Citation for published version (APA):
Pupillometry tracks errors in interval timing

Shamini Warda\textsuperscript{1}, Jaana Simola\textsuperscript{2}, & Devin B. Terhune\textsuperscript{3}

\textsuperscript{1}Department of Humanities and Social Sciences, Indian Institute of Technology Bombay, Mumbai, India
\textsuperscript{2}Department of Education, University of Helsinki, Finland
\textsuperscript{3}Department of Psychology, Goldsmiths, University of London, London, UK

Correspondence:
Dr. Devin B. Terhune
Department of Psychology
Goldsmiths, University of London
New Cross, London, UK
d.terhune@gold.ac.uk

Word count: 3715
Words in abstract: 159
Number of Tables: 2
Number of Figures: 2

Acknowledgments
We thank Jake Sullivan for assistance with data collection. DBT is currently supported by the BBSRC (BB/R01583X/1). JS is currently supported by The Strategic Research Council (SRC) at the Academy of Finland (SA 1335236).
Abstract

Recent primate studies suggest a potential link between pupil size and subjectively elapsed duration. Here, we sought to investigate the relationship between pupil size and perceived duration in human participants performing two temporal bisection tasks in the subsecond and suprasecond interval ranges. In the subsecond task, pupil diameter was greater during stimulus processing when shorter intervals were overestimated but also during and after stimulus offset when longer intervals were underestimated. By contrast, in the suprasecond task, larger pupil diameter was observed only in the late stimulus offset phase prior to response prompts when longer intervals were underestimated. This pattern of results suggests that pupil diameter relates to an error monitoring mechanism in interval timing. These results are at odds with a direct relationship between pupil size and the perception of duration but suggest that pupillometric variation might play a key role in signifying errors related to temporal judgments.

Keywords: interval timing; pupillometry; error monitoring; attentional lapses
Accurate perception of time in the subsecond to suprasecond range is critical for many aspects of learning and behavior (Mauk & Ruiz, 1992; Savastano & Miller, 1998). Whereas an objective clock runs at a steady pace, our subjective time is prone to various distortions arising from factors such as attention (Brown et al., 1985; Mattes & Ulrich, 1998) and arousal (Penton-Voak et al., 1996). A promising biomarker to track attentional and arousal states in humans is pupil size (Murphy et al., 2011; Mathôt et al., 2013). Converging evidence suggests that changes in pupil size are an indicator of cognitive functioning reflecting a wide range of processes from visual processing to decision making (Hess & Polt, 1960; Murphy et al., 2014; Mathôt & Stigchel, 2015; Mathôt, 2018; Joshi & Gold, 2020).

Pupil-linked brain states, which are primarily associated with the activity in the locus coeruleus-norepinephrine system (Murphy et al., 2014; Joshi et al., 2016) but may also involve cholinergic (Reimer et al., 2016), dopaminergic (de Gee et al., 2017) and serotonergic (Schmid et al., 2015) systems, can potentially influence neural activity related to the monitoring of elapsed time (Suzuki et al., 2016; Suzuki & Tanaka, 2017). In a study by Suzuki et al. (2016), a negative correlation was observed between pupil size and latency of self-timed saccades. Monkeys made a memory-guided saccade after a previously trained specified interval and analysis of pupil size with respect to saccadic latency revealed a significantly larger pupil diameter in the early saccadic latency group. Early saccadic latency suggests an expansion of perceived duration, and therefore, its correlation with pupil size was considered to be indicative of an association between pupil size and perceived duration. However, it is not yet clear whether these findings generalize to tasks that require manual responses regarding time judgments in human participants.

