophysiological mechanisms underlying these effects, potentially integrating other markers of cognitive effort such as the CNV component to provide a more comprehensive understanding of how reward and efficacy influence preparatory effort.

Author Contributions

Joshua O. Eayrs: conceptualization, data curation, formal analysis, investigation, methodology, project administration, software, validation, visualization, writing – original draft, writing – review and editing. **Haya Serena Tobing:** conceptualization, data curation, formal analysis, investigation, methodology, project administration, writing – review and editing. **S. Tabitha Steendam:** conceptualization, writing – review and editing. **Nicoleta Prutean:** conceptualization, writing – review and editing. **Wim Notebaert:** conceptualization, funding acquisition, writing – review and editing. **Jan R. Wiersema:** conceptualization, funding acquisition, writing – review and editing. **Ruth M. Krebs:** conceptualization, funding acquisition, writing – review and editing. **C. Nico Boehler:** conceptualization, funding acquisition, methodology, project administration, supervision, writing – review and editing.

Acknowledgments

This work was supported by a concerted research project awarded to R.M.K., C.N.B., W.N., and J.R.W. (Special Research Fund Ghent University, grant number: BOF20-GOA-004).

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data and code pertaining to this research will be made available upon publication on the Open Science Framework (OSF; <https://osf.io/9yjd/>).

References

- Alnaes, D., M. H. Sneve, T. Espeseth, T. Endestad, S. H. P. Van De Pavert, and B. Laeng. 2014. “Pupil Size Signals Mental Effort Deployed During Multiple Object Tracking and Predicts Brain Activity in the Dorsal Attention Network and the Locus Coeruleus.” *Journal of Vision* 14, no. 4: 1. <https://doi.org/10.1167/14.4.1>.
- Aston-Jones, G., and J. D. Cohen. 2005. “An Integrative Theory of Locus Coeruleus-Norepinephrine Function: Adaptive Gain and Optimal Performance.” *Annual Review of Neuroscience* 28, no. 1: 403–450. <https://doi.org/10.1146/annurev.neuro.28.061604.135709>.
- Boehler, C. N., J.-M. Hopf, R. M. Krebs, et al. 2011. “Task-Load-Dependent Activation of Dopaminergic Midbrain Areas in the Absence of Reward.” *Journal of Neuroscience* 31, no. 13: 4955–4961. <https://doi.org/10.1523/JNEUROSCI.4845-10.2011>.
- Braver, T. S. 2012. “The Variable Nature of Cognitive Control: A Dual Mechanisms Framework.” *Trends in Cognitive Sciences* 16, no. 2: 106–113. <https://doi.org/10.1016/j.tics.2011.12.010>.
- Bugg, J. M., and A. Smallwood. 2016. “The Next Trial Will Be Conflicting! Effects of Explicit Congruency Pre-Cues on Cognitive Control.” *Psychological Research* 80, no. 1: 16–33. <https://doi.org/10.1007/s00426-014-0638-5>.
- Campbell, J. I. D., and V. A. Thompson. 2012. “MorePower 6.0 for ANOVA With Relational Confidence Intervals and Bayesian Analysis.” *Behavior Research Methods* 44, no. 4: 1255–1265. <https://doi.org/10.3758/s13428-012-0186-0>.
- Clay, G., C. Mlynski, F. M. Korb, T. Goschke, and V. Job. 2022. “Rewarding Cognitive Effort Increases the Intrinsic Value of Mental Labor.” *Proceedings of the National Academy of Sciences of the United States of America* 119, no. 5: e2111785119. <https://doi.org/10.1073/pnas.2111785119>.
- Cousineau, D. 2005. “Confidence Intervals in Within-Subject Designs: A Simpler Solution to Loftus and Masson’s Method. Tutorials in Quantitative Methods for.” *Psychology* 1, no. 1: 42–45. <https://doi.org/10.20982/tqmp.01.1.p042>.
- Denison, R. N., J. A. Parker, and M. Carrasco. 2020. “Modeling Pupil Responses to Rapid Sequential Events.” *Behavior Research Methods* 52, no. 5: 1991–2007. <https://doi.org/10.3758/s13428-020-01368-6>.
- Diedenhofen, B., and J. Musch. 2015. “Cocor: A Comprehensive Solution for the Statistical Comparison of Correlations.” *PLoS One* 10, no. 4: e0121945. <https://doi.org/10.1371/journal.pone.0121945>.
- Eayrs, J., and N. Lavie. 2018. “Establishing Individual Differences in Perceptual Capacity.” *Journal of Experimental Psychology: Human Perception and Performance* 44, no. 8: 1240–1257.
- Eayrs, J. O., N. Kukkonen, N. Prutean, et al. 2024. “Attentional Set and Explicit Expectations of Perceptual Load Determine Flanker Interference.” *Journal of Experimental Psychology: Human Perception and Performance* 50: 769–784.
- Eayrs, J. O., and N. Lavie. 2019. “Individual Differences in Parietal and Frontal Cortex Structure Predict Dissociable Capacities for Perception and Cognitive Control.” *NeuroImage* 202: 116148.
- Engle, R. W., and M. J. Kane. 2003. “Executive Attention, Working Memory Capacity, and a Two-Factor Theory of Cognitive Control.” *Psychology of Learning and Motivation* 44: 145–199. [https://doi.org/10.1016/S0079-7421\(03\)44005-X](https://doi.org/10.1016/S0079-7421(03)44005-X).

