Pupil dilation offers a time-window on prediction error

Authors: Olympia Colizoli[1]*, Tessa van Leeuwen[1,2], Danaja Rutar[1,3], Harold Bekkering[1]

Affiliations:
[1] Donders Institute for Brain, Cognition and Behaviour, Radboud University, Nijmegen, The Netherlands
[2] Department of Communication and Cognition, Tilburg University, Tilburg, The Netherlands
[3] Department of Psychology, Sigmund Freud University, Ljubljana, Slovenia

Corresponding Author:
*Olympia Colizoli
Radboud University
Donders Institute for Brain, Cognition and Behaviour
Thomas van Aquinostraat 4
6525 GD Nijmegen, The Netherlands
Email: olympia.colizoli@donders.ru.nl

Materials & Correspondence: olympia.colizoli@donders.ru.nl (O.C.)

Abstract: 240 words
Task-evoked pupil dilation has been linked to many cognitive variables, perhaps most notably unexpected events. Zénon (2019) proposed a unifying framework stating that pupil dilation related to cognition should be considered from an information-theory perspective. In the current study, we investigated whether the pupil's response to decision outcome in the context of associative learning reflects a prediction error defined formally as information gain, while also exploring the time course of this prediction error signal. To do so, we adapted a simple model of trial-by-trial learning of stimulus probabilities based on information theory from previous literature. We analyzed two data sets in which participants performed perceptual decision-making tasks that required associative learning while pupil dilation was recorded. Our findings consistently showed that a significant proportion of variability in the post-feedback pupil response during decision-making can be explained by a formal quantification of information gain shortly after feedback presentation in both task contexts. In the later time window, the relationship between information-theoretic variables and the post-feedback pupil response differed per task. For the first time, we present evidence that whether the pupil dilates or constricts along with information gain seems to be context dependent, specifically related to increasing or decreasing average uncertainty (entropy) across trials. This study offers empirical evidence showcasing how the pupil's response can offer valuable insights into the process of model updating during learning, highlighting the promising utility of this readily accessible physiological indicator for investigating internal belief states.

Keywords: Prediction error, pupillometry, associative learning, information theory, decision making

Acknowledgments: We would like to thank and acknowledge Yibrán Amador Pacheco Sáez, Filip Novický, Francesco Poli, Luke Miller, Lieke van Lieshout, and the Karl Friston Lab for helpful insight and discussions at various stages of the research presented here.

Author contributions: O.C. contributed Conceptualization, Methodology, Investigation, Formal Analysis, Visualization, Writing - original draft, Writing – review and editing; T.v.L. contributed Conceptualization, Methodology, Investigation, Writing - review and editing; D.R.

contributed Methodology, Investigation, Writing - review and editing; H.B. contributed Conceptualization, Supervision, Writing - review and editing.

Funding: This research was funded by the Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, The Netherlands.

Conflict of interest disclosure: The authors declare no competing interests.

Data and code availability: The analysis code is publicly available (https://github.com/colizoli/pupil_associative_learning). Upon final publication in a peer-reviewed journal, all code and data will be publicly available and be given a unique digital object identifier (DOI).

Ethics approval: All experiments were conducted on human participants and were approved by the Ethical Committee of the Faculty of Social Sciences at the Radboud University, the Netherlands.

1. Introduction

The human brain is constantly predicting its environment by forming expectations based on the meaningful environmental statistics we encounter in life (e.g., bananas are yellow, sometimes green, but never blue or red)[1–3]. Furthermore, perception is largely driven by these expectations[4,5]. In the predictive processing framework, the brain constructs and updates internal models of the world that aim to accurately predict incoming sensory data by minimizing prediction errors[6]. Prediction errors are abstractly defined as the difference between an expected and obtained outcome and are crucial concepts in models of learning[7–9]. Pupil dilation may be a reliable biomarker marker of neural prediction error signals. Such a biomarker would be advantageous for investigating the neural computations involved in learning and decision-making due to the relative ease of measuring pupil size with a standard eye-tracking device. To achieve this aim, we must investigate the computational signatures reflected in pupil dilation to determine whether they genuinely represent prediction error signals during learning.

While reaction times (RT) scale with uncertainty, confidence, and reward expectation during perceptual decision-making[10–13], there is no overt behavioral marker that reflects the brain's processing of a prediction error following *feedback* on a decision outcome when the brain supposedly updates its internal model(s) of the world[6]. Pupil dilation under constant luminance is a peripheral marker of the brain's central arousal system[14–18]. Evidence suggests that the brain leverages its arousal system to relay computational variables to circuits that execute inference and action selection[8,14,19–24]. A growing body of literature suggests that pupil dilation may be a reliable physiological marker reflecting a prediction error signal in the post-feedback pupil response[10,11,25–33]. Whether a prediction error signal is detectable in pupil dilation will depend on the specific definition of "prediction error", how it is operationalized, as well as the temporal nature of the signal itself. For instance, an unsigned prediction error is defined as the difference between expected and unexpected events but is agnostic to whether the events are better or worse than expected. In contrast, a signed prediction error indicates whether an outcome was better or worse than expected. Relatedly, reward prediction errors are a type of signed prediction error indicating whether an obtained reward was better or worse than expected[7].

In recent years, pupil dilation has received increased attention in psychology and human neuroscience research for its ability to reflect cognitive and computational variables involved in memory, attention, perception, and decision-making. For instance, pupil dilation has been shown to reflect stimulus expectancy and surprise[11,29,30,34–48], decision uncertainty[10,13,28,29,32,36,49–57], the updating of belief states in internal models[10,11,25–32,58–60], and reward prediction errors[10,32]. In 2019, Zénon[61] proposed that the plethora of cognitive phenomena reflected in pupil dilation can be unified under an information-theoretic framework. Under Zénon's hypothesis, the common factor driving all cognitive processes reflected in pupil dilation can be quantified in terms of information gain. In information theory, information gain quantifies the reduction of uncertainty about a random variable given the knowledge of another variable. In other words, information gain measures how much knowing about one variable improves the prediction or understanding of another variable. A unified framework that relates pupil dilation to cognition through information theory would be beneficial for several reasons. First, a unified framework would enable us to more accurately quantify cognitive processes by allowing us to connect physiological responses like arousal with cognitive functions. Second, such an approach could reveal how effectively an individual integrates new information and adjusts their predictions. Finally, a unified framework would allow researchers to apply consistent metrics across different contexts and tasks, facilitating comparisons between studies and enhancing our overall understanding of cognitive processes linked to pupil dilation.

If pupil dilation reflects information gain during the observation of an outcome event, such as feedback on decision accuracy, then pupil size will be expected to increase in proportion to how much novel sensory evidence is used to update current beliefs[29,61].

Information gain can be operationalized within information theory as Kullback-Leibler (KL) divergence (sometimes termed "Bayesian surprise")[29,62]. KL divergence is a measure of the difference between two probability distributions. In the context of predictive processing, KL divergence can be used to quantify the mismatch between the probability distributions corresponding to the brain's expectations about incoming sensory input and the actual sensory input received, in other words, the prediction error[63,64]. KL divergence is thereby a formalized quantity of the amount of information gained during model updating based on both the sensory expectations and sensory evidence[29]. To our knowledge, only a few studies to date have directly investigated whether the pupil response was correlated with information gain defined as KL divergence, and even fewer studies have specifically focused in the interval following outcome feedback presentation. Using a saccadic planning task, O'Reilly et al. (2013)[29] found that pupil dilation scaled negatively with KL divergence in the interval 0.73-1.24 s following target onset. While a signification correlation between post-feedback pupil dilation and information gain was obtained, the direction of this result seems at odds with the hypothesis that an increase in information gain would lead to greater pupil dilation. It was suggested then by O'Reilly et al. that the direction of the scaling between the post-feedback pupil response and KL divergence might depend on the effect of uncertainty (i.e., entropy) on learning, an issue we will return to in the Discussion. Testing infants, Kayhan et al. (2019)[26] used a statistical learning paradigm and recorded pupil dilation during the event outcome interval in each trial. Kayhan et al. reported that pupil dilation responses of 24-month-old (but not 18-month-old) infants followed a pattern consistent with a prediction error calculated as KL divergence between the prior and posterior probability distributions. However, the time course of the pupil response of the infants was not reported nor was it directly correlated to the Bayesian model parameters.

Different types of prediction errors, for example, by being linked with a reward, may show distinct computational signatures in the brain's neurophysiology. Furthermore, the temporal dynamics of prediction error signals in pupil dilation are likely informative because the brain's process of updating internal models may contain an inherent temporal dimension[65]. Different temporal components of the pupil signal may correspond to different stages of predictive processing, for instance, as proposed by the hybrid predictive coding model[66]. These temporal dynamics can shed light on the mechanisms of predictive processing by clarifying the timing of learning updates in relation to feedback presentation[67]. The temporal dynamics of prediction error signals may also help differentiate between types of learning, indicate how attention and cognitive load are allocated during tasks, and implicate specific brain regions or neural processes involved in learning[68,69].

Previous studies have shown different temporal response dynamics of prediction error signals in pupil dilation following feedback on decision outcome. Similar to the time window obtained by O'Reilly et al., de Gee et al. (2021)[11] found that the mean pupil dilation in a window 0.5-1.5 s following feedback reflected "surprise" about choice outcome, defined as an interaction between confidence and accuracy, during an orientation discrimination task. The results of these two studies suggest that the prediction error signals arise around the peak (~1 s) of the canonical impulse response function of the pupil[70,71]. Other studies have shown evidence that prediction error signals arise considerably later than the pupil's canonical peak with respect to feedback on choice outcome. For instance, Browning et al. (2015)[25] found that the Shannon surprise about decision outcome on pupil dilation was significant within a 1.5-2.5 s interval during an aversive learning task. In a reinforcement learning paradigm, van Slooten et al. (2018)[32] found that early pupil dilation scaled with uncertainty about the value of the upcoming reward until about 2 s following feedback, whereas later pupil dilation (from about 2-3 s) scaled with signed reward prediction errors. By implementing long inter-stimulus intervals administered during functional magnetic resonance imaging, Colizoli et al. (2018)[10] found that a signed prediction error signal arose in pupil dilation starting at 3 s following reward-linked feedback and was sustained until 6 s in a motion discrimination task. A relatively slower prediction error signal following feedback presentation may suggest deeper cognitive

processing, increased cognitive load from sustained attention or ongoing uncertainty, or that the brain is integrating multiple sources of information before updating its internal model.

