

Open your eyes for prediction errors

Senne Braem · Ena Coenen · Klaas Bombeke ·
Marlies E. van Bochove · Wim Notebaert

Published online: 15 January 2015
© Psychonomic Society, Inc. 2015

Abstract Previous studies have demonstrated that autonomic arousal is increased following correct task performance on a difficult, relative to an easy, task. Here, we hypothesized that this arousal response reflects the (relative) surprise of correct performance following a difficult versus an easy task. Following this line of reasoning, we would expect to find a reversed pattern following erroneous responses, because errors are less expected during an easy than during a difficult task. To test this, participants performed a flanker task while pupil size was measured online. As predicted, the results demonstrated that pupil size was larger following difficult (incongruent) correct trials than following easy (congruent) correct trials, but smaller following difficult than following easy incorrect trials. Moreover, participants with larger congruency effects, and hence a larger difference in outcome expectancies between the two trial types, showed larger differences in pupil size after both correct and incorrect responses, further corroborating the idea that pupil size increased as a measure of performance prediction errors.

Keywords Cognitive control · Prediction error · Performance monitoring · Pupil dilation · Norepinephrine

Senne Braem and Ena Coenen contributed equally to the present work.

S. Braem (✉) · E. Coenen · K. Bombeke · W. Notebaert
Department of Experimental Psychology, Ghent University,
Henri Dunantlaan 2, 9000 Ghent, Belgium
e-mail: senne.braem@ugent.be

S. Braem
Department of Experimental Clinical and Health Psychology,
Ghent University, Ghent, Belgium

M. E. van Bochove
Donders Institute for Brain, Cognition and Behaviour,
Radboud University, Nijmegen, The Netherlands

In order to achieve our goals more efficiently, cognitive control processes continuously monitor our environment and adjust information processing. Central to this remarkable feature of human behavior is our ability to constantly make predictions about the environment and our own actions. In light of this idea, recent computational models have suggested that we monitor and act upon these so-called *prediction errors* (Alexander & Brown, 2011; Holroyd & Coles, 2002; Silvetti, Alexander, Verguts, & Brown, 2014; Silvetti, Seurinck, & Verguts, 2011). For example, when you are handed a small gift-wrapped box, you probably expect to find a small present. In this situation, a prediction error will occur when you find either an expensive watch (*positive* prediction error) or a small candy (*negative* prediction error). The abovementioned models suggest that we generate predictions and monitor prediction errors not only for our environment, but also for our own behavior. For example, failing at walking down the stairs (unexpected failure in a simple task) or hitting the bull's eye in darts (unexpected success in a hard task) can also result in prediction errors (i.e., *performance* prediction errors). In the present study, we tested this hypothesis by investigating the interaction between task difficulty and task performance on pupil dilation—as a measure of cognitive surprise (e.g., Nassar et al., 2012; Preuschoff, 't Hart, & Einhäuser, 2011; Ratsig, Welke, Hagendorf, & van der Meer, 2010; Silvetti, Seurinck, van Bochove, & Verguts, 2013).

The role of prediction errors and their influence on behavior is inherited from the reinforcement learning literature (Montague, Hyman, & Cohen, 2004; Sutton & Barto, 1998). The reward prediction error can be conceptualized as the difference between the expected and actually received reward. However, although such errors have been convincingly demonstrated in reward-learning studies (Schultz, 2002, 2004), these reward prediction errors may also be central to how we monitor our task performance in the absence of reward. Along these lines, Holroyd and Coles (2002) stated that the anterior

cingulate cortex (ACC) uses reward and error signals to improve task performance. According to their model, the ACC learns which response is best for a specific task through reinforcement learning signals (i.e., prediction errors). Inspired by the model of Holroyd and Coles, Silvetti and colleagues (2011) further suggested that ACC activity could be summarized by one function—namely, *value estimation*. In this way, Silvetti and colleagues (2011) used their model to explain reward processing, conflict monitoring, error processing, and volatility estimation. What is especially interesting for the present purpose is that both models implicitly assume that people keep track of their performance, and do so for different task conditions or task contexts separately. For example, when presented with difficult and easy tasks, people quickly learn the separate outcome expectancies (i.e., mean accuracy) for both tasks, which will result in different prediction errors during task performance. Consequently, when making an error on an easy task, one would experience a larger (negative) prediction error than during an error on a difficult task. Conversely, being correct on a difficult task would result in a larger (positive) prediction error than would being correct on an easy task.