Using pupillometry, Toscano-Zapién et al. (2016) aimed to understand the attentional mechanisms involved in the timing of subsecond durations. Although it was not predicted in advance, pupil diameter was larger when participants incorrectly judged the duration of stimulus intervals (including “short” responses for long intervals and “long” responses for short intervals). An independent body of evidence linking pupil size and temporal information processing comes from studies of temporal expectation. Pupil size has been shown to track temporal regularities, exhibiting different preparatory activity for different delay conditions (Akdoğan et al., 2016). For example, the pupil dilates at a higher rate when targets are expected to appear after a shorter relative to a longer delay period. Additionally, pupil dilation, given its involuntary nature, has been proposed to represent a potentially valuable measure of infants’ interval timing abilities (Addyman et al., 2014).
Despite these promising links between pupil diameter and timing, there has not yet been an explicit attempt to investigate if pupillary variability relates to intra-individual variability in perceived duration in human participants. Toward this end, the present study examined whether interval timing performance covaried with pupillometry by having participants complete subsecond and suprasecond visual temporal bisection tasks whilst their eye movements were recorded. In turn, we sought to examine the relationship between pupil size and perceived duration of visual stimulus intervals. Following Suzuki et al (2016), we expected that pupil diameter would be larger when intervals were perceived to be longer and we explored this association at multiple phases of stimulus processing. We also evaluated the competing hypothesis that pupil size indexes error processing (Toscano-Zapién et al., 2016), rather than perceived duration per se, and tested the alternative prediction that pupil diameter would be larger when participants make incorrect interval judgments.

Methods

Participants
A sample of 31 adults (10 females; $M_{Age}=23.9$, $SE=0.9$) consented to take part in the study in accordance with local ethical approval from the Division of Medical Sciences, University of Oxford. All participants were right-handed, self-reported normal or corrected-to-normal vision, and had completed secondary school, with an average of 3.7±0.5 years of higher education. Participants were recruited via fliers and word-of-mouth for an eye-tracking study on perception. These data consist of a re-analysis of a previous study that did not involve the analysis of pupillometric data (Terhune et al., 2016) and thus no formal a priori power analysis was undertaken for the present study. The original study was run to detect weak effect sizes in the range of 0.20 and above and optional stopping was not performed. Following preprocessing steps and data segregation (see below), the data from 6 participants were excluded from one of the two tasks (one participant had an excessively large number of blinks in both tasks, three participants had a large number of missing data points in the stimulus offset phase in suprasecond task, and two different participants in each of the tasks had perfect accuracy for at least one of the stimulus intervals), resulting in a final sample of 29 and 26 participants in the subsecond and suprasecond tasks, respectively.
Materials

Temporal bisection tasks
Participants completed two temporal bisection tasks in subsecond and suprasecond interval ranges. All visual stimuli were presented against a purple-grey background. Intervals in the two tasks ranged from 300 to 700ms in 63ms increments (7 intervals) and 1400 to 2600ms in 200ms increments, respectively. Participants were initially trained to distinguish between two anchor durations (subsecond: 300 vs. 700; suprasecond: 1400 vs. 2600). These intervals were selected to index subsecond (~500ms) and suprasecond (~2000ms) interval timing. They were subsequently presented with variable intervals and judged whether they were closer to the trained short or long anchor intervals. Each trial consisted of a blank jittered interstimulus interval (ISI; 1250-1450ms), the target interval stimulus (a centrally located light green circle), another jittered ISI (subsecond: 800-1200ms; suprasecond: 900-2100ms), and a response screen (S L; S=“short”; L=“long”) to which participants judged the duration of the interval relative to the anchor intervals using the index and middle fingers of their right hand (see Supplementary Figure 1).

Procedure
Participants were first seated comfortably in a light-controlled, sound-attenuated room and their head was placed on a chin and forehead rest at a distance of ~75cm from the monitor. Eye movements and pupil data were recorded using an Eye Link 1000 Desktop Mount eye tracker (SR Research, Ontario, Canada). Data were monocularly recorded at a rate of 500Hz from the right eye. The eyetracker was calibrated for each participant using a 9-point calibration procedure and the calibration was accepted if the average error was less than 0.5º. Participants then received onscreen and oral instructions regarding the completion of the two tasks, which were administered in counterbalanced order. They were instructed to fixate on the centre of the monitor throughout the task. Each participant completed a training session of 20 trials followed by four blocks of 70 trials with randomly presented stimulus intervals. Finger-response mappings were counterbalanced across participants. The tasks were presented using Experiment Builder® (v. 1.6.121; SR Research, Ontario, Canada). Stimuli subtended a visual angle of 1.73º x 1.73º. The background and stimuli were matched for luminance using a ColorCAL MkII colorimeter (Cambridge Research Systems Ltd: Rochester, UK). All participants were thanked and compensated for their time at a rate of £10 per hour.
Pupillometry tracks timing errors