Taken together, the literature indicates that prediction error signals in pupil dilation were obtained anywhere between 0.5 to 6 s following feedback on decision outcome, and the results do not converge to produce a consistent temporal signature. It is clear that the specific time window analyzed across different tasks can affect whether a prediction error signal is detected at all. The emergence of consistent results is important for validating the pupil as a biomarker of prediction error, by facilitating comparative research, informing predictive models, uncovering neural mechanisms, as well as improving practical applications. Many factors could potentially explain these discrepant results, such as different task contexts (e.g., probabilistic learning vs. perceptual discrimination), different approaches to the pupil analyses such as using simple contrasts or model-based regression, and different interpretations of what constitutes a "prediction error." Given these discrepancies, it is crucial to investigate the specific conditions under which pupil dilation reflects a prediction error.

Aims of the current study

The literature summarized above prompted us to investigate whether the pupil's response to decision outcomes during learning aligns with a prediction error signal defined within an information-theoretic framework. Whilst Zénon (2019) theoretically proposed a direct link between pupil dilation and information gain, this hypothesis has not been thoroughly tested in empirical studies. We sought to fill this gap in the literature and shed light on the relationship between information gain and uncertainty during learning as reflected in pupil dilation.

In the current study, we investigated whether the pupil's response to decision outcome (i.e., feedback) in the context of associative learning reflects a prediction error as defined by KL divergence, while also exploring the time course of this prediction error signal. To do so, we adapted a simple model of trial-by-trial learning of stimulus probabilities based on information theory from previous literature[29,72,73]. For completeness, Shannon surprise and entropy were also computed and related to the post-feedback pupil response. We analyzed two independent datasets featuring distinct associative learning paradigms, one characterized by increasing entropy and the other by decreasing entropy as the tasks progressed. By examining these different tasks, we aimed to identify commonalities in the results across varying contexts. Additionally, the contrasting directions of entropy in the two tasks enabled us to explore the hypothesis that the direction of the relationship between the post-feedback pupil response and information gain is influenced by the average uncertainty during learning.

In the first data set, participants were instructed to predict the upcoming orientation (left vs. right) of a visual target based on the probability of visual and auditory cues. In the second data set, participants were first exposed to letter-color pairs of stimuli in different frequency conditions during an odd-ball detection task. The letter-color pair contingencies were irrelevant to the odd-ball task performance. The participants subsequently completed a decision-making task in which they had to decide which letter was presented together most often with which color during the previous odd-ball detection task (match vs. no match). Pupil dilation was recorded during the decision-making tasks in both data sets and the post-feedback pupil response was the event of interest. We did not formally compare the results across the two data sets given substantial differences between these two task contexts. We expected the post-feedback pupil dilation to scale with KL divergence in both tasks in a relatively early time window, following the results of O'Reilly et. al. We explored whether later prediction error components in the post-feedback pupil dilation might reflect other information-theoretic variables, such as Shannon surprise or entropy.

To preview, the results show for the first time that whether the pupil dilates or constricts along with information gain seems to be context dependent, specifically related to increasing or decreasing average uncertainty (entropy). Our findings are overall in line with Zénon's hypothesis that pupil dilation reflects information-theoretic processing and furthermore suggest that these signatures in pupil dilation are complex and multifaceted. This study

provides empirical evidence that the pupil's response can shed light on model updating during learning, demonstrating the potential of this easily measured physiological indicator for exploring internal belief states.

2. Materials and methods

Data sets: decision-making tasks in associative learning paradigms

We analyzed two data sets related to associative learning in which pupil dilation was recorded during decision making; one of these data sets was publicly available[33] (see section 2.1), the other was collected for the study purposes. In the decision-making tasks of both data sets, participants had to make a two-alternative forced choice (2AFC) on each trial based on visual and/or auditory information. For the current analyses, the dependent variables of interest were participants' accuracy, reaction time (RT), and the feedback-locked pupil response. In the first data set (see section 2.1), participants learned probabilistic contingencies between stimuli during the 2AFC decision-making task itself. In the second data set (see section 2.2), there was an implicit learning phase prior to the 2AFC decision-making task during which the participants completed an odd-ball detection task that was irrelevant to the probabilistic contingencies between pairs of stimuli being presented. Pupil dilation was continuously measured in both data sets during each of the 2AFC decision-making tasks. The data sets consisted of independent samples of participants.

These data sets were chosen to analyze because we were able to quantify the post-feedback pupil dilation as the interaction between stimulus-pair frequency and accuracy as well as adapt an information-theoretic model of trial-by-trial learning in both task paradigms. The pupil is known to scale positively with errors as compared with correct responses[10,11,13,74–79]. In addition, studies have shown that pupil dilation positively scales with unexpected as compared with expected events[11,29,30,34–48]. Prediction errors can be operationalized as a function of task conditions related to stimulus expectations. For instance, a main effect of stimulus frequency reflects an unsigned prediction error, since the different frequency conditions correspond to different levels of expectancy (in the simplest form, a contrast of unexpected vs. expected). A main effect of accuracy (categorized post-hoc based on task performance) indicates an error signal about the binary outcome of a decision (correct vs. incorrect). However, an error signal is not the same as a prediction error signal, because an error alone does not necessarily convey information regarding expectation. In other words, errors on accuracy do not contain quantitative information regarding a difference between what was expected and what occurred. Expectations can modulate a main effect of accuracy in the post-feedback pupil response, indicating whether the outcome of the participant's accuracy (correct or incorrect) is better or worse than expected[10,11]. A signed prediction error in the context of associative learning would therefore be evidenced by an interaction between stimulus-pair frequency and accuracy.

In the current data sets, we were interested in comparing the condition when a participant makes an error, and the outcome was expected, with the condition in which they make an error and the outcome was unexpected (and similarly for when they are correct about their decision). If the signed prediction error signal reflects information gain, then we would expect a larger difference between error and correct trials for the condition with weaker expectations as compared with stronger expectations. In other words, more information is gained from a decision outcome in conditions with more uncertainty as compared with less uncertainty. Finally, we explored whether a definition of prediction error as an interaction between stimulus-pair frequency and accuracy will relate to information gain defined as KL divergence.

2.1 Data set #1: Cue-target 2AFC task

Independent analyses that focused on the relationship between the participants' pupil responses and Bayesian learning mechanisms have been previously published based on the same data set[33]. The data are publicly available and have been re-analyzed (including pre-

processing) in the current paper to answer a complementary but conceptually distinct research question compared with the Rutar et al. (2023) paper. The relevant methods that have been previously published are summarized here.

Participants and informed consent

From the thirty participants included in the published data set, six participants missed responses in at least one of the conditions required for the main two-way repeated-measures ANOVA and were therefore excluded from the statistical analysis. The final sample consisted of 24 participants aged 19-42 years (M = 23.3, SD = 4.9, 18 women). All participants gave written informed consent before participating and were compensated for participation.

Task and procedure

Participants performed a 2AFC decision-making task on the expected orientation direction (left vs. right) of the target stimulus (Gabor patches; spatial frequency = 0.033, opacity = 0.5, 400 x 400 pixels) while pupil dilation was recorded (Figure 1A). A chin rest was used to keep the distance (50 cm) to the computer screen constant (1920 x 1080 pixels, 120 Hz). This setup resulted in a Gabor patch with a visual angle of 12.1°. Participants were instructed to use the visual and auditory cues to predict the orientation of the upcoming target stimulus in each trial and to respond (by left or right button press corresponding to a left or right prediction, respectively) as soon as they knew which target orientation would appear. Participants were instructed to wait until cue offset to make their predictions. The orientation of the target was probabilistically determined by the combination of the preceding visual and auditory cues. The cue-target contingencies were not communicated in advance to the participants. The cue-target mappings were counterbalanced between participants in such a way that half of the participants saw the square followed by a right-oriented grating and a diamond followed by a left-oriented grating in 80% of the trials and for the other half of the participants, this mapping was reversed (i.e., square –> left and diamond –> right). In the remaining 20% of the trials, the participants received the reversed cue-target mapping with respect to their 80% mapping condition. On half of the trials, an auditory tone ("C" octave 4; 300 ms) was presented together with the onset of the visual cue. In the first half of the experiment (phase 1), this auditory tone was uninformative of the upcoming target orientation and could essentially be ignored. After phase 1 was completed, participants took a short break and were informed that the cue-target contingency rule would change; for the purposes of the current analysis, we only inspected phase 1 of the original experiment. Phase 1 consisted of 200 trials per participant. The order of the trials was randomly presented to each participant. The entire session took about 1.5 hours to complete (~45 minutes for phase 1). Stimuli were isoluminant and had a grey background.

Trial structure

Each trial of the cue-target task consisted of a fixation period (1.5 s), a cue period (1 s), a decision period during which the participant responded by pressing either the left or right button (RT), a delay period (3 s), and the target period (3 s) which served as feedback for the cue-target contingencies. A vertically oriented Gabor patch was presented on screen except for the target period. Participants were instructed to keep their gaze centered at the fixation cross in the middle of the screen. The fixation cross remained on screen except for the cue period, during which either a square or diamond appeared. During the cue period, a square or diamond would indicate the upcoming orientation direction of the target stimulus with a certain probability (20% vs. 80%). Note that in the cue-target task, the target period served as trial-by-trial feedback on the accuracy of the participants' cue-target predictions.

Data acquisition and preprocessing

Pupil dilation of the right eye was continuously recorded during phase 1 of the cue-target task using an SMI RED500 eye-tracker (SensoMotoric Instruments, Teltow/Berlin, Germany). The sampling rate was 500 Hz. Using custom Python code, the following steps were applied to the entire pupil dilation time series: i) linear interpolation around missing samples (0.15 s before and after each missing event), ii) linear interpolation around blinks or saccade events based on spikes in the temporal derivative (0.15 s before and after each nuisance event; note that blinks and saccades were considered as a single nuisance event), iii) band-pass filtering (third-order Butterworth, 0.01 to 6 Hz), iv) responses to nuisance events were removed using linear regression (nuisance responses were estimated using deconvolution [39] and v) the residuals of the nuisance regression were converted to percent signal change with respect to the temporal mean.

Trials that were faster or longer than 3 times the standard deviation of the *Z*-transformed RT distribution of each participant were excluded from the analysis, because there was no maximum response window (1.5% of the total number of trials were excluded)[80].