Using pupil dilation as a marker of cognitive surprise, we set out to investigate this hypothesis in an Eriksen flanker task (Eriksen & Eriksen, 1974). When doing a flanker task, participants have to respond to a centrally presented target, flanked by a number of distractors. These distractors (i.e., flankers) can be either the same as the central symbol (i.e., the target), resulting in a congruent trial (e.g., HHHHH), or different from the target, resulting in an incongruent trial (e.g., HHSHH). Typically, people tend to react faster and make fewer errors on congruent than on incongruent trials. Therefore, on the basis of the abovementioned models, we can make the following three predictions. First, since errors are less frequent than correct trials, there should be a larger pupil size following erroneous responses than following correct performance. Second, correct performance is less expected on incongruent than on congruent trials, whereas errors are less expected on congruent than on incongruent trials. Therefore, pupil size should be larger following correct incongruent trials than following correct congruent trials, but this congruency effect should reverse when errors are being made, resulting in an interaction between congruency and accuracy on pupil dilation. Finally, if this interaction is really driven by performance prediction errors, we would expect this pattern to depend on how individuals experience the difference in outcome expectancies between the two congruency conditions. Therefore, individual differences in the congruency effect (i.e., a larger congruency effect would mean a larger difference in outcome expectancies between congruency conditions) should result in a more pronounced interaction between congruency and accuracy in pupil size.

Method

Participants

Twenty-nine persons participated in this experiment in exchange for either course credits or €15. All participants signed an informed consent disclosure. For optimal online pupil size measures, participants were not allowed to wear glasses or hard lenses. One participant did not make a sufficient number of errors on congruent trials for analyses (only 11; the remaining participants made more than 20 errors). Seven other participants made too many errors, resulting in chance-level performance, probably due to the lack of online feedback during task performance and to difficult task conditions (i.e., short stimulus presentation times and stimulus masks; see below). The absence of online feedback was used to promote self-monitoring of performance, and the difficult task was used to induce sufficient error data on both incongruent and congruent trials for pupil dilation analyses. One extra participant was excluded due to excessive blinking behavior and insufficient pupil data. The remaining 20 participants (16 female, four male) were 21.3 years old ($SD = 4.74$, range = 18–32) and were all right-handed.

Material and procedure

The four-choice flanker task consisted of nine little squares in a 3×3 matrix. Four different types of squares were formed by removing one of the four sides, resulting in differently oriented U-shapes. The target stimulus was the central stimulus, which was always surrounded by eight distractor stimuli. The stimuli were presented in dark gray on a light gray background. The target stimulus could be either the same as the eight distractor stimuli (a congruent trial) or different (an incongruent trial). Importantly, all of the stimuli were equiluminant, ensuring that changes in pupil size could not be attributed to light reflexes or luminance effects. The four middle keys on a keyboard—“F,” “G,” “H,” and “J”—were assigned to the four stimulus types, and the participant’s task was to respond quickly and accurately to the target stimulus with its corresponding response button.

After a practice block of 24 trials, participants performed five blocks of 120 trials. Equal numbers of congruent and incongruent trials were randomly presented. On each trial, a stimulus was presented for 100 ms, followed by a mask (##) for 150 ms (see Fig. 1 for an example of an incongruent trial). Participants could react until 550 ms after mask offset (a total of 800 ms of response time). We used a short stimulus duration and strict response deadline to increase the task difficulty and promote speeded and erroneous responses. In this way, we could study pupil size following correct and erroneous task performance for each congruency condition separately. During the practice block, participants received feedback about

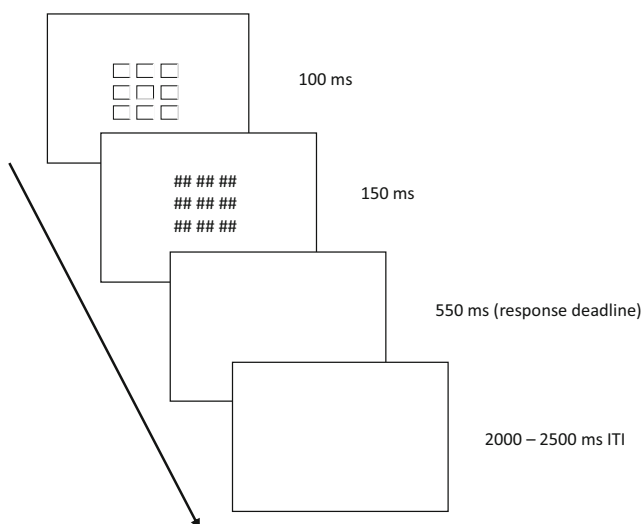


Fig. 1 General paradigm and trial procedure. Participants had to identify the direction of the central figure (target) while ignoring the surrounding shapes, and to respond as quickly as possible. The stimuli were presented in dark gray on a light gray background. No feedback was provided during the experimental trials

their performance: presentation of the word JUIST (“correct”) followed correct responses, FOUT (“wrong”) followed erroneous responses, and TE TRAAG (“too slow”) followed the response deadline when no response was registered. During the experimental blocks, participants only received feedback whenever they were too slow. The intertrial interval (ITI) was jittered between 2,000 and 2,500 ms.