Analyses

Pupillometric pre-processing

All analyses, including statistical analyses, were performed using MATLAB® (2018b, Math Works, Natick, MA). Each participant's pupillometric data were initially segregated into 2 phases (pre-stimulus: -1000ms to 0ms [stimulus onset]; post-stimulus: 0ms to response prompt [variable]) and 7 stimulus intervals. For each phase, a novel noise-based blink detection algorithm was used to identify start and end-points for each blink (Hershman et al., 2018). Each blink period was subsequently replaced by linear interpolation of pupil diameter values prior to and after the blink period. Trials with unidentifiable blink onset/offset markers and participants missing more than 20% of data points in any of the phases were removed from the sample. There were three participants in the suprasecond task who did not have a complete set of data points in the stimulus phase and one participant in both of the tasks with a large number of missing data points in both prestimulus and stimulus phases. This resulted in the exclusion of data from one and four participant(s) in the subsecond and suprasecond tasks, respectively.

Baseline correction was performed for the stimulus phase using the median pupil diameter of the final 100ms prior to stimulus onset. Trial-level outliers (median±3SD) were subsequently removed using a Hampel identifier (Hampel, 1974). Data were then down-sampled by a factor of 20 (computing the mean pupil diameter for every 20ms bin), such that each 1000ms window of data were represented by 50 data points per participant. The stimulus phase was subsequently segregated into three phases: stimulus onset (subsecond: 0 to 300ms; suprasecond: 0-1400ms), stimulus pre-offset (subsecond: -300 to 0ms; suprasecond: -1400 to 0ms), and stimulus offset (subsecond: 0 to 800ms; suprasecond: 0 to 900ms).

Statistical analyses

Pre-processed data for both tasks were next segregated according to responses and interval ranges collapsing across three stimulus intervals (subsecond: short: 300, 367, 433ms, long: 567, 633, 700ms; suprasecond: short: 1400, 1600, 1800ms, long: 2200, 2400, 2600ms). For both the pre-stimulus and each post-stimulus (baseline)-corrected phase, and for each stimulus interval range, the mean pupil diameter corresponding to short and long responses was computed for each time bin within the respective phase. Two participants who had perfect accuracy for either short or long intervals in either task were excluded from further analyses, resulting in a final sample size of 29 and 26 for the subsecond and suprasecond tasks respectively.
Pupillometry tracks timing errors

Pupil diameter in each time bin was subsequently analysed using a series of 2 x 2 repeated measures analyses of variance (ANOVAs) with Response (short vs. long) and Interval (short vs. long) as independent variables. Additionally, separate sets of ANOVAs with Response (correct vs error) and Interval (short vs. long) were performed on data in the pre-stimulus and three stimulus phases for both tasks. For each family of tests, statistical significance of main and interaction effects was corrected for multiple analyses using a false discovery rate (FDR) correction (Benjamini & Hochberg, 1995). Effect sizes (Hedges’s g) and corresponding Bootstrap 95% confidence intervals (3,000 samples) were computed for significant time windows (Hentschke & Stüttgen, 2011). This was done using the means computed from the baseline-corrected pupil diameter averaged across the significant (interaction) time window and all reported mean differences (MDs) reflect the mean of short responses subtracted from the mean of long response.

Results

Figure 1 shows baseline-corrected average pupil diameter as a function of stimulus interval and response in each of the three phases of the two temporal bisection tasks. The analyses of the subsecond task revealed no significant main effects of Response (short vs. long) at any of the time points across the three stimulus phases (Table 1). By contrast, significant main effects of Interval (short<long) were found from -300 to -160ms prior to stimulus offset, \( F_s(1,28)=4.82-12.55, ps<.020, \eta_p^2=0.15-0.31 \), but not in any other phases. However, there was a clear Response \times\ Interval interaction across all time points beginning at 20ms from stimulus onset, \( F_s(1,28)=5.12-8.78, ps<.031, \eta_p^2=0.16-0.24 \), and continuing through the stimulus pre-offset and stimulus offset phases.