Differences with Rutar et al. (2023)
The main differences between the current work and Rutar et al. (2023) are the following. First, in the current analysis, we are only considering the first 200 trials (referred to as "phase 1") out of the 400 trials in total. After 200 trials (referred to as "phase 2"), the cue-target contingencies switched according to a specific rule. The change in rule-based contingencies prevented us from applying the ideal learner model to both phases of the task. Crucially, trials in phase 1 were independent from trials in phase 2 as they took place earlier in time, and therefore, discarding the second half of the experiment would not affect the associative learning processes taking place in the first half of the experiment. Second, we did not include six of the 30 participants that are in the publicly available data set provided by Rutar et al. (2023) due to missing cases in the repeated-measures ANOVAs in phase 1 of the experiment. Finally, Rutar et al. (2023) only tested for signed prediction errors within an early time window (see their Supplementary Materials) but did not investigate any later time windows as we do in the current experiment.

2.2 Data set #2: Letter-color 2AFC task

Participants and informed consent
The final sample consisted of 47 participants aged 17-45 years (M = 23.8, SD = 6.16, 34 women and 13 men). Fifty participants completed the experiment. Three participants had to be excluded due to technical error or human error on the part of the researcher. All participants gave written informed consent before participating and were compensated for participation.

Tasks and procedure
The experiment consisted of two separate tasks, the odd-ball detection and 2AFC decision tasks, which corresponded to a learning and decision-making phase, respectively (Figure 2A). Participants were not instructed about the decision-making task until the learning task was completed. A chin rest was used to keep the distance (58 cm) to the computer screen constant (1920 x 1080 pixels, 120 Hz). This setup resulted in visual stimuli spanning visual angles between 2.6° - 3.1°. We exposed participants to pairs of stimuli presented together in different frequency conditions during the learning phase of the experiment. We aimed for having low, medium, and high levels of stimulus-pair frequency to correspond to different amounts of exposure to associations between the individual letters and the colors. More exposure to consistent associations between letters and colors was expected to result in stronger expectations of the specific letter-color pairs as compared with less exposure. After this odd-ball task in which participants were exposed to the letter-color pairs in different frequency

conditions, participants were asked to make a 2AFC decision based on the presented stimuli pairs in the decision-making phase. Participants completed a questionnaire at the end of the computer tasks (data not reported here). Participants were given self-paced breaks between each task. The entire session took about 2.5 hours to complete.

Independent learning phase: odd-ball detection task

We hypothesized that statistical learning would take place within an odd-ball detection paradigm, during which participants had to monitor both the identity of a letter or number and the background color it was presented on (see Figure 2A). The stimuli used for the statistical-learning hypothesis were six letters ("A", "D", "I", "O", "R", "T", 100 pixels, Bookman Old Style font) and six shades of green as the background color in the shape of a square (120 x 120 pixels; see Figure 2A and Supplementary Figure 1 for hexadecimal codes). Shades of a single hue were chosen to help minimize verbal heuristic strategies for naming the six different colors. Three different frequency conditions were used (20%, 40%, and 80% of trials in which a specific letter was presented against a square background with a specific color), meaning that two letter-color pairs were in each frequency condition. For example, if the letter "A" was assigned to the 80% frequency condition, then "A" was shown with its associated shade of green as the background color in eight out of 10 trials in which "A" was presented. On the remaining two out of 10 trials in which an "A" was presented, the background shade of green was randomly drawn from all six shades of green. This sampling with replacement resulted in the letters being show together with their associated color in actual frequency conditions of 33%, 50% and 84%. Each letter was shown together with the other five unassociated colors on average in 13%, 10%, and 3% of trials (i.e., the noise around the letter-color pair signal) with respect to the 20%, 40% and 80% frequency conditions, respectively. The six letter-color pair combinations as well as their frequency condition were randomly assigned to each participant at the start of the experiment, meaning that participants received individual combinations of stimuli. For clarity, we will only refer to the actual frequency conditions in the letter-color 2AFC task (i.e., 33%, 50% and 84%) from here on instead of the intended frequency conditions (20%, 40% and 80%).

The odd-ball stimuli were numbers 1-9 (randomly drawn) and a non-green hue (one of four colors was chosen by random for each participant at the start of the experiment). Note that odd-balls could consist of i) a number with a non-green background, ii) a number with a green background, and iii) a letter with a non-green background. Participants completed a short round of 10 practice trials of the odd-ball task, which they could repeat until they understood the odd-ball rule. During the practice round, an equal number of odd-ball and regular stimuli were presented. Participants were instructed to indicate whether the stimulus on each trial was an odd-ball or not with either a left or a right button-press (button order was counterbalanced between participants). They could respond as soon as the stimulus appeared on screen. Participants were not informed about the frequency conditions of the letter-color pairs. The odd-ball task consisted of 660 trials in total (9% were odd-balls), during which participants had two self-paced breaks. The order of the trials was randomly presented to each participant. The odd-ball task took about 30 minutes to complete.

Oddball-task trial structure

Each trial of the odd-ball task consisted of a stimulus period (0.75 s), a response period (< 1.25 s), and an inter-trial interval that included feedback on accuracy (0.5 to 1 s jittered). A black fixation cross was presented in the center of the screen and changed to green, red, or blue in the inter-trial interval for correct, incorrect, or missed responses, respectively. Errors and missed trials were also accompanied by a short auditory tone (3rd octave "D" for 0.3 s). All stimuli were presented in the center of the screen against a grey background.

Letter-color visual decision task
After the odd-ball task, participants performed a 2AFC decision-making task on the occurrence of letter-color pairs that were shown during the preceding odd-ball task. Specifically, participants were instructed to indicate whether specific letter-color pairs occurred most often together in the preceding odd-ball task. They were instructed to guess if they were unsure. If a letter occurred most often together with a particular color in the preceding oddball task, then this was considered a "match." The 2AFC response options were "match" or "no match." For example, if the letter "A" was most often shown together with a specific shade of green as the background color during the preceding odd-ball task, then "A" would *match* this shade of green. Match and no-match conditions were presented in a 1:1 ratio of the number of trials. Participants could respond as soon as the color was presented on screen. Participants were instructed to indicate whether the letter-color pairs matched or not with either a left or a right button-press (button order was counterbalanced between participants). The letter-color visual decision task consisted of 250 trials in total, during which participants had three self-paced breaks. The order of the trials was randomly presented to each participant.

Each trial of the letter-color visual decision task consisted of a new trial cue period (0.2 s), a letter-stimulus period (0.75 s), a short delay period (0.1 s), a response period during which a colored square appeared (< 2.5 s), a longer delay period to give sufficient time for the pupil to return to baseline following a colored impulse (3.5 – 5.5 s, uniform distribution; see Supplementary Figure 1A,B) [81], and an inter-trial interval that included feedback on accuracy (3.5 – 5.5 s, uniform distribution). A black fixation cross was presented in the center of the screen during the longer delay period preceding feedback as well as in the inter-trial interval. Trial-by-trial feedback was presented to the participants by means of two auditory tones (0.3 s) for errors (3rd octave "D") and correct trials (4th octave "B"). We verified that the tone-locked pupil response was not differentially affected by the two feedback tones irrespective of the task context (see Supplementary Figure 1C,D). Participants were familiarized with the feedback tones before the decision task began. All stimuli were presented in the center of the screen against a grey background. The letter-color decision 2AFC task took about one hour to complete.

Data acquisition and preprocessing
Pupil dilation of the left or right eye was continuously recorded during the letter-color visual decision task using an EyeLink eye-tracker (SR Research, Ottawa, Ontario, Canada) with a sampling rate of 1000 Hz. The eye-tracker was calibrated once at the start of the decision task. A drift correction was performed after participants took breaks and could move their heads before continuing. Eye blinks and saccades were detected using the standard EyeLink software algorithms (default settings). Using custom Python code, the following steps were applied to the entire pupil dilation time series: i) linear interpolation around blinks (0.15 s before and after each blink), ii) band-pass filtering (third-order Butterworth, 0.01 to 6 Hz), iii) responses to blink and saccade events were estimated using deconvolution [39] and subsequently removed using linear regression, and v) the residuals of the nuisance regression were converted to percent signal change with respect to the temporal mean of the time series within each block of trials between the breaks.

Missed trials were excluded from all analyses (0.9%).

2.3 Quantification of the feedback-locked pupil response
For data from both datasets, each trial time course was first baseline corrected with the mean pupil size during the 0.5 s before the feedback event (target onset or auditory feedback for the cue-target or letter-color 2AFC tasks, respectively)[10,11,33]. The same time windows were used for both tasks with respect to the feedback event. Feedback-locked pupil responses were

defined by the mean pupil response within two time windows of interest: An early time window was defined to be 0.75 to 1.25 seconds after feedback onset to be centered around the peak of a transient event based on the canonical impulse response function of the pupil[10,11,71,82]. A late time window was defined to be 2.5 to 3 s after feedback onset and was determined by the shortest feedback interval within both decision tasks to make sure that the pupil response was uncontaminated by the subsequent trial: in the cue-target 2AFC task, a new trial started 3 s after the feedback onset.

2.4 Ideal learner models

Following previous literature[29,72,73], we adapted an ideal learner model to each of the two datasets separately. To model the learning of the task by participants, we assumed that they acted as ideal observers who learn the probability of seeing each of the stimulus pair types as the experiment progressed across trials. For each decision-making task, trial-by-trial quantifications of KL divergence was estimated using information theory to formally quantify information gain. In addition, surprise and entropy were computed and assumed to reflect the subjective probability and average uncertainty on each trial, respectively.

The ideal learner model for both data sets follows from the algorithms described by[72,73]. The ideal learner model represents a set of discrete events x, which can range from one to K. The events follow each other in a sequence with its length, j, equal to the number of trials in the task and denoted by $X^j = \{x^1, \dots, x^j\}$. We assume the participants learn the probabilities of X^j occurring, and following each trial, they update their estimate of the probability that each event type will occur based on the previously observed events. The distribution of probabilities can be parametrized by the vector $P(x) = [p_1, \dots, p_K]$, the elements of which sum to one and will be abbreviated by $P(x) = p$. Note that the probability of the kth event occurring, $P(x = k)$, is denoted by p_k.

Following previous literature[72,73], we assumed a Bayesian paradigm of giving the observer prior knowledge represented by a prior distribution of beliefs. This takes the form of a Dirichlet distribution which indicates the belief in all parameters prior to any observations of events. A Dirichlet prior over p is parameterized by the vector $\alpha = [\alpha_1, \dots, \alpha_K]$ and denoted by $P(p|\alpha) = Dir(p; \alpha_k)$. All elements of α are positive and the magnitude of each element corresponds to the relative expectation of each element. When the ideal learner has no previous expectations about which event will occur before the first trial of the task begins, all events are considered equally likely to occur and all the elements of α would be set to one (i.e., a uniform prior distribution). For example, if $\alpha = [100, 1, 1, 1]$, then the ideal learner model assumes that the event represented by the first element in α is much more likely to occur than any other event.