An EyeLink 1000 eyetracking device was used to measure the spontaneous eye blink rate (EBR) and pupil diameter. Before the experiment, the spontaneous eye blink rate was measured for each participant. Participants had to look at a light gray screen with a central fixation cross for 3 min. They were asked not to gaze, but just to look casually at the fixation cross. The EBR is believed to be a measurement of tonic dopamine (Taylor et al., 1999). During the experiment, participants were requested to blink less than usual, but not to refrain from blinking (they were encouraged to blink during breaks). Calibration and validation of gaze position were carried out with a 9-point grid. Viewing was binocular throughout the experiment, but pupil dilation was recorded for the right eye only. A chinrest and a brace at forehead height were used to restrict head movements. Participants had to look at the computer screen through a pane that reflected their eyes into the camera, and they were not allowed to move their head during the entire experiment. The flanker task was presented using Tscope software (Stevens, Lammertyn, Verbruggen, & Vandierendonck, 2006) on a Pentium PC. After the experiment, participants completed the BIS/BAS questionnaire (Carver & White, 1994), measuring their reward and punishment sensitivity. However, none of the correlation analyses with BIS/BAS measures or spontaneous EBR reached significance

after controlling for multiple comparisons, and therefore are not discussed further.

Results

Behavioral results

Trials with no registered responses (within the 800-ms deadline) were excluded from all analyses, which did not differ between congruent (5.2 %) and incongruent (5.4 %) trials, $t(19) < 1$. On the remaining trials, we found a mean accuracy of 72.7 % (note that chance level was 25 %). As expected, an overall congruency effect emerged for both error rates (Fig. 2b), $t(19) = 3.1$, $p < .01$, and reaction times, $t(19) = 9.4$, $p < .01$. Participants made more errors and were slower on incongruent (36.8 % and 640 ms, respectively) than on congruent (18.0 % and 604 ms) trials.

Pupil measures

Pupil size was measured at 1000 Hz. Blinks and missing data points due to recording failure were corrected for by means of a linear interpolation procedure, which allowed us to use all relevant trials for the analyses. However, removing trials with blinks from the analyses rendered similar results. Next, the mean pupil size during a 200-ms time window before stimulus onset was subtracted from the maximum pupil size within a 2,000-ms time window following stimulus onset, ensuring a baseline-corrected measure of pupil size for each trial separately.¹ Using these measures, the mean (maximum) pupil size was calculated for each congruency and accuracy condition. Finally, pupil size was analyzed using a 2×2 ANOVA with Congruency (congruent or incongruent) and Accuracy (correct or error) as within-subjects factors.

The main effect of accuracy was significant, $F(1, 19) = 9.6$, $p < .01$, indicating enhanced pupil dilation following errors (235 a.u.), as opposed to following correct (198 a.u.) trials. The main effect of congruency was not significant, $F(1, 19) < 1$, $p = .810$. However, most importantly, the two-way interaction between accuracy and congruency was significant, $F(1, 19) = 11.4$, $p < .01$ (see Fig. 2a and c). Post-hoc t tests between the two congruency conditions for each accuracy condition

¹ Because the stimulus-locked analysis and graph offers the earliest unbiased onset and most comprehensive visualization of our data, and because the peak-based analysis is most commonly used, we report this analysis. However, there are few standardized guidelines for analyzing pupil dilation data. Therefore, we also did similar analyses that were response-locked. Furthermore, for both the stimulus-locked and response-locked data, peak-based (maximum) as well as mean pupil size analyses were performed. Importantly, all of these analyses resulted in similar significant results.