Subsidiary analyses in the stimulus onset phase of the subsecond task revealed that there was a trend for larger pupil diameter when participants gave a long (error) response relative to a short (correct) response for the short intervals, \textit{mean difference (MD)=2.92, g=0.37, 95\% CI [0.06, 0.76]}. This trend continued up to almost 40ms prior to stimulus offset, \( MD=3.81, g=0.36, 95\% CI [0.11, 0.73] \), but was significant only in the stimulus pre-offset phase (-280 to -40ms). The converse trend was also observed for long intervals (greater pupil diameter for short [error] responses than long [correct] responses), but it did not achieve statistical significance following an FDR correction, \( -140\text{ms prior to stimulus pre-offset: } MD=-12.49, g=-0.36, 95\% CI [-0.67, -0.07]; \) stimulus offset to response prompt: \( MD=-15.39, g=-0.38, 95\% CI [-0.68, -0.07] \). For short intervals, there was no specific trend observed in the stimulus offset phase, \( MD=0.13, g=0.00, 95\% CI [-0.33, 0.32] \).
Figure 1. Baseline corrected pupil diameter during stimulus onset, stimulus-pre-offset, and stimulus offset phases of subsecond (top row) and suprasecond (bottom row) temporal bisection tasks as a function of Stimulus interval (short vs. long) and Response (short vs. long). In the three phases, 0 corresponds to stimulus onset (left), offset (middle), and offset (right), respectively. Horizontal bars denote significant (FDR corrected) interval effects (light gray) and interval × response interactions (dark gray).
Table 1

Summary inferential statistics for analyses of variance on pupil diameter ranges for different phases of the subsecond temporal bisection task (N=29)

<table>
<thead>
<tr>
<th>Stimulus phase</th>
<th>F(1,28)</th>
<th>p</th>
<th>η_p²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-stimulus (-1000-0ms)</td>
<td>0.27-1.34</td>
<td>.25-.60</td>
<td>.01-.04</td>
</tr>
<tr>
<td>Stimulus onset (0-300ms)</td>
<td>0.00-0.43</td>
<td>.51-.98</td>
<td>.00-.01</td>
</tr>
<tr>
<td>Stimulus pre-offset (-300-0ms)</td>
<td>0.55-12.55</td>
<td>.001-.46*</td>
<td>.01-.31</td>
</tr>
<tr>
<td>Stimulus offset (0-800ms)</td>
<td>0.00-1.17</td>
<td>.28-.97</td>
<td>.00-.04</td>
</tr>
<tr>
<td><strong>Response (short vs. long)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-stimulus (-1000-0ms)</td>
<td>0.99-2.24</td>
<td>.14-.32</td>
<td>.03-.07</td>
</tr>
<tr>
<td>Stimulus onset (0-300ms)</td>
<td>0.00-0.14</td>
<td>.71-.99</td>
<td>.00-.004</td>
</tr>
<tr>
<td>Stimulus pre-offset (-300-0ms)</td>
<td>0.98-1.65</td>
<td>.21-.33</td>
<td>.03-.05</td>
</tr>
<tr>
<td>Stimulus offset (0-800ms)</td>
<td>1.13-2.79</td>
<td>.10-.29</td>
<td>.04-.09</td>
</tr>
<tr>
<td><strong>Interval × Response</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-stimulus (-1000-0ms)</td>
<td>1.44-4.78</td>
<td>.037-.24</td>
<td>.04-.14</td>
</tr>
<tr>
<td>Stimulus onset (0-300ms)</td>
<td>1.75-8.78</td>
<td>.006-.19*</td>
<td>.06-.24</td>
</tr>
<tr>
<td>Stimulus pre-offset (-300-0ms)</td>
<td>5.46-9.64</td>
<td>.004-.026*</td>
<td>.16-.26</td>
</tr>
<tr>
<td>Stimulus offset (0-800ms)</td>
<td>5.38-9.92</td>
<td>.003-.027*</td>
<td>.16-.26</td>
</tr>
</tbody>
</table>

Notes. Values reported included ranges of Fs, ps, and η_p²s in the respective stimulus phases.
* At least 53.3% or more of p-values are significant after a false discovery rate (FDR) correction, p<.05.