After an event is observed, the estimated probabilities p will be updated. The belief after j trials, X^j, is given by the posterior distribution

$$P(p \mid X^j, \alpha) = Dir(p^j; n_k^j + \alpha_k) = D^j, \tag{1}$$

in which n_k^j refers to the number of occurrences of event type k until trial j, and the parameter, α_k, determines prior expectations of event type k occurring. The posterior distribution given in equation (1) is updated after each new observation of an event and is abbreviated as D^j. For example, if $\alpha = [100, 1, 1, 1]$ and the participant observes an event represented by the second element in α on the first trial, then the posterior distribution becomes $[100, 2, 1, 1]$ for $j = 1$.

The probability of a certain event occurring at trial j can be computed directly from (1) as

$$p(x^j = k \mid X^{j-1}, \alpha) = \frac{(n_k^{j-1} + \alpha_k)}{(j - 1 + K)} = \tilde{p}_k^j \tag{2}$$

In words, equation (2) states that the probability that a certain event k will occur on trial j is denoted by $\tilde{p}_k^j$ (note the tilde indicates a prediction), which is equal to the total number of times event *k* occurred in previous trials $j-1$ plus the value of α_k, which is known to be fixed in the prior distribution, then divided by the total number of observations up to and including $j-1$ plus the total number of possible event types K.

Information-theoretic variables

Information theory[72,73,83–85] is used to estimate the KL divergence, surprise, and entropy at each trial. The subjective probability of each event occurring was quantified as surprise in terms of Shannon information, *I*

$$I(x^j = k) = -log_2\tilde{p}_k^j. \tag{3}$$

Note that $\tilde{p}_k^j$ is given by equation (2) and represents the prediction of the probability of event k occurring at trial j based on what has been observed up to and including trial $j-1$. The negative logarithm ensures that highly probable events are considered as less surprising.

The average uncertainty of the event at each trial was quantified as the entropy, H,

$$H(p^j) = -\sum_{k=1}^{k} p_k^j \, log_2 p_k^j. \tag{4}$$

Note that entropy was calculated including the observation of the event at trial j, denoted by p_k^j, not just up to trial $j-1$, because the participants received the information related to trial j at the time of feedback presentation.

Finally, the information gain at each trial was quantified as the KL divergence, D_{KL},

$$D_{KL}^j(p^j \mid p^{j-1}) = \sum_{k=1}^{k} p_k^j \, log_2 \frac{p_k^j}{p_k^{j-1}}, \tag{5}$$

and represents the difference between the prior and posterior distributions at trial j. Note that KL divergence is sometimes referred to as a 'Bayesian surprise'[86]. Here, when referring to 'surprise', we are always referring to the Shannon information as given by equation (3).

Ideal learner model assumptions for data set #1: cue-target 2AFC task

In the cue-target 2AFC task, an event was defined as one of the four possible cue-target pairs determined by the two cues (square and diamond) and two target orientations (left and right). Therefore, in the cue-target 2AFC task, $K = 4$, and x could take on values from 1 to 4. We assumed that the participants had no prior knowledge about the probabilities of events occurring at the start of the cue-target 2AFC task, and therefore, all events are assumed to be equally likely in the prior distribution. Since there are four possible events that can occur in the cue-target 2AFC task, then $P(x) = [p_1, p_2, p_3, p_4]$ and $\alpha = [1,1,1,1]$. Note that this resulted in a belief that all four cue-target pairs were likely to occur 25% of the time.

Ideal learner model assumptions for data set #2: letter-color 2AFC task

For the letter-color 2AFC task, during which pupil dilation was recorded, there were six letters and six shades of green used as stimuli. An event was defined as one of the 36 possible letter-color pairs; therefore, $K = 36$, and x could take on values from one to 36 and $P(x) = [p_1, \dots, p_{36}]$. We assumed that the participants started the letter-color 2AFC task with prior beliefs about the probabilities of the 36 letter-color pairs based on their observations during the preceding odd-ball task. The prior distribution for the letter-color 2AFC ideal learner model

was given by $\alpha = P_o(x) = [p_1, \dots, p_{36}]$, where P_o refers to the final probabilities of all 36 event types occurring at the end of the odd-ball task. Note that the probabilities of the odd-ball stimuli (i.e., an odd color or number) occurring were not estimated in the model. Unlike in the odd-ball task, the 36 letter-color pairs occurred at equal frequency in the letter-color 2AFC task; the occurrence of response options (match vs. no match) were also balanced. Learning about the letter-color probabilities could still occur however through the presentation of trial-by-trial feedback on accuracy.

2.5 Software and statistical analysis

All tasks were presented with PsychoPy (cue-target 2AFC task, version 1.81; letter-color 2AFC, version 1.82). Custom software in Python (version 3.6) was used for the preprocessing and data analysis unless otherwise stated. Note that the Rutar et al. (2023)[33] data set has been re-analyzed here and the above refers to the corresponding code for the current study. Python-specific packages and versions used are listed in the code repository.

Statistical inference on the evoked pupil responses was conducted with cluster-based permutation test from the MNE-Python package[87] unless otherwise stated. When the cluster-based permutation test could not be applied to a simple difference between two conditions, the evoked pupil responses were corrected with the false discovery rate instead. Repeated-measures ANOVAs and paired-samples t-tests were conducted using JASP version 0.16.3[88]. Wilcoxon signed-ranked t-tests were used when normality was deviated, and effect size is indicated by the matched rank biserial correlation (r_{rb}). For ANOVA results, Greenhouse-Geisser statistics are reported for violations of the assumption of sphericity and generalized eta squared (η^2_G) is reported as the effect size. In repeated-measures designs, η^2_G is useful as it can be more easily compared with between-subject designs[89,90].

Comparing the information-theoretic variables to the feedback-locked pupil response

We sought to examine the overall relationship each of the different information-theoretic variables has with the post-feedback pupil response. Furthermore, it is expected that these explanatory variables will be correlated with one another. For this reason, we did not adopt a multiple regression approach to test the relationship between the information-theoretic variables and pupil response in a single model[91]. For each decision task, we computed the trial-by-trial correlations between the feedback-locked pupil response at each time point in the time course (i.e., in the interval 0 to 3 s) and the KL divergence, surprise, and entropy in bits. This correlation analysis was done separately for each of the information-theoretic variables for all trials and then separately for error and correct trials. Correlation coefficients were normalized using the Fisher transform[92], and statistical significance of the coefficients was assessed using non-parametric cluster-based permutation tests.

All data and code used are publicly available (see link in data availability statement).

3. Results

The current study aimed to investigate whether the pupil's response to decision outcome in the context of associative learning reflects a prediction error as defined by KL divergence, while also exploring the time course of this prediction error signal. In two independent data sets, the frequency of pairs of stimuli were modulated in different conditions to induce a gradient of uncertain states (associative learning) upon which the participants had to make decisions. Within each data set, we inspected task performance, the evoked pupil time course, and the averaged pupil response in an early as compared with a late time window. A signed prediction error was defined as the interaction between stimulus-pair frequency and performance accuracy averaged across conditions. Finally, we tested the linear relationship

between the post-feedback pupil response and KL divergence in a trial-by-trial correlation analysis across the pupil time course.

3.1 Results from the cue-target 2AFC task (data set #1)

While pupil dilation was recorded, participants predicted the orientation (left or right) of the upcoming target stimulus (Gabor patch) based on the visual and/or auditory cues (Figure 1A). The target stimulus served as the feedback event of interest and was always presented after the participants made a prediction of the upcoming orientation of a Gabor patch by button press with the corresponding finger on the left or right hand. Dependent variables of interest were response accuracy, RT, and feedback-locked pupil response. Note that the results of the statistical tests are reported in the figure alongside the data illustrations (see Figure 1).

Behavioral performance

We evaluated the behavioral performance of the participants on the cue-target 2AFC task (data set #1). On average, participants were more accurate on the trials in the 80% condition as compared with the 20% condition tested with a Wilcoxon signed-ranked t-test (Figure 1B). Notably, the average accuracy within these two frequency conditions approximated the frequency itself but with substantial individual differences across the sample (see individual lines in Figure 1B). An interaction between frequency condition and accuracy was obtained in RT tested with a two-way repeated-measures ANOVA (Figure 1C): post-hoc t-tests showed that participants were faster to respond on correct trials as compared with incorrect trials in the 80% condition, while in contrast, participants were slower to respond on correct trials as compared with incorrect trials for the 20% condition. We note that it was impossible for participants to determine from the cues whether the trial was in the 80% or 20% condition at the time when responses were given. Average RTs did not differ between frequency or accuracy conditions overall as indicated by a lack of main effects in the two-way ANOVA (Figure 1C).