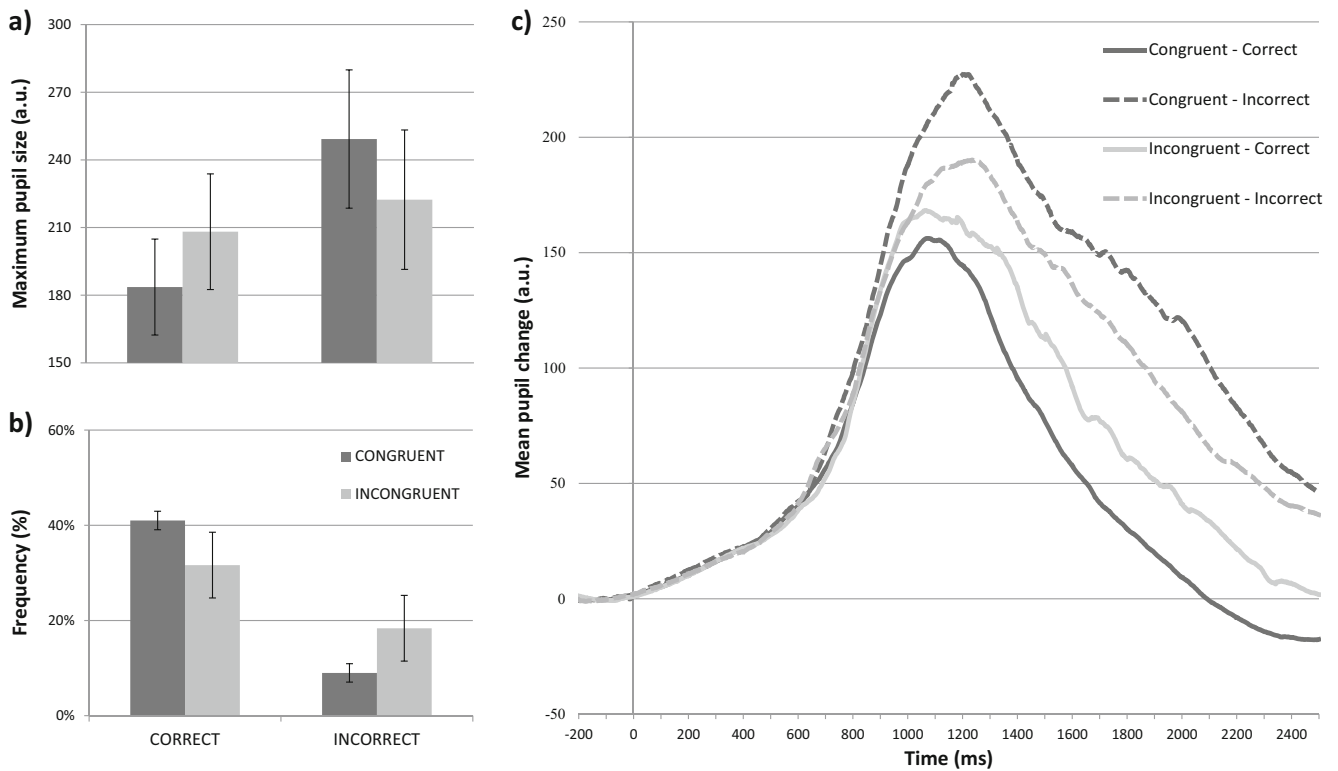


Fig. 2 (a) Maximum pupil size increases within the 2,000 ms following stimulus onset (baseline-corrected to 200 ms before stimulus onset) on congruent (dark gray) and incongruent (light gray) trials, for correct and erroneous responses separately. (b) Frequency of each Congruency by

Accuracy event across all registered responses. (c) Time courses of pupil size from 200 ms before until 2,500 ms after the stimulus onset for each accuracy and congruency condition separately. Error bars represent ± 1 standard error

separately indicated significantly larger pupil dilation on incongruent (208 a.u.) than on congruent (184 a.u.) trials following correct responses, $t(19) = 2.9, p < .01$, and larger pupil dilation on congruent (249 a.u.) than on incongruent (222 a.u.) trials following erroneous responses, $t(19) = 2.8, p = .011$.

Pupil-behavior correlations

The number of errors attributable to incongruent trials (i.e., a relative measure of the congruency effect on error rates that controls for individual differences in overall accuracy) was calculated for each participant separately with the following formula: $[P(\text{Error} | \text{Incongruent}) - P(\text{Error} | \text{Congruent})] / [P(\text{Error} | \text{Incongruent}) + P(\text{Error} | \text{Congruent})]$. Next, this normalized congruency effect was correlated with the two-way interaction observed in the pupil data [computed by subtracting the congruency effect for erroneous responses from the congruency effect for correct responses: (incongruent correct – congruent correct) – (incongruent incorrect – congruent incorrect)]. As predicted, the congruency effect correlated positively with the two-way interaction between congruency and accuracy in the pupil dilation data (see Fig. 3a), as was indicated by both Pearson’s $r, r = .684, p < .01$, and the rank-ordered Spearman’s $\rho, \rho = .621, p < .01$.

This relation between response event probability (per Congruency \times Accuracy condition) and pupil size is also illustrated by Fig. 2a and b. To further demonstrate that event probability and pupil dilation were proportionally, and not just

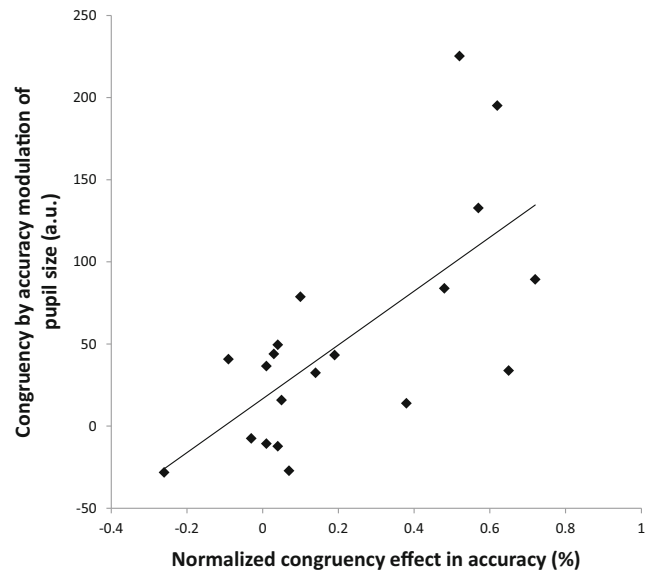


Fig. 3 Correlations between the normalized congruency effect in accuracy and the two-way interaction between congruency and accuracy in pupil size