In the suprasecond task, there were no significant main effects of Interval or Response (Table 2 and Figure 1). However, there was a significant interaction in the stimulus offset phase just prior to the response prompt (580-900ms), F(1,25)=6.86-10.2, p<.01, η_p²=0.21-0.29. The direction of the interaction is consistent with that observed in the subsecond task, but it did not survive an FDR correction. In particular, pupil diameter was greater for short than long responses for long intervals, MD=-16.38, g=-0.49, 95% CI [-0.99, -0.13]. For short intervals, there was no specific trend observed, MD=6.25, g=0.15, 95% CI [-0.22, 0.43].
Table 2

*Summary inferential statistics for analyses of variance on pupil diameter ranges for different phases of the suprasecond temporal bisection task (N=26)*

<table>
<thead>
<tr>
<th>Stimulus Phase</th>
<th>F (1,25)</th>
<th>p</th>
<th>η²p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Interval</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-stimulus (-1000-0ms)</td>
<td>0.82-2.05</td>
<td>.16-.37</td>
<td>.03-.07</td>
</tr>
<tr>
<td>Stimulus onset (0-1400ms)</td>
<td>0.00-2.41</td>
<td>.13-.99</td>
<td>.00-.09</td>
</tr>
<tr>
<td>Stimulus pre-offset (-1400-0ms)</td>
<td>0.00-1.19</td>
<td>.28-.95</td>
<td>.00-.04</td>
</tr>
<tr>
<td>Stimulus offset (0-900ms)</td>
<td>0.00-5.01</td>
<td>.034-.99</td>
<td>.00-.16</td>
</tr>
<tr>
<td><strong>Response (short vs. long)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-stimulus (-1000-0ms)</td>
<td>0.81-2.26</td>
<td>.14-.37</td>
<td>.03-.08</td>
</tr>
<tr>
<td>Stimulus onset (0-1400ms)</td>
<td>0.00-1.78</td>
<td>.19-.97</td>
<td>.00-.06</td>
</tr>
<tr>
<td>Stimulus pre-offset (-1400-0ms)</td>
<td>0.25-1.00</td>
<td>.32-.62</td>
<td>.00-.03</td>
</tr>
<tr>
<td>Stimulus offset (0-900ms)</td>
<td>0.00-1.25</td>
<td>.27-.99</td>
<td>.00-.05</td>
</tr>
<tr>
<td><strong>Interval × Response</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-stimulus (-1000-0ms)</td>
<td>0.01-0.59</td>
<td>.44-.91</td>
<td>.00-.02</td>
</tr>
<tr>
<td>Stimulus onset (0-1400ms)</td>
<td>0.00-0.52</td>
<td>.47-.99</td>
<td>.00-.02</td>
</tr>
<tr>
<td>Stimulus pre-offset (-1400-0ms)</td>
<td>0.00-0.83</td>
<td>.37-.99</td>
<td>.00-.03</td>
</tr>
<tr>
<td>Stimulus offset (0-900ms)</td>
<td>0.21-10.21</td>
<td>.003-.64*</td>
<td>.01-.29</td>
</tr>
</tbody>
</table>

*Notes. Values reported included ranges of F’s, ps, and η²p's in the respective stimulus phases.
* At least 33.3% or more of p-values are significant after a false discovery rate (FDR) correction, p<.05.*

A separate series of ANOVAs with Response (Correct vs Error) and Interval (Short vs Long) as independent variables revealed no significant main effects or interactions for the 1000ms pre-stimulus window in the subsecond or suprasecond task (see Supplementary Tables 1 and 2 for inferential statistics). However, as observed in the previous analysis of the subsecond task, there was a significant main effect of Interval in the stimulus pre-offset phase from -300 to -160ms prior to stimulus offset, $F$s(1,28)=6.08-12.55, $p$s<.02, $η²p$s=0.18-0.31, with larger pupil diameter for long compared to short intervals (see Figure 1). In addition, significant main effects of Response were found from 20ms following stimulus onset, $F$s(1,28)=5.12-8.78, $p$s<.03, $η²p$s=0.15-0.24, and continuing through the stimulus pre-offset, $F$s(1,28)=5.46-9.64, $p$s<.026, $η²p$s=0.16-0.25, and stimulus offset phases, $F$s(1,28)=5.37-9.91, $p$s<.027, $η²p$s=0.16-0.26 (Figure 2). Throughout these phases, pupil diameter was reliably larger for error than correct responses. A similar effect, albeit in in a narrower time window, was observed in the suprasecond task. The ANOVAs including Response (Correct vs Error) and Interval (Short vs Long) as independent variables revealed a main effect of Response in the stimulus offset phase (580-900ms), $F$s(1,25)=6.86-10.21, $p$s<.01, $η²p$s=0.21-0.29, with larger pupil diameter observed on error trials.
Figure 2. Pupil diameter at baseline (and baseline-corrected) during pre-stimulus onset, stimulus onset, stimulus pre-offset, and stimulus offset phase of subsecond (top row) and suprasecond (bottom row) temporal bisection tasks as a function of Response (correct vs. error). In the pre-stimulus and stimulus onset phases, 0 corresponds to stimulus onset and in the pre-offset and offset phases, 0 corresponds to stimulus offset. Horizontal bars denote significant (FDR corrected) response effects (dark gray).
**Discussion**