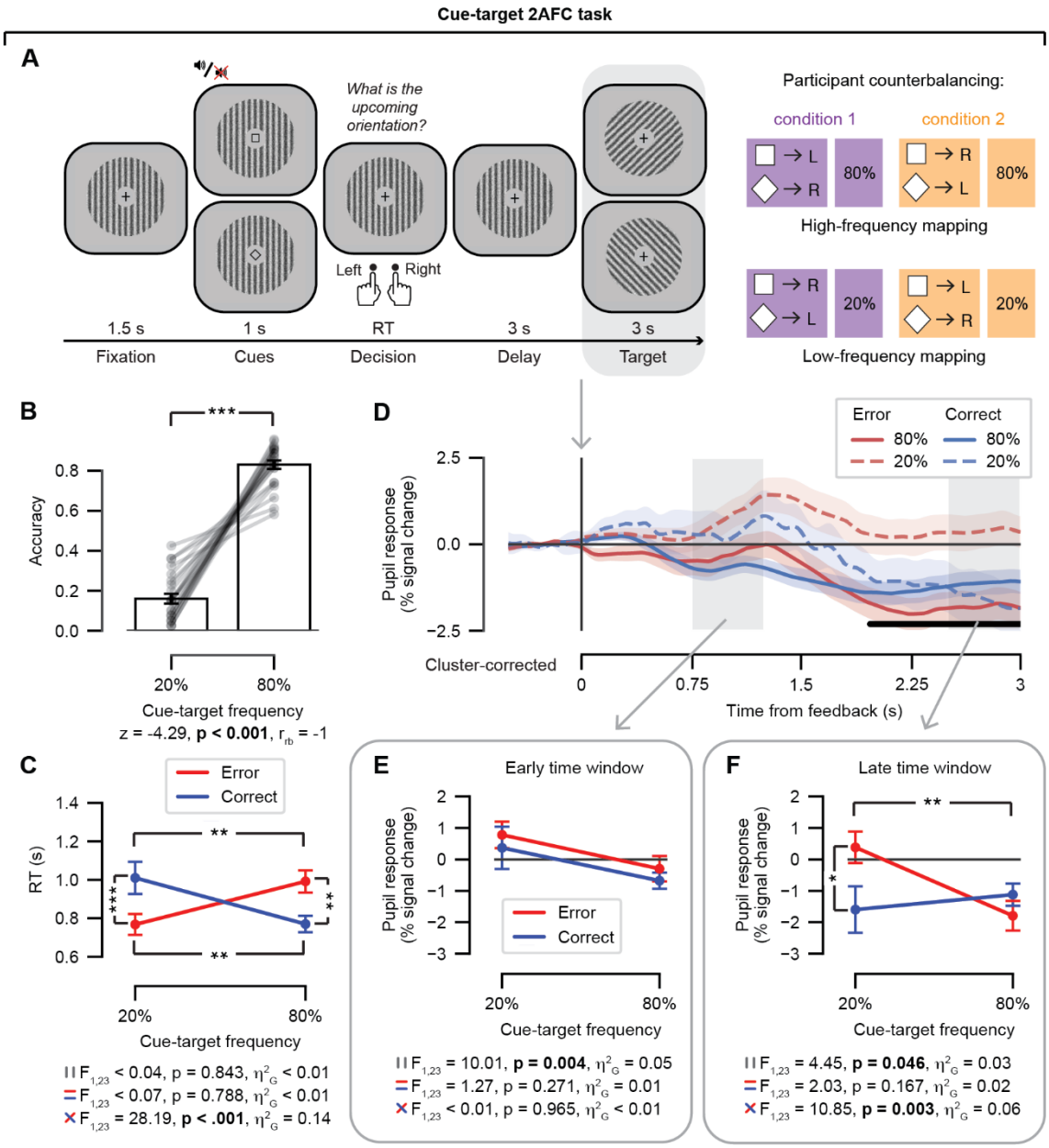


Figure 1. Data set #1: Cue-target 2AFC task and results. (A) Events during a single trial. While pupil dilation was recorded, participants predicted the orientation (left/right) of the upcoming target (Gabor patch) based on the visual and/or auditory cues (in the data analyzed here, only the visual cue had predictive validity). Predictions were given by a button press with the corresponding finger on the left or right hand. Two mapping conditions (condition 1 or condition 2) were counterbalanced across participants such that a participant in condition 1 was shown the square cue followed by a left-oriented target on 80% of the trials, while the square cue was followed by a right-oriented target on 20% of the trials. Gray box indicates the feedback event of interest. (B) Accuracy (fraction of correct responses) as a function of cue-target frequency. Data points are individual participants; stats, paired-samples Wilcoxon signed-rank t-test. (C) RT as a function of both cue-target frequency and accuracy (error/correct); stats, repeated-measures ANOVA. (D) Feedback-locked pupil response time course, plotted as a function of cue-target frequency and accuracy. Shading represents the standard error of the mean. Gray boxes, time windows of interest; early time window, [0.75, 1.25]; late time window, [2.5, 3.0]. The black horizontal bar indicates a significant interaction term (cluster-corrected, permutation test). (E) Early time window, average feedback-locked pupil response as a function of cue-target frequency and accuracy. stats, repeated-measures ANOVA. (F) As *E*, for the late time window. ANOVA results (multiple panels): top, main effect of frequency; middle, main effect of accuracy; bottom, frequency x accuracy interaction. Error bars, standard error of the mean. *$p<0.05$, **$p<0.01$, *** $p<0.001$.

Comparing the feedback-locked pupil response between time windows

We were interested in evaluating the pupil's response to the feedback event: in the cue-target 2AFC task, feedback occurred with the presentation of the target stimulus following the prediction given by the participant with a button press. The pupil response time course locked to the onset of the target stimulus is shown in Figure 1D (gray boxes indicated the early and late time windows of interest) with the four conditions of interested defined by the two-way interaction between stimulus-pair frequency and accuracy plotted separately for the 3 s interval of target presentation. A significant interaction between frequency and accuracy emerged later in the trial around 2 s and was sustained until the next trial occurred (3 s; the black bar in Figure 1D refers to significant time points based on the cluster-based permutation test). In Supplementary Figure 2, the feedback-locked pupil response time courses are plotted for the main effects of stimulus-pair frequency and accuracy.

We formally tested for a difference on the average feedback locked pupil response within the early as compared with late time windows in a three-way repeated-measures ANOVA with factors: time window (levels: early vs. late), frequency (levels: 20% vs. 80%) and accuracy (levels: error vs. correct). The results of the three-way repeated-measures ANOVA are presented in Table 1. Main effects of time window and frequency were obtained, in addition to a three-way interaction between the time window, frequency, and accuracy factors.

Table 1. Results of the three-way repeated-measures ANOVA on the feedback-locked pupil response in the cue-target 2AFC task (data set #1). The three-way repeated-measures ANOVA included factors: time window (levels: early vs. late), frequency (levels: 20% vs. 80%) and accuracy (levels: error vs. correct).

Factor	**F(1,23)**	**p**	**η^2_G**
Time window	7.52	**0.012**	0.05
Frequency	9.79	**0.005**	0.04
Accuracy	2.67	0.116	0.01
Time window x Frequency	0.25	0.621	< 0.01
Time window x Accuracy	0.26	0.614	< 0.01
Frequency x Accuracy	3.45	0.076	0.02
Time window x Frequency x Accuracy	24.97	**< 0.001**	0.02

To break down the three-way interaction, the two-way interactions between stimulus-pair frequency and accuracy were tested in independent repeated-measures ANOVAs for the early and late time window (see Figure 1E & 1F, respectively). As suggested from the time course analysis (see Figure 1D), frequency and accuracy did not interact for the average feedback-locked pupil response within the early time window (Figure 1E). In contrast, the interaction between frequency and accuracy was significant for the average feedback-locked pupil response within the late time window (Figure 1F): post-hoc t-tests showed that the error trials drove the two-way interaction between frequency and accuracy such that pupils dilated more for error trials as compared with correct trials only in the 20% frequency condition; pupils also dilated more during errors in the 20% frequency condition than errors in the 80% condition. Taken together, the data suggest that the post-feedback pupil response may reflect unsigned prediction errors in the early time window and signed prediction errors in the late time window in the cue-target 2AFC task.

For both the early and late time windows, a main effect of frequency was obtained in each two-way ANOVA, while no main effect of accuracy was evident (see Figures 1E and 1F). Post-hoc t-tests showed that pupils dilated on average more for the 20% frequency condition (M = 0.57%, SE = 0.45) as compared with the 80% frequency condition (M = -0.49%, SE = 0.31) for the early time window. The frequency effect was in the same direction in the late time window, with larger pupil dilation for the 20% frequency condition (M = -0.61%, SE = 0.51) as compared with the 80% frequency condition (M = -1.46%, SE = 0.36). Larger pupil dilation in

response to errors as compared with correct trials is a consistently reported effect in the literature[10,11,13,74–79]; therefore, it is worth noting here that accuracy and frequency were highly correlated in the cue-target 2AFC task (see Figure 1B) which could explain the lack of a main effect of accuracy obtained here. This is further illustrated by comparing the 'Error' time course in Supplementary Figure 2B with the '20%' time course in Supplementary Figure 2C (likewise, compare the 'Correct' time course in Supplementary Figure 2B with the '80%' time course in Supplementary Figure 2C). We note that within the early time window, the frequency effect in accuracy scaled with the frequency effect in the post-feedback pupil dilation across individual participants tested with a Spearman correction. Participants who showed a larger mean difference between the 80% as compared with the 20% frequency conditions in accuracy also showed smaller differences in pupil responses between frequency conditions (see Supplementary Figure 3). This monotonic relationship between the frequency effect in accuracy and post-feedback pupil response indicates that the improvement in accuracy (as measured behaviorally) across trials was also reflected in the change in pupil dilation.

3.2 Results from the letter-color 2AFC task (data set #2)

While pupil dilation was recorded, participants were administered a letter-color decision 2AFC: participants indicated whether each letter presented "matched" the subsequently presented colored square with a button press (Figure 2A; right-hand side). A match was correct when the letter and color had occurred most often together in the preceding odd-ball detection task. Dependent variables of interest were response accuracy, RT, and feedback-locked pupil response. Note that the results of the statistical tests are reported in the figure alongside the data illustrations (see Figure 2).

Behavioral performance during the odd-ball detection task

The odd-ball detection task served as an independent learning phase for the letter-color pairs. Participants performed the task as expected: accuracy was lower in identifying trials with an odd-ball present (M = 84.0%, SD = 9.9) as compared to regular trials (M = 99.8%, SD = 0.3; (t_{46} = 11.04, $p < 0.001$, d = 1.61). Likewise, RTs were slower for trials with an odd-ball present (M = 0.49 s, SD = 0.01) as compared to regular trials (M = 0.42, SD = 0.01; t_{46} = 16.53, $p <$ 0.001, d = 2.41).

Behavioral performance on the letter-color 2AFC task

We evaluated the behavioral performance of the participants on the letter-color 2AFC task (data set #2). During the letter-color 2AFC task, participants could accurately indicate whether a letter was presented most often with a given color in the preceding odd-ball detection task: response accuracy (around 80%) was higher than chance level (50%) in each of the three stimulus-pair frequency conditions on average (see bars in Figure 2B). This was also true for most participants (see individual lines in Figure 2B); however, neither accuracy (Figure 2B) nor RTs (Figure 2C) differed across the frequency conditions tested with repeated-measures ANOVAs (levels: 33%, 50%, and 84%). A main effect of accuracy as well as an interaction between frequency condition and accuracy was obtained in RT tested in a two-way repeated-measures ANOVA (Figure 2C). Post-hoc t-tests showed that participants were slower on error trials (M = 0.97 s, SE = 0.04) as compared with correct trials (M = 0.72 s, SE = 0.03). Breaking down the interaction in RT, post-hoc t-tests indicated that the accuracy difference between correct and error trials in the 50% frequency condition (M = 0.18 s, SE = 0.03) was significantly smaller as compared with that for the 33% frequency condition (M = 0.27 s, SE = 0.04; z = -2.15, p = 0.031, r_{rb} = 0.36) as well as the 84% frequency condition (M = 0.28 s, SE = 0.04; z = -2.55, p = 0.010, r_{rb} = -0.43), while the accuracy difference between the 33% and 84% frequency conditions did not differ on average (z = -0.78, p = 0.440, r_{rb} = -0.13).