qualitatively, related, we tested the correlation coefficients for each of the event log probabilities and its corresponding mean pupil size (normalized per participant). Specifically, the Pearson's r correlation coefficients were $r = -.696$, $p < .01$, for congruent correct, $r = -.589$, $p < .01$, for incongruent correct, $r = -.275$, $p > .1$, for congruent incorrect, and $r = -.788$, $p < .001$, for incongruent incorrect trials. Three of the four correlations demonstrate a significant negative correlation between event probability and pupil dilation, in line with our hypothesis that the pupil dilates as a function of the unexpectedness of response outcomes. The nonsignificant correlation for congruent incorrect trials was numerically in the same direction (and possibly nonsignificant because this least frequent condition, naturally, had the least number of data points—i.e., half the number, as compared to incongruent incorrect trials).

Discussion

In this study, we investigated the interaction between congruency and accuracy on pupil dilation. The results indicate that pupil size is increased on incongruent, relative to congruent, trials during correct task performance. Crucially, this pattern reversed during error responses, resulting in larger pupil size on congruent, relative to incongruent, trials. Furthermore, this two-way interaction between congruency and accuracy in pupil size correlated strongly with individual differences in the congruency effect on error rates.

Whereas these differences in pupil size between correct and incorrect performance (Critchley, Tang, Glaser, Butterworth, & Dolan, 2005; Wessel, Danielmeier, & Ullsperger, 2011) and between correct congruent and correct incongruent trials (Brown et al., 1999; Laeng, Ørbo, Holmlund, & Miozzo, 2011; Siegle, Steinhauer, & Thase, 2004; van Bochove, Van der Haegen, Notebaert, & Verguts, 2013; van Steenbergen & Band, 2013) have been well documented before, we are not aware of empirical studies that have investigated the interaction between accuracy and congruency. We believe that there are two important reasons for this. First, earlier experiments did not obtain a sufficient number of errors on congruent trials to investigate this. Here, we used a strict response deadline and short stimulus presentation time to ensure sufficient numbers of errors in both congruency conditions. This allowed us to study pupil size following correct and erroneous task performance for each congruency condition separately.

A second reason is more theoretical: Earlier studies did not investigate this prediction because their measurement of pupil dilation was used for other reasons. Specifically, it was originally proposed that pupil size could be used as a measure of cognitive effort, since it is often found to increase with increasing task demands (Hess & Polt, 1964; Kahneman & Beatty, 1966). Similarly, in contrast to the recent models of

Silvetti and colleagues (2011) and Alexander and Brown (2011), earlier models of conflict-related neural activity had focused on its effortful processing demands (Botvinick, Braver, Barch, Carter, & Cohen, 2001), promoting hypotheses in terms of cognitive effort, rather than cognitive surprise. Importantly, our study did not intend to dissociate the different interpretations of this univariate psychophysiological measure. The interpretation of pupil dilation in terms of cognitive effort could perhaps account for our present results as well, although we believe such an explanation might be less straightforward. For instance, it is not evident whether and how making an error could be understood as being more effortful, and why erroneous responses are more cognitively demanding on congruent than on incongruent trials. Instead, we would argue that the role of prediction errors in task performance processing offers a more promising avenue for future studies of conflict and/or error processing and their related neural signatures. In this study, analyzing prediction errors led to the straightforward prediction that the earlier observed difference between incongruent and congruent correct trials (Brown et al., 1999; Laeng et al., 2011; Siegle et al., 2004; van Bochove et al., 2013; van Steenbergen & Band, 2013) should reverse when making an error. This suggests that this earlier found difference can be reinterpreted in terms of (positive) cognitive surprise following correct performance on incongruent, relative to congruent, trials.

Corroborating evidence for this idea can be found in a recent study by Schouppe and colleagues (in press). There, participants had to perform a flanker task (Exp. 2a) followed by an affective judgment task with positive and negative words. Interestingly, the authors demonstrated that correct performance on incongruent, relative to congruent, trials led to a significant benefit in reaction times for the evaluation of positive, relative to negative, words. Similarly, the authors predicted and interpreted this finding by suggesting that people find it more positively surprising to solve a difficult than an easy task (for a similar reasoning, see Alessandri, Darceville, Delevoye-Turrell, & Zentall, 2008; Braem, Verguts, Roggeman, & Notebaert, 2012; Satterthwaite et al., 2012).