- Eriksen, B. A., and C. W. Eriksen. 1974. "Effects of Noise Letters Upon the Identification of a Target Letter in a Nonsearch Task." *Perception & Psychophysics* 16, no. 1: 143–149. <https://doi.org/10.3758/BF03203267>.
- Ferguson, K. A., and J. A. Cardin. 2020. "Mechanisms Underlying Gain Modulation in the Cortex." *Nature Reviews Neuroscience* 21, no. 2: 80–92. <https://doi.org/10.1038/s41583-019-0253-y>.
- Frömer, R., H. Lin, C. K. Dean Wolf, M. Inzlicht, and A. Shenhav. 2021. "Expectations of Reward and Efficacy Guide Cognitive Control Allocation." *Nature Communications* 12, no. 1: 1030. <https://doi.org/10.1038/s41467-021-21315-z>.
- Gilzenrat, M. S., S. Nieuwenhuis, M. Jepma, and J. D. Cohen. 2010. "Pupil Diameter Tracks Changes in Control State Predicted by the Adaptive Gain Theory of Locus Coeruleus Function." *Cognitive, Affective, & Behavioral Neuroscience* 10, no. 2: 252–269. <https://doi.org/10.3758/CABN.10.2.252>.
- Goldstein, R. Z., L. A. Cottone, Z. Jia, T. Maloney, N. D. Volkow, and N. K. Squires. 2006. "The Effect of Graded Monetary Reward on Cognitive Event-Related Potentials and Behavior in Young Healthy Adults." *International Journal of Psychophysiology* 62, no. 2: 272–279. <https://doi.org/10.1016/j.ijpsycho.2006.05.006>.
- Grahek, I., R. Frömer, M. Prater Fahey, and A. Shenhav. 2023. "Learning When Effort Matters: Neural Dynamics Underlying Updating and Adaptation to Changes in Performance Efficacy." *Cerebral Cortex* 33, no. 5: 2395–2411. <https://doi.org/10.1093/cercor/bhac215>.
- Hess, E. H., and J. M. Polt. 1964. "Pupil Size in Relation to Mental Activity During Simple Problem-Solving." *Science* 143, no. 3611: 1190–1192.
- Hoeks, B., and W. J. M. Levelt. 1993. "Pupillary Dilation as a Measure of Attention: A Quantitative System Analysis." *Behavior Research Methods, Instruments, & Computers* 25, no. 1: 16–26. <https://doi.org/10.3758/BF03204445>.
- Kahneman, D., and J. Beatty. 1966. "Pupil Diameter and Load on Memory." *Science* 154, no. 3756: 1583–1585.
- Kool, W., J. T. McGuire, Z. B. Rosen, and M. M. Botvinick. 2010. "Decision Making and the Avoidance of Cognitive Demand." *Journal of Experimental Psychology: General* 139, no. 4: 665–682. <https://doi.org/10.1037/a0020198>.
- Kostandyan, M., K. Bombeke, T. Carsten, R. M. Krebs, W. Notebaert, and C. N. Boehler. 2019. "Differential Effects of Sustained and Transient Effort Triggered by Reward – A Combined EEG and Pupillometry Study." *Neuropsychologia* 123: 116–130. <https://doi.org/10.1016/j.neuropsychologia.2018.04.032>.
- Krebs, R. M., C. N. Boehler, K. C. Roberts, A. W. Song, and M. G. Woldorff. 2012. "The Involvement of the Dopaminergic Midbrain and Cortico-Striatal-Thalamic Circuits in the Integration of Reward Prospect and Attentional Task Demands." *Cerebral Cortex* 22, no. 3: 607–615. <https://doi.org/10.1093/cercor/bhr134>.
- Kukkonen, N., S. Braem, J. Allaert, et al. 2023. "The Cost of Regulating Effort: Reward and Difficulty Cues With Longer Prediction Horizons Have a Stronger Impact on Performance."
- Mathôt, S., E. Aarts, M. Verhage, J. Veenvliet, C. Dolan, and S. van der Sluis. 2013. "A Simple Way to Reconstruct Pupil Size During Eye Blinks."
- Mittner, M. 2020. "Pupillometry: A Python Package for Pupillometric Analyses." *Journal of Open Source Software* 5, no. 51: 2348. <https://doi.org/10.21105/joss.02348>.
- Murphy, P. R., R. G. O'Connell, M. O'Sullivan, I. H. Robertson, and J. H. Balsters. 2014. "Pupil Diameter Covaries With BOLD Activity in Human Locus Coeruleus." *Human Brain Mapping* 35, no. 8: 4140–4154. <https://doi.org/10.1002/hbm.22466>.
- Pfeffer, T., C. Keitel, D. S. Kluger, et al. 2022. "Coupling of Pupil- and Neuronal Population Dynamics Reveals Diverse Influences of Arousal on Cortical Processing." *eLife* 11: e71890. <https://doi.org/10.7554/eLife.71890>.