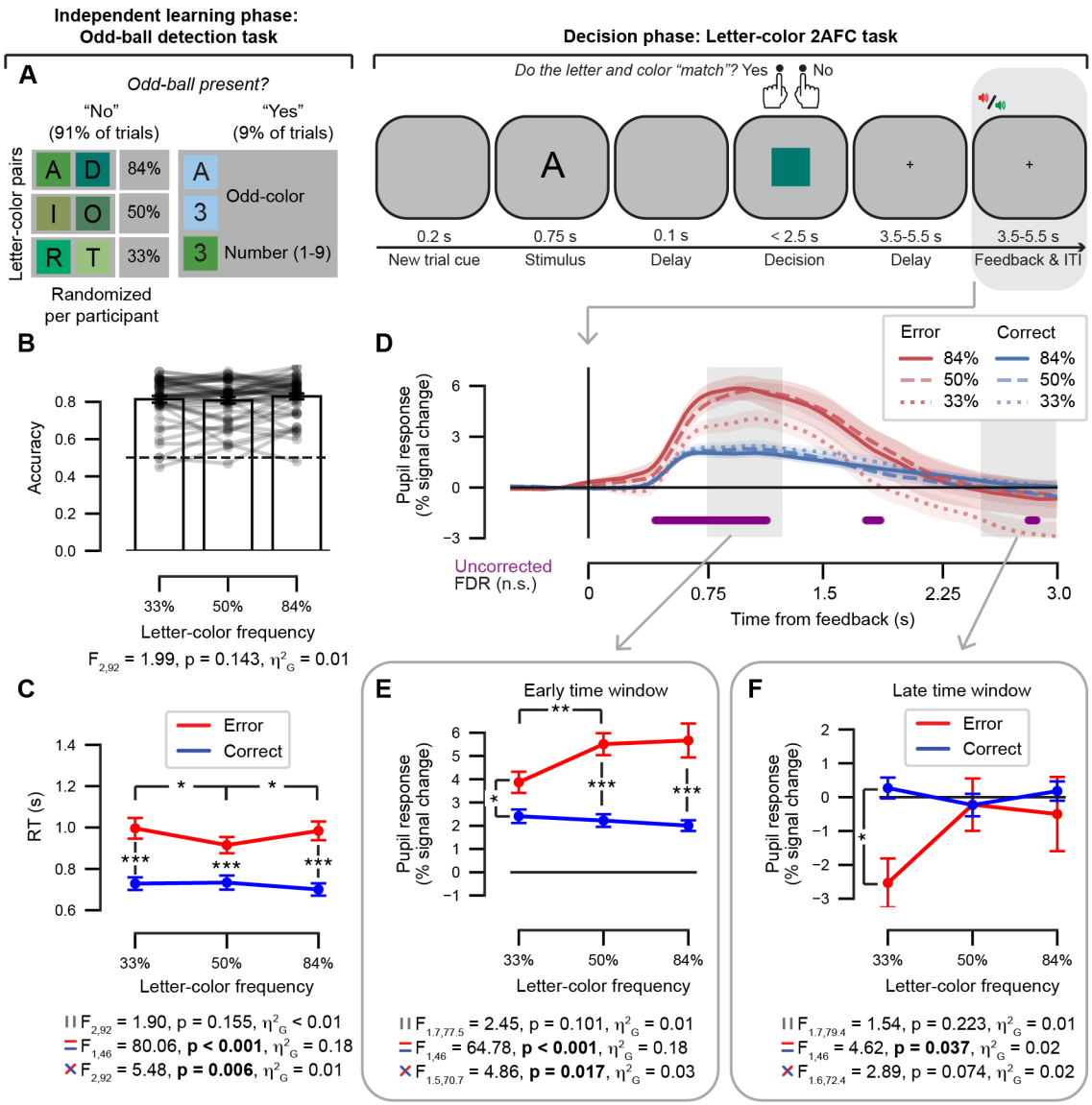


Figure 2. Data set #2: Letter-color 2AFC task and results. (A) *Left*, an independent learning phase was administered in the form of an odd-ball detection task during which six letters together with six shades of green colors as background (squares) were presented in three frequency conditions (33%, 50%, and 84%) on most trials (91%). Participants had to quickly respond to odd-ball targets (numbers and/or non-green color, 9% of trials). Letter-color mapping conditions were randomized per participant. *Right*, events during a single trial of the subsequent letter-color decision 2AFC task. While pupil dilation was recorded, participants indicated whether the letter "matched" the colored square with a button press. A match was correct when the letter and color had occurred most often together in the preceding odd-ball task. Gray box indicates the feedback event of interest. (B) Accuracy (fraction of correct responses) as a function of letter-color frequency. Dashed line represents chance level. Data points are individual participants; stats, paired-samples t-test. (C) RT as a function of both letter-color frequency and accuracy; stats, repeated-measures ANOVA. (D) Feedback-locked pupil response time course, plotted as a function of letter-color frequency and accuracy. Shading represents the standard error of the mean. Gray boxes, time windows of interest; early time window, [0.75, 1.25]; late time window, [2.5, 3.0]. The purple horizontal bar indicates a significant two-way interaction effect (uncorrected for multiple comparisons). No significant time points remained after correction using the false discovery rate (FDR). (E) Early time window, average feedback-locked pupil response as a function of letter-color frequency and accuracy. (F) As *E*, for the late time window. ANOVA results (multiple panels): top, main effect of frequency; middle, main effect of accuracy; bottom, frequency x accuracy interaction. Error bars, standard error of the mean. **p*<0.05, ***p*<0.01, *** *p*<0.001.

Comparing the feedback-locked pupil response between time windows

We were interested in evaluating the pupil's response to the feedback event: in the letter-color 2AFC task, explicit feedback was administered on each trial in the form of an auditory tone (error vs. correct) following the prediction given by the participant with a button press. The pupil response time course locked to the onset of the auditory feedback stimulus is shown in Figure 2D (gray boxes indicated the early and late time windows of interest) with the six conditions of interested defined by the two-way interaction between stimulus-pair frequency and accuracy plotted separately for a 3 s post-feedback interval. A two-way repeated-measures ANOVA was computed independently for each time point in the feedback-locked pupil time course in Figure 2D. Clusters indicating an interaction between frequency and accuracy emerged at three distinct time points, but none of these clusters survived corrections for multiple comparisons using the False Discovery Rate (see purple bar, Figure 2D). In Supplementary Figure 2, the feedback-locked pupil response time courses are plotted for the main effects of stimulus-pair frequency and accuracy. Using cluster-based permutation tests, we found that a robust main effect of accuracy spanned the early time window (Supplementary Figure 2E), while no main effect of frequency was obtained at any time points (Supplementary Figure 2F).

We formally tested for a difference on the average feedback locked pupil response averaged within the early as compared with late time windows in a three-way repeated-measures ANOVA with factors: time window (levels: early vs. late), frequency (levels: 33%, 50%, and 84%) and accuracy (levels: error vs. correct). The results of the three-way repeated-measures ANOVA are presented in Table 2. Main effects of time window and accuracy were obtained. The two-way interactions between time window and accuracy as well as frequency and accuracy were obtained. The three-way interaction between the time window, frequency and accuracy factors was not significant.

Table 2. Results of the three-way repeated-measures ANOVA on the feedback-locked pupil response in the letter-color 2AFC task (data set #2). The three-way repeated-measures ANOVA included factors: time window (levels: early vs. late), frequency (levels: 33%, 50%, and 84%) and accuracy (levels: error vs. correct). Greenhouse-Geisser statistics are reported when assumptions of sphericity were violated.

Factor	df	F	p	η^2_G
Time window	(1, 46)	130.16	**< 0.001**	0.23
Frequency	(1.639, 75.415)	2.11	0.137	0.01
Accuracy	(1, 46)	4.76	**0.034**	0.01
Time window x Frequency	(2, 92)	0.22	0.806	< 0.01
Time window x Accuracy	(1, 46)	61.11	**< 0.001**	0.06
Frequency x Accuracy	(1.480, 68.095)	3.91	**0.036**	0.02
Time window x Frequency x Accuracy	(2, 92)	1.05	0.354	< 0.01

We continued with a post-hoc exploration of the two-way interactions between stimulus-pair frequency (levels: 33%, 50%, 84%) and accuracy (levels: error vs. correct) by testing separate repeated-measures ANOVAs for the early and late time window given our a priori hypotheses about the nature of the three-way interaction. In the early time window, the two-way interaction between frequency and accuracy was obtained for the average feedback-locked pupil response (Figure 2E). To breakdown this two-way interaction between frequency and accuracy in the early time window (Figure 2E), we compared the difference between error as compared with correct trials across each pair of frequency conditions using t-tests: the accuracy difference in the 33% frequency condition (M = 1.46%, SE = 0.45) was significantly smaller as compared with that for the 50% frequency condition (M = 3.38%, SE = 0.43; z = -2.94, p = 0.003, r_{rb} = -0.49) as well as the 84% frequency condition (M = 3.66%, SE = 0.74; z = -2.32, p = 0.020, r_{rb} = -0.39), while the accuracy difference between the 50% and 84% frequency conditions did not differ on average (z = 0.65, p = 0.525, r_{rb} = 0.11). In the late time window, only a trend towards an interaction between frequency and accuracy was evident

(Figure 2F). In both the early and late time windows, the two-way ANOVAs showed that a main effect of accuracy was obtained for the average feedback-locked pupil responses while no main effect of frequency was obtained. Interestingly the direction of this main effect of accuracy differed per time window indicated by post-hoc t-tests (compare Figure 2F with Figure 2G): In the early time window, pupil dilation was larger on average for error trials (M = 5.01%, SE = 0.39) as compared with correct trials (M = 2.21%, SE = 0.23), while this effect reversed in direction during the late time window with larger pupil dilation for correct trials (M = 0.07%, SE = 0.26) as compared with error trials (M = -1.09%, SE = 0.55). Taken together, the data suggest that the post-feedback pupil response may reflect signed prediction errors, albeit more strongly within the early as compared with late time window; most striking is the fact that the direction of this interaction reversed across these time intervals.

In the letter-color 2AFC task, no scaling was evident between the frequency effect in accuracy and the frequency effect in the post-feedback pupil dilation across individual participants tested with a Spearman correlation (see Supplementary Figure 3). In contrast to the cue-target 2AFC task, no relationship between behaviorally accuracy measures and changes in pupil dilation was obtained in the letter-color 2AFC task.