More broadly, our results demonstrate that pupil dilation can act as a marker of cognitive surprise—not only about external events outside a participant's control (e.g., Preuschoff et al., 2011), but also about his or her own performance. Interestingly, similar observations have been made in the electroencephalography (EEG) and functional magnetic resonance imaging (fMRI) literatures. For example, studying the feedback-related negativity, a presumed marker of ACC activity following performance errors, Oliveira, McDonald, and Goodman (2007) demonstrated that this component was not exclusive to error performance, but could be elicited by unexpected positive feedback, as well (for similar results, see Ferdinand, Mecklinger, Kray, & Gehring, 2012; Jessup, Busemeyer, & Brown, 2010; Silvetti, Nuñez Castellar, Roger,

& Verguts, 2014). In a similar vein, Wessel, Danielmeier, Morton, and Ullsperger (2012) took a different approach, by demonstrating that the neural correlates of error processing also show remarkable similarities to those of novelty processing, as evidenced by both EEG and fMRI data (see also Desmet, Deschrijver, & Brass, 2014). Together, these and our results suggest that neural indices of performance monitoring can be best understood in terms of more general processes that signal the violation of expectancies (prediction errors), in line with recent models of performance monitoring (Alexander & Brown, 2011; Silvetti, Alexander et al., 2014; Silvetti et al., 2011).

These similarities between ACC activity during performance monitoring and the present pupil dilation results seem to suggest that our data tap into ACC–locus coeruleus interactions (Aston-Jones & Cohen, 2005). The locus coeruleus is mainly known for its role in norepinephrine release and the orienting response, thought to be marked by the widening of pupils and an event-related electrophysiological component called the P3 (De Taeye et al., 2014; Geva, Zivan, Warsha, & Olchik, 2013; Murphy, Robertson, Balsters, & O’Connell, 2011; Nieuwenhuis, Aston-Jones, & Cohen, 2005; Nieuwenhuis, De Geus, & Aston-Jones, 2011). This orienting response is typically elicited by infrequent and motivationally significant stimuli and serves to facilitate further behavioral adaptations (Lynn, 1966). Performance prediction errors, we believe, are infrequent and motivationally significant in and of themselves. Therefore, the present results might not just reflect prediction errors, but rather the orienting response that follows these prediction errors, in order to facilitate further task performance or learning strategies.

Indeed, other theories of pupil dilation that similarly stress the role of surprise and the locus coeruleus–norepinephrine system on pupil size (Aston-Jones & Cohen, 2005; Gilzenrat, Nieuwenhuis, Jepma, & Cohen, 2010; Jepma & Nieuwenhuis, 2011) go one step further, by emphasizing its importance in task engagement and cognitive adaptation. Specifically, these studies have looked at tonic changes in pupil size (where our study focused on stimulus-evoked phasic changes) and suggested that pupil size decreases as task engagement increases. Conversely, increases in pupil size would be associated with task disengagement or decreases in task utility. In this regard, it remains an important research endeavor to identify what roles prediction errors, pupil dilation, and autonomic arousal might have in driving cognitive adaptations and strategies that serve performance optimization (Aston-Jones & Cohen, 2005; Brown, Van Steenbergen, Kedar, & Nieuwenhuis, 2014; Nassar et al., 2012; Silvetti et al., 2013; Verguts & Notebaert, 2009).

Author note S.B., E.C., and W.N. were supported by FWO-Vlaanderen G.0098.09N. We thank Joshua Brown and an anonymous reviewer for interesting suggestions to improve both the analyses and discussion of our results.