- Reimer, J., M. J. McGinley, Y. Liu, et al. 2016. "Pupil Fluctuations Track Rapid Changes in Adrenergic and Cholinergic Activity in Cortex." *Nature Communications* 7, no. 1: 13289. <https://doi.org/10.1038/ncomm513289>.
- Rondeel, E. W. M., H. Van Steenbergen, R. W. Holland, and A. Van Knippenberg. 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. <https://doi.org/10.3389/fnhum.2015.00494>.
- Schevernels, H., R. M. Krebs, P. Santens, M. G. Woldorff, and C. N. Boehler. 2014. "Task Preparation Processes Related to Reward Prediction Precede Those Related to Task-Difficulty Expectation." *NeuroImage* 84: 639–647. <https://doi.org/10.1016/j.neuroimage.2013.09.039>.
- Schneider, M., L. Leuchs, M. Czisch, P. G. Sämann, and V. I. Spormaker. 2018. "Disentangling Reward Anticipation With Simultaneous Pupillometry/fMRI." *NeuroImage* 178: 11–22. <https://doi.org/10.1016/j.neuroimage.2018.04.078>.
- Shenhav, A., M. M. Botvinick, and J. D. Cohen. 2013. "The Expected Value of Control: An Integrative Theory of Anterior Cingulate Cortex Function." *Neuron* 79, no. 2: 217–240. <https://doi.org/10.1016/j.neuron.2013.07.007>.
- Shenhav, A., J. D. Cohen, and M. M. Botvinick. 2016. "Dorsal Anterior Cingulate Cortex and the Value of Control." *Nature Neuroscience* 19, no. 10: 1286–1291. <https://doi.org/10.1038/nn.4384>.
- Stroop, J. R. 1935. "Studies of Interference in Serial Verbal Reactions." *Journal of Experimental Psychology* 18, no. 6: 643–662.
- Tromp, J., S. Nieuwenhuis, and P. Murphy. 2022. "The Effects of Neural Gain on Reactive Cognitive Control." *Computational Brain & Behavior* 5, no. 3: 422–433. <https://doi.org/10.1007/s42113-022-00140-7>.
- Unsworth, N., and A. L. Miller. 2023. "Pupillary Correlates of Preparatory Control in the Stroop Task." *Attention, Perception, & Psychophysics* 85, no. 7: 2277–2295. <https://doi.org/10.3758/s13414-023-02751-z>.
- Unsworth, N., A. L. Miller, and S. Aghel. 2022. "Effort Mobilization and Lapses of Sustained Attention." *Cognitive, Affective, & Behavioral Neuroscience* 22, no. 1: 42–56. <https://doi.org/10.3758/s13415-021-00941-6>.
- Unsworth, N., and M. K. Robison. 2017. "A Locus Coeruleus-Norepinephrine Account of Individual Differences in Working Memory Capacity and Attention Control." *Psychonomic Bulletin & Review* 24, no. 4: 1282–1311. <https://doi.org/10.3758/s13423-016-1220-5>.
- Van Den Berg, B., R. M. Krebs, M. M. Lorist, and M. G. Woldorff. 2014. "Utilization of Reward-Prospect Enhances Preparatory Attention and Reduces Stimulus Conflict." *Cognitive, Affective, & Behavioral Neuroscience* 14, no. 2: 561–577. <https://doi.org/10.3758/s13415-014-0281-z>.
- Van Den Brink, R. L., P. R. Murphy, and S. Nieuwenhuis. 2016. "Pupil Diameter Tracks Lapses of Attention." *PLoS One* 11, no. 10: e0165274. <https://doi.org/10.1371/journal.pone.0165274>.
- Van Der Wel, P., and H. Van Steenbergen. 2018. "Pupil Dilation as an Index of Effort in Cognitive Control Tasks: A Review." *Psychonomic Bulletin & Review* 25, no. 6: 2005–2015. <https://doi.org/10.3758/s13423-018-1432-y>.
- Wang, C.-A., D. C. Brien, and D. P. Munoz. 2015. "Pupil Size Reveals Preparatory Processes in the Generation of Pro-Saccades and Anti-Saccades." *European Journal of Neuroscience* 41, no. 8: 1102–1110.

Wang, C.-A., H. McInnis, D. C. Brien, G. Pari, and D. P. Munoz. 2016. "Disruption of Pupil Size Modulation Correlates With Voluntary Motor Preparation Deficits in Parkinson's Disease." *Neuropsychologia* 80: 176–184. <https://doi.org/10.1016/j.neuropsychologia.2015.11.019>.

Westbrook, A., and T. S. Braver. 2015. "Cognitive Effort: A Neuro-economic Approach." *Cognitive, Affective, & Behavioral Neuroscience* 15, no. 2: 395–415. <https://doi.org/10.3758/s13415-015-0334-y>.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.