3.3 Ideal learner model fits to feedback-locked pupil response

The ideal learner model used the stimulus information on each trial to estimate the KL divergence, surprise, and entropy during each of the decision-making tasks. We fit the feedback-locked pupil response to the resulting model parameters at each time point in the pupil time course to investigate which theoretic variable the post-feedback pupil signal may be reflecting (if any) and furthermore to see how these patterns developed over time (within the 0 to 3 s window following feedback). The relationship between the ideal learner models and the post-feedback pupil response was assessed by a correlation analysis (see Figure 3).

Correlation analysis for the cue-target 2AFC task

The information-theoretic variables are shown as function of task trial in Figure 3A and as a function of stimulus-pair frequency in Figure 3B. Multicollinearity was evidenced by the correlations between the model parameters (KL divergence vs. surprise: $r = 0.09$, $p < 0.001$; KL divergence vs. entropy: $r = 0.36$, $p < 0.001$; surprise vs. entropy: $r = 0.17$, $p < 0.001$). We independently evaluated the trial-by-trial correlations between the information-theoretic variables and the post-feedback pupil response during the cue-target 2AFC task to investigate the variance explained by each of the three model parameters. The time course of the resulting correlation coefficients showed a pattern in which the pupil scaled negatively with KL divergence almost immediately after feedback onset until about 1 s into the feedback interval extending into the early time window (Figure 3C, purple line). The post-feedback pupil response also scaled with surprise starting around 0.5 s with respect to feedback onset and lasted throughout the duration of the feedback interval, notably spanning both the early and late time windows (Figure 3C, teal line). No scaling between the post-feedback pupil response and entropy was obtained.

Note that this trial-by-trial analysis considers both error and correct trials simultaneously and is therefore not sensitive to the actual behavioral performance of each participant. To see whether the correlation between the information-theoretic variables and pupil response differed as a function of behavioral performance, we repeated the same correlation analysis of the model parameters to the feedback-locked pupil response separately for error and correct trials. Results showed that the post-feedback pupil response during correct trials scaled negatively with KL divergence from feedback onset until 1 s within the feedback interval (Figure 3D, blue line). The correlation between surprise and the feedback-locked pupil response for the error and correct trials diverged around 1.75 s with respect to feedback onset, and this difference persisted into the late time window (Figure 3E, black line). This suggests that Shannon surprise might be linked to the signed prediction errors (defined by the interaction between accuracy and stimulus-pair frequency) evident in the late time

window of the cue-target 2AFC task. No scaling between the post-feedback pupil response and entropy was obtained for either correct or error trials (Figure 3F).

Correlation analysis for the letter-color 2AFC task

We first confirmed that the odd-ball task yielded a gradient of probability distributions according to the task as designed: the 33%, 50%, and 84% stimulus-pair frequency conditions had mean final probabilities of 0.034 ($SD = 0.001$), 0.050 ($SD = 0.003$) and 0.112 ($SD = 0.004$), respectively. For each participant, the final probabilities of each letter-color pair at the end of the odd-ball task corresponded to the prior distribution entered into the ideal learner model for the letter-color 2AFC task.

For the letter-color 2AFC task, the three information-theoretic variables are shown as a function of task trial in Figure 3G and as a function of stimulus-pair frequency in Figure 3H. Multicollinearity was again evidenced by the correlations between the model parameters (KL divergence vs. surprise: $r = 0.33$, $p < 0.001$; KL divergence vs. entropy: $r = -0.51$, $p < 0.001$; surprise vs. entropy: $r = 0.05$, $p < 0.001$). The trial-by-trial correlation of the post-feedback pupil response with the model parameters was repeated for the letter-color 2AFC task. The time course of the resulting correlation coefficients showed a pattern in which the pupil scaled positively with KL divergence shortly (~0.5 s) after feedback onset until about 1.75 s into the feedback interval spanning across the early time window (Figure 3I, purple line). The post-feedback pupil response also scaled negatively with entropy across the same interval as for KL divergence (Figure 3I, yellow line). No scaling between the post-feedback pupil response and surprise was obtained.

We repeated the correlation analysis of the information-theoretic variables to the feedback-locked pupil response separately for error and correct trials. The post-feedback pupil response scaled positively with KL divergence across the early time window for both error and correct trials (Figure 3J, red and blue lines). While no scaling between the post-feedback pupil response with surprise was obtained for error or correct trials (Figure 3K), the relationship with KL divergence was mirrored in a negative scaling with entropy for correct trials only (Figure 3L).

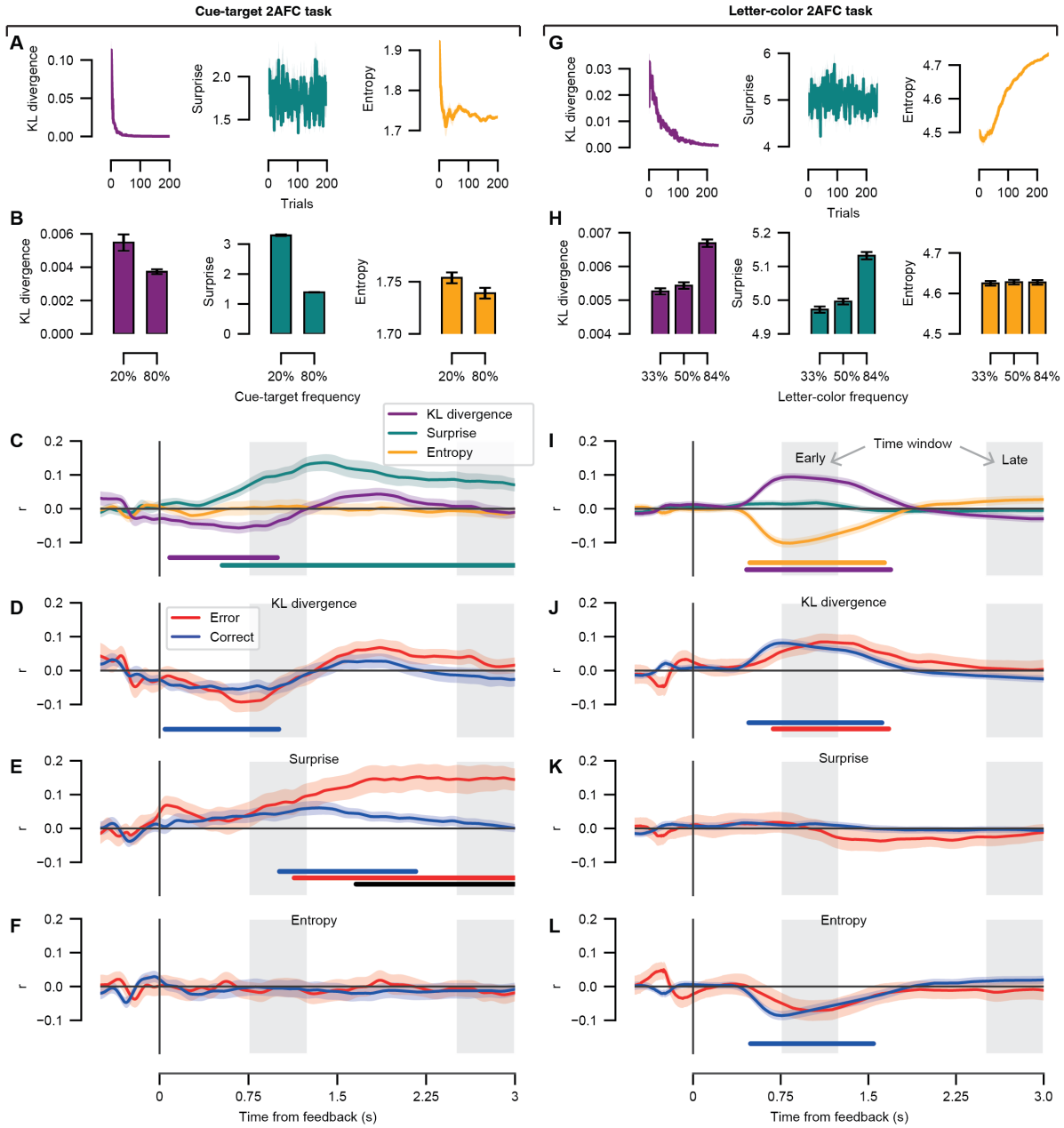


Figure 3. Correlations between the feedback-locked pupil response time course and the information-theoretic variables. Left column, results for the cue-target 2AFC task. Right column, results for the letter-color 2AFC task. (A) The KL divergence, surprise, and entropy parameters are shown as a function of task trial. Model parameter units are in bits. (B) The mean KL divergence, surprise, and entropy parameters are shown as a function of frequency condition. (C) Trial-by-trial correlations between the ideal learner model parameters (KL divergence, surprise, and entropy) at each time point in the feedback-locked pupil response. Shading represents the standard error of the mean. Gray boxes, time windows of interest; early time window, [0.75, 1.25]; late time window, [2.5, 3.0]. (D) Trial-by-trial correlations between the KL divergence parameter and the feedback-locked pupil response separately for the error and correct trials. (E) As *C*, for the surprise parameter. (F) As *D*, for the entropy parameter. (G-L) As *A-F* for the letter-color 2AFC task. All panels, the colored horizontal bars indicate time periods of significant correlation coefficients tested against zero for each model parameter or condition of interest (cluster-corrected, permutation test). The black horizontal bar indicates a difference between conditions (cluster-corrected, permutation test).

4. Discussion

In the current study, we investigated whether the pupil's response to decision outcome (i.e., feedback) in the context of associative learning reflects a prediction error as defined by KL divergence. We also explored how prediction error signals changed over time (3 s) with respect to the trial-wise feedback on decision outcome across two independent associative learning paradigms. Information-theoretic variables were derived from an ideal learner model and fit to the post-feedback pupil response at the trial-by-trial level. For completeness, we also computed Shannon surprise and entropy and examined their relationship with the post-feedback pupil response.

Results showed that the post-feedback pupil response correlated with KL divergence in the early time window across both tasks, while the direction of this scaling (qualitatively) differed per task. We speculate that this difference in the direction of scaling between information gain and the pupil response may depend on whether entropy was increasing or decreasing across trials. In further analyses, we noted that signed prediction error signals, which illustrate the relationship between frequency and accuracy, were evident in distinct time windows: during the later time window for the cue-target 2AFC task and in the early window for the letter-color 2AFC task. Shannon surprise appeared to be directly associated with the later component in the cue-target 2AFC task, while both KL divergence and entropy seemed to relate to the early interaction between frequency and accuracy in the letter-color 2AFC task. Our findings offer novel insights into the relationship between prediction error signals in post-feedback pupil responses and information processing by investigating how performance accuracy and information-theoretic variables reveal the underlying computational processes driving the interactions between stimulus-pair frequency and accuracy.