References

- Alessandri, J., Darcheville, J.-C., Delevoeye-Turrell, Y., & Zentall, T. R. (2008). Preference for rewards that follow greater effort and greater delay. *Learning & Behavior*, *36*, 352–358. doi:10.3758/LB.36.4.352
- Alexander, W. H., & Brown, J. W. (2011). Medial prefrontal cortex as an action–outcome predictor. *Nature Neuroscience*, *14*, 1338–1344. doi:10.1038/nn.2921
- Aston-Jones, G., & Cohen, J. D. (2005). An integrative theory of locus coeruleus–norepinephrine function: Adaptive gain and optimal performance. *Annual Review of Neuroscience*, *28*, 403–450. doi:10.1146/annurev.neuro.28.061604.135709
- Botvinick, M. M., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*, 624–652. doi:10.1037/0033-295X.108.3.624
- Braem, S., Verguts, T., Roggeman, C., & Notebaert, W. (2012). Reward modulates adaptations to conflict. *Cognition*, *125*, 324–332. doi:10.1016/j.cognition.2012.07.015
- Brown, G. G., Kindermann, S. S., Siegle, G. J., Granholm, E., Wong, E. C., & Buxton, R. B. (1999). Brain activation and pupil response during covert performance of the Stroop Color Word task. *Journal of the International Neuropsychological Society*, *5*, 308–319.
- Brown, S. B., Van Steenbergen, H., Kedar, T., & Nieuwenhuis, S. (2014). Effects of arousal on cognitive control: Empirical tests of the conflict-modulated Hebbian-learning hypothesis. *Frontiers in Human Neuroscience*, *8*, 23. doi:10.3389/fnhum.2014.00023
- Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS scales. *Journal of Personality and Social Psychology*, *67*, 319–333. doi:10.1037/0022-3514.67.2.319
- Critchley, H. D., Tang, J., Glaser, D., Butterworth, B., & Dolan, R. J. (2005). Anterior cingulate activity during error and autonomic response. *NeuroImage*, *27*, 885–895.
- Desmet, C., Deschrijver, E., & Brass, M. (2014). How social is error observation? The neural mechanisms underlying the observation of human and machine errors. *Social Cognitive and Affective Neuroscience*, *9*, 427–435. doi:10.1093/scan/nst002
- De Taeye, L., Vonck, K., van Bochove, M., Boon, P., Van Roost, D., Mollet, L., & Raedt, R. (2014). The P3 event-related potential is a biomarker for the efficacy of vagus nerve stimulation in patients with epilepsy. *Neurotherapeutics*, *11*, 612–622.
- Eriksen, B. A., & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task. *Perception & Psychophysics*, *16*, 143–149. doi:10.3758/BF03203267
- Ferdinand, N. K., Mecklinger, A., Kray, J., & Gehring, W. J. (2012). The processing of unexpected positive response outcome in the medial frontal cortex. *Journal of Neuroscience*, *32*, 12087–12092.
- Geva, R., Zivan, M., Warsha, A., & Olchik, D. (2013). Alerting, orienting or executive attention networks: Differential patterns of pupil dilations. *Frontiers in Behavioral Neuroscience*, *7*, 145.
- Gilzenrat, M. S., Nieuwenhuis, S., Jepma, M., & Cohen, J. D. (2010). Pupil diameter tracks changes in control state predicted by the adaptive gain theory of locus coeruleus function. *Cognitive, Affective, & Behavioral Neuroscience*, *10*, 252–269. doi:10.3758/CABN.10.2.252
- Hess, E. H., & Polt, J. M. (1964). Pupil size in relation to mental activity during simple problem-solving. *Science*, *143*, 1190–1192.
- Holroyd, C. B., & Coles, M. G. H. (2002). The neural basis of human error processing: Reinforcement learning, dopamine, and the error-related negativity. *Psychological Review*, *109*, 679–709. doi:10.1037/0033-295X.109.4.679
- Jepma, M., & Nieuwenhuis, S. (2011). Pupil diameter predicts changes in the exploration–exploitation trade-off: Evidence for the adaptive gain theory. *Journal of Cognitive Neuroscience*, *23*, 1587–1596.