We sought to assess whether, and to what extent, pupillometry can be used to track the subjective perception of time (Suzuki et al., 2016; Suzuki & Tanaka, 2017). Pupil diameter did not reliably vary across stimulus intervals or temporal judgments, which is at odds with previous results pointing to pupil diameter as a potential index of perceived duration (Suzuki et al., 2016). Rather, our results suggest instead that pupil diameter tracks errors in subsecond timing, with larger pupil diameter observed on error trials. Collectively, these results suggest that pupil dilation during interval timing reflects an error monitoring mechanism but that pupil diameter does not robustly track subjective variability in the perception of duration.

The principal result of this study is that pupil diameter was greater when participants made errant temporal judgments regarding subsecond intervals. In particular, we found a reliable pattern of enlarged pupil diameter both when shorter intervals were overestimated (judged to be “long”) and when longer intervals were underestimated (judged to be “short”). This pattern was present both during stimulus presentation (stimulus onset: 20-300ms; stimulus pre-offset: -300-0) as well as after stimulus presentation (stimulus offset: 0-800). However, the temporal locus of these effects varied as a function of stimulus interval: whereas pupil dilation was observed for short interval errors during stimulus presentation it was present during the stimulus pre-offset and offset phases for long interval errors. By contrast, corresponding effects were not reliably observed in a suprasecond temporal bisection task. For example, the corresponding interaction of stimulus interval and perceptual response was replicated only in the late stimulus offset phase just prior to the response prompt (580-900ms) and was driven by differential responses for long stimulus intervals. Taken together, these results suggest that pupil diameter during interval timing reliably indexes subsecond error monitoring mechanisms.

These results partially align with those of Toscano-Zapién et al. (2016), which suggested that pupil dilation predicts incorrect interval judgments. In their study, pupil diameter was found to be larger when participants’ gave long responses for short intervals and short responses for long intervals. More broadly, prior research suggests that pupil diameter is a reliable marker of performance prediction errors (Braem et al., 2015) and is sensitive to different types of errors according to their behavioral relevance (Maier et., 2019). In particular, pupil dilation is observed following difficult correct trials relative to easy correct trials whereas pupil contraction is greater following difficult incorrect trials than easy incorrect trials (Braem et al., 2015). Moreover, error-related pupil dilation has been shown to be larger for perceived than unperceived errors (Wessel et al., 2011). Applied to the present data, and considering the
Pupillometry tracks timing errors

results of Toscano-Zapién et al. (2016), a pupil dilation error monitoring mechanism might have been expected for both short and long stimulus intervals. However, as noted above, pupillary tracking of errors in these interval ranges varied across the phases of stimulus presentation. This potentially suggests that pupillary changes in response to timing errors are not being robustly tracked for shorter subsecond intervals, potentially because these errors are less accessible to awareness. In contrast, it is plausible that pupil dilation was more easily apparent for long subsecond intervals because of the greater ease of detecting errors on such trials.

One another plausible explanation for the difference in pupil diameter observed during the stimulus offset phase with long subsecond intervals is that the difference may have emerged as a consequence of lapses in attention. Previous research demonstrated that relative to self-reported on-task states, mind wandering states were associated with temporal underestimation and increased error rates (Terhune et al., 2017). Mind-wandering is often characterized by enlarged pupil diameter (Pelagatti et al., 2018; Smallwood et al., 2011) and thus the present results potentially reflect in part errors attributable to mind wandering or attentional lapses, such as through perceptual decoupling during mind wandering states. Insofar as the present effect was most pronounced for longer subsecond intervals, this interpretation is arguably further bolstered by the previous finding that increased timing error rates during mind wandering were mostly specific to long subsecond intervals (Terhune et al., 2017). It is possible that timing errors due to attentional lapses are more likely when the response is to be given after a certain amount of time has elapsed. Further research using concurrent eye-tracking and mind wandering state probes during interval timing is required to more rigorously evaluate this possibility.