Information gain was reflected in the early post-feedback pupil response

The results supported our hypothesis that prediction error signals in the post-feedback pupil dilation reflected KL divergence, indicating that the pupil's response to decision outcome reflects the amount of information gained during associative learning. Across both task contexts, the post-feedback pupil response correlated with KL divergence within the early (0.75-1.25 s) but not late time window (2.5-3 s) (see Figures 3C and 3I). This early scaling between the pupil response following feedback onset and KL divergence suggest that the difference between the prior and posterior belief distributions are transiently reflected in pupil dilation shortly after observing the decision outcome about the stimulus pairs.

For the first time, we show that the *direction* of the relationship between pupil dilation and information gain (defined as KL divergence) may depend on whether the average uncertainty (i.e., entropy) was increasing or decreasing while learning was taking place. Specifically, in the cue-target 2AFC task, there was a negative effect of information gain on pupil dilation: the pupil response was smaller for larger values of KL divergence. The results from the cue-target 2AFC task are in-line with those reported by O'Reilly et. al. (2013) in which pupil dilation was also smaller for larger values of KL divergence centered around 1 s following target onset in a saccadic planning task. In both the cue-target 2AFC task as well as the saccadic planning task in O'Reilly et al., the moment when the observer is presented with the outcome of their decision, the average uncertainty is reduced. In contrast, during the letter-color 2AFC task, the pupil response was *larger* for larger values of KL divergence and information gain in fact seems to be driven by increased uncertainty. The entropy as a function of task trial differed between these task-contexts (compare Figure 3A right panel with Figure 3G right panel) which could explain the opposite direction of the relationship between pupil dilation and information gain: At the end of the odd-ball task, participants were exposed to the letter-color pairs in the high-frequency (84%) more often as compared with the lower-frequency conditions (33% and 50%). Therefore, stronger expectations about letter-color pairs for the 84% letter-color condition are represented by larger priors at the start of the letter-color 2AFC task. This increasing entropy in the letter-color 2AFC task can be attributed to the fact

that the letter-color pair conditions are balanced in terms of frequency of presentation while the prior distribution was not uniform. In other words, there is increasing average uncertainty driven by the stronger prior expectations in the subsequent fully balanced letter-color 2AFC task. To verify that the direction of entropy depended on the prior distribution chosen, we ran the ideal learning model on the letter-color 2AFC data using a uniform prior distribution (see Supplementary Figure 4).

Although we did not explicitly dissociate surprise and KL divergence in the cue-target 2AFC task design, as did O'Reilly et al., we found converging results related to the pupil's response to information gain (or "updating"). While O'Reilly et al. also found a positive relationship between pupil dilation and surprise, this surprise effect in their saccadic planning task emerged earlier than the scaling with KL divergence, unlike in the cue-target 2AFC task here. In the cue-target 2AFC task, we furthermore see a sustained surprise effect that spans most of the post-feedback interval, while the surprise effect in the previous saccadic planning task was transient. The discrepant results further illustrate the importance of task context for interpreting the relationship between pupil dilation and information processes. For instance, saccadic RTs were reported to scale with target expectancy; however, in the current study, we are investigating the post-feedback intervals that did not contain any motor responses from the participants. Furthermore, while we found early scaling between the post-feedback pupil response and KL divergence in both the cue-target and letter-color 2AFC tasks, the presence of a relationship between the post-feedback pupil response with surprise and entropy differed across tasks (compare Figures 3C with 3I).

O'Reilly et al. (2013) suggested that "pupil increases during learning are driven by uncertainty, or the influence of uncertainty on learning, rather than by learning or change per se." The results taken together across these two associative learning paradigms are in line with their suggestion. The results of the current study are also in line with the proposition of Zénon (2019) that "the ensemble of phenomena that trigger changes in pupil-linked arousal all depend on a basic underlying information theoretic process: the update of the brain internal models." Furthermore, from the contrast of the two associative learning tasks presented here, we can confirm that the pupil does not respond simply to surprise, because it does not always follow the frequency of occurrence of the stimulus-pairs, independently of the task. Instead, the pupil seems to respond to the amount of information provided by stimuli about the task variables. However, Zénon discusses the negative scaling of pupil dilation with KL divergence reported by O'Reilly et al. as contradicting their hypothesis that the pupil dilation will increase in proportion to how much novel sensory evidence is used to update current beliefs. Here we provide evidence that the direction of this relationship between pupil dilation and KL divergence needs more context and seems to relate to the direction of the entropy as learning progresses (i.e., either increasing or decreasing average uncertainty). One study testing children found that pupil dilation positively correlated with KL divergence, but not Shannon surprise, only when children were actively making predictions about water displacement, but not when they evaluated outcomes about water displacement without having to make predictions[68]. However, the authors did not report the direction of entropy across the trials in the water displacement task and it should be noted that the interval investigated corresponded to the prediction cue interval and not the outcome (i.e., feedback) interval.

Interestingly, only in the letter-color 2AFC task, results showed that the post-feedback pupil response negatively correlated with entropy in the early but not late time window, mirroring the positive scaling between the pupil response and KL divergence (see Figure 3I). While volatility and entropy describe different phenomena, they are interconnected through their relationship with uncertainty and predictability. High volatility typically corresponds to higher entropy, reflecting a system's complexity and unpredictability[93]. Previous work has shown that both tonic and phasic fluctuations in pupil dilation may track volatility in the environment, sometimes referred to as unexpected uncertainty[28,57–60,94]. Furthermore, noradrenaline plays a crucial role in how organisms track and respond to volatility in their environments, as it can signal when to update beliefs and expectations, enhancing the brain's

ability to adapt to fluctuations[67,95]. Only in the letter-color 2AFC task, in which the entropy increases across trials, do we also see that the post-feedback pupil response reflects the trial-by-trial estimates of entropy, but this is not the case for the cue-target 2AFC task. Using a probabilistic reversal learning task, Pajkossy et al. (2023)[60] reported that state entropy positively predicted post-feedback pupil size changes and interacted with the reversal probability of stimulus-reward contingencies across three different variations of the experiment. The scaling on entropy with the post-feedback pupil response reported by Pajkossy et al. occurred overall later in time with respect to feedback onset (ranging from feedback onset to 6 s depending on experimental variation) as compared with the letter-color 2AFC task. While more research is needed to understand these discrepant results, certainly differences between the reversal learning tasks in Pajkossy et al. and "simple" associative learning required in the letter-color 2AFC task could play a role, such as the presence or lack of changing stimulus probabilities and the inclusion of a point-based reward system for feedback.

The relationship between signed prediction error signals and information gain

As the term "ideal learner" suggests, the information-theoretic variables are computed based on the stimulus events shown to participants but are not sensitive to their actual behavioral performance. Of course, participants do not always act as ideal learners and make errors of observation, inference, and motor responses. Therefore, as a complementary analysis, we sought to examine the relationship between the performance accuracy and the information-theoretic variables to understand which computational processes may be underlying the stimulus-pair frequency and accuracy interactions reflected in the post-feedback pupil response.

In the cue-target 2AFC task, a prediction error should occur when the target orientation did not match the expected orientation based on the learned contingencies. A signed prediction error signal was obtained in the late time window, with the low-frequency (20%) error trials driving the interaction effect (see Figure 1F). Converging with this late signed prediction error signal, the correlations between surprise and the post-feedback pupil response differed for the error as compared with correct trials in the late time window (see Figure 3E). Specifically, the post-feedback pupil response during error trials showed larger correlation coefficients with surprise as compared with correct trials from about 1.75 to 3 s following the target onset. Thus, a signed prediction error signal defined by the interaction between frequency and accuracy in the late time window task seems to be driven by *surprise* and not KL divergence for the cue-target 2AFC. The direction of this effect could be interpreted in relation to sensory evidence. For instance, Colizoli et al. (2018) used a random dot discrimination paradigm with hard and easy levels of motion coherence that did not involve probabilistic learning. It was found that the errors during hard trials (corresponding to weaker sensory evidence) were driving the signed prediction error signal in a late time window (in that case, 3-6 s) following feedback onset similarly as the cue-target 2AFC task (see Figures 1D and 1F). The relationship between surprise and prediction error signal is partially in line with reward-linked feedback signals in van Slooten et al. (2018). Using a probabilistic value-based reinforcement learning task, van Slooten et al. reported that the early post-feedback pupil response (< ~2 s) was modulated by uncertainty about the value of options (with smaller differences between value options resulting in larger pupil dilation) but was not affected by violations of value beliefs (i.e., surprise). In contrast, the later post-feedback pupil response (around 2-3 s) positively reflected the degree to which outcomes violated current value beliefs. However, the direction of the late prediction error signal indicated that worse-than-expected outcomes were related to smaller pupil sizes, which seems to be at odds with other work showing that pupils generally dilate more when performance is worse than expected such as during errors[10,11,13,74–79]. The authors speculated that the late reward prediction error signal may reflect the firing pattern of phasic dopamine neurons, and other work supports the notion

of a significant component of dopamine signaling being reflected in pupil dilation[10,15,82]. A key difference between the current study and van Slooten et al. is the absence of reward-driven feedback during associative learning in the current study.

In the letter-color 2AFC task, a prediction error should occur when the participant expected that the letter-color pair did "match", but in fact they did not match, or vice versa. A signed prediction error signal was significant in the early time window (see Figure 2E). The direction of this interaction effect indicated that the pupil response difference between errors and correct trials increased as letter-color pair frequency increased (see Figure 2E). We note that the direction of the signed prediction error might seem counter intuitive as it relates to information gain, because stronger predictions (i.e., higher-frequency observations) often result in less information gain following outcome observation. As discussed above, in the letter-color 2AFC task, the amounts of surprise and KL divergence are highest for the high-frequency as compared with the lower-frequency conditions related to the increasing entropy across trials (see Figure 3H, left-hand and middle panels; see Supplementary Figure 4). Converging with this early signed prediction error signal, correlations between both KL divergence and entropy with the post-feedback pupil response were obtained in the early time window (see Figure 3I); however, no differences between correlations on error as compared with correct trials were obtained for either information-theoretic variable (see Figures 3J and 3L). Understanding the information processing in relation to performance accuracy may be crucial for disentangling the early signed prediction error signal. An alternative contributing factor that we did not explicitly test