- Jessup, R. K., Busemeyer, J. R., & Brown, J. W. (2010). Error effects in anterior cingulate cortex reverse when error likelihood is high. *Journal of Neuroscience*, *30*, 3467–3472.
- Kahneman, D., & Beatty, J. (1966). Pupil diameter and load on memory. *Science*, *154*, 1583–1585.
- Laeng, B., Ørbo, M., Holmlund, T., & Miozzo, M. (2011). Pupillary Stroop effects. *Cognitive Processing*, *12*, 13–21.
- Lynn, R. (1966). *Attention, arousal, and the orientation reaction*. Oxford, UK: Pergamon Press.
- Montague, P. R., Hyman, S. E., & Cohen, J. D. (2004). Computational roles for dopamine in behavioural control. *Nature*, *431*, 760–767.
- Murphy, P. R., Robertson, I. H., Balsters, J. H., & O'Connell, R. G. (2011). Pupillometry and P3 index the locus coeruleus–noradrenergic arousal function in humans. *Psychophysiology*, *48*, 1532–1543.
- Nassar, M. R., Rumsey, K. M., Wilson, R. C., Parikh, K., Heasly, B., & Gold, J. I. (2012). Rational regulation of learning dynamics by pupil-linked arousal systems. *Nature Neuroscience*, *15*, 1040–1046.
- Nieuwenhuis, S., Aston-Jones, G., & Cohen, J. D. (2005). Decision making, the P3, and the locus coeruleus–norepinephrine system. *Psychological Bulletin*, *131*, 510–532. doi:10.1037/0033-2909.131.4.510
- Nieuwenhuis, S., De Geus, E. J., & Aston-Jones, G. (2011). The anatomical and functional relationship between the P3 and autonomic components of the orienting response. *Psychophysiology*, *48*, 162–175.
- Oliveira, F. T., McDonald, J. J., & Goodman, D. (2007). Performance monitoring in the anterior cingulate is not all error related: Expectancy deviation and the representation of action–outcome associations. *Journal of Cognitive Neuroscience*, *19*, 1994–2004.
- Preuschoff, K., 't Hart, B. M., & Einhäuser, W. (2011). Pupil dilation signals surprise: Evidence for noradrenaline's role in decision making. *Frontiers in Neuroscience*, *5*, 115. doi:10.3389/fnins.2011.00115
- Raisig, S., Welke, T., Hagendorf, H., & van der Meer, E. (2010). I spy with my little eye: Detection of temporal violations in event sequences and the pupillary response. *International Journal of Psychophysiology*, *76*, 1–8.
- Satterthwaite, T. D., Ruparel, K., Loughhead, J., Elliott, M. A., Gerraty, R. T., Calkins, M. E., & Wolf, D. H. (2012). Being right is its own reward: Load and performance related ventral striatum activation to correct responses during a working memory task in youth. *NeuroImage*, *61*, 723–729. doi:10.1016/j.neuroimage.2012.03.060
- Schoupe, N., Braem, S., De Houwer, J., Silvetti, M., Verguts, T., Ridderinkhof, K. R., & Notebaert, W. (2015). No pain, no gain: The affective valence of congruency conditions changes following a successful response. *Cognitive, Affective, & Behavioral Neuroscience*. doi:10.3758/s13415-014-0318-3
- Schultz, W. (2002). Getting formal with dopamine and reward. *Neuron*, *36*, 241–263.
- Schultz, W. (2004). Neural coding of basic reward terms of animal learning theory, game theory, microeconomics and behavioural ecology. *Current Opinion in Neurobiology*, *14*, 139–147.
- Siegle, G. J., Steinhauer, S. R., & Thase, M. E. (2004). Pupillary assessment and computational modeling of the Stroop task in depression. *International Journal of Psychophysiology*, *52*, 63–76. doi:10.1016/j.ijpsycho.2003.12.010
- Silvetti, M., Alexander, W., Verguts, T., & Brown, J. W. (2014). From conflict management to reward-based decision making: Actors and critics in primate medial frontal cortex. *Neuroscience & Biobehavioral Reviews*, *46*, 44–57. doi:10.1016/j.neubiorev.2013.11.003
- Silvetti, M., Nuñez Castellar, E., Roger, C., & Verguts, T. (2014). Reward expectation and prediction error in human medial frontal cortex: An EEG study. *NeuroImage*, *84*, 376–382.
- Silvetti, M., Seurinck, R., van Bochove, M. E., & Verguts, T. (2013). The influence of the noradrenergic system on optimal control of neural plasticity. *Frontiers in Behavioral Neuroscience*, *7*, 160. doi:10.3389/fnbeh.2013.00160
- Silvetti, M., Seurinck, R., & Verguts, T. (2011). Value and prediction error in medial frontal cortex: Integrating the single-unit and systems levels of analysis. *Frontiers in Human Neuroscience*, *5*, 75. doi:10.3389/fnhum.2011.00075
- Stevens, M., Lammertyn, J., Verbruggen, F., & Vandierendonck, A. (2006). Tscope: AC library for programming cognitive experiments on the MS Windows platform. *Behavior Research Methods*, *38*, 280–286. doi:10.3758/BF03192779
- Sutton, R. S., & Barto, A. G. (1998). *Introduction to reinforcement learning*. Cambridge, MA: MIT Press.
- Taylor, J. R., Elsworth, J. D., Lawrence, M. S., Sladek, J. R., Roth, R. H., & Redmond, D. E. (1999). Spontaneous blink rates correlate with dopamine levels in the caudate nucleus of MPTP-treated monkeys. *Experimental Neurology*, *158*, 214–220.
- van Bochove, M. E., Van der Haegen, L., Notebaert, W., & Verguts, T. (2013). Blinking predicts enhanced cognitive control. *Cognitive, Affective, & Behavioral Neuroscience*, *13*, 346–354. doi:10.3758/s13415-012-0138-2
- van Steenbergen, H., & Band, G. P. H. (2013). Pupil dilation in the Simon task as a marker of conflict processing. *Frontiers in Human Neuroscience*, *7*, 215. doi:10.3389/fnhum.2013.00215
- Verguts, T., & Notebaert, W. (2009). Adaptation by binding: A learning account of cognitive control. *Trends in Cognitive Sciences*, *13*, 252–257. doi:10.1016/j.tics.2009.02.007
- Wessel, J. R., Danielmeier, C., & Ullsperger, M. (2011). Error awareness revisited: Accumulation of multimodal evidence from central and autonomic nervous systems. *Journal of Cognitive Neuroscience*, *23*, 3021–3036. doi:10.1162/jocn.2011.21635
- Wessel, J. R., Danielmeier, C., Morton, J. B., & Ullsperger, M. (2012). Surprise and error: Common neuronal architecture for the processing of errors and novelty. *Journal of Neuroscience*, *32*, 7528–7537. doi:10.1523/JNEUROSCI.6352-11.2012