The pattern of pupil diameter in relation to erroneous interval judgments differed across subsecond and suprasecond intervals. Although the pupil diameter was larger for incorrect interval judgments in the subsecond task, this was present in suprasecond tasks only for the post-stimulus decisional phase prior to response prompts. Similar findings have been reported previously in the perceived duration of gaze shift (Binetti et al., 2017); specifically, pupil dilation associated with temporal judgments in gaze shift significantly differed only for subsecond intervals. By contrast, with suprasecond intervals, recent research suggests that there is no robust association between between pupil size and temporal judgment errors (Suárez-Pinilla et al., 2019). Taken together, our results are consistent with a wealth of evidence for a functional dissociation in the mechanisms subserving subsecond and suprasecond timing
(Lewis & Miall, 2003; Rammsayer & Ulrich, 2011; Hayashi et al., 2014). This work suggests that sensory-based automatic processing underlies the former whereas the latter is supported to a greater extent by executive cognitive processes. In humans, pupillary fluctuations in response to timing errors are more apparent in the subsecond range of intervals which predominantly involves lower level sensory processes.

Although the present results appear to be at odds with previous findings suggesting that pupil diameter covaries with perceived duration (Suzuki et al., 2016), these dissimilar results can potentially be reconciled. In the latter study, monkeys performed an oculomotor version of the temporal reproduction task and saccadic latency after a self-timed interval was used as an index of reproduced duration. Monkeys were trained to generate a saccade within 1000-1700ms of the cue onset, of which latencies below 1000ms constituted nearly 30% of trials. Saccadic latencies were segregated into early, middle, and late latencies, and pupil diameter was compared between early and late latencies. Given that nearly one third of trials included latencies shorter than the minimum trained interval (1000ms) and rarely any latencies were longer than the maximum interval range (1700ms), a large fraction of saccades clustered under the short-latency group included latencies shorter than 1000ms. Those saccades were not impulsive but rather were produced earlier in time, and therefore constitute errors. Should this be the case, the purported effect of increased pupil diameter in shorter latencies, suggesting subjective temporal dilation, might be indicative of an error monitoring mechanism similar to that observed here.

Our findings on interval-specific, error-related pupillary variations indirectly support the possible involvement of norepinephrine modulators in signifying errors related to the perception of duration. As such, it is reasonable to assume that LC-NE might play a key role in signifying errors related to temporal judgments. Given that NE is a key neuromodulator in probing attentional lapses (Smith & Nutt, 1996) and error monitoring mechanisms (Murphy et al., 2011), NE might likely mediate their influence on timing errors. The LC-NE system could possibly have different duration-sensitive channels which may have contributed to the interval-specific difference in pupil size. However, given that there was no control task to test whether the purported effect reflects a temporal error monitoring mechanism or a generic error monitoring mechanism (see, e.g., Coull et al., 2004), more research is required to determine whether the effect observed here is specific to temporal judgment errors or reflective of a broader error monitoring mechanism. Additionally, we did not analyse for any confounding or interacting effect of saccades; nevertheless, it seems unlikely that, trial-by-trial variability in saccades would be a potential confound in the error-related pupil dilation observed here.
Finally, the present data are unable to determine whether the observed effect is specific to the LC-NE system given that pupillometric variation has also been linked to the cholinergic, dopaminergic, and serotonergic systems (Schmid et al., 2015; Reimer et al., 2016). Further research coupling eye tracking and pharmacological interventions targeting these systems (e.g., Coull et al., 2012) is required to further distinguish the roles of these systems in the effects observed here.

In summary, our results suggest that pupil diameter does not track the subjective perception of duration, but instead tracks errors in interval timing. This effect appears to be more specific to the processing of subsecond intervals and was most pronounced for longer subsecond intervals. We propose that LC-NE signaling underlies this pattern of pupillary dilation by its modulatory influence on mechanisms such as error monitoring and attentional lapses.

References
Pupillometry tracks timing errors


Pupillometry tracks timing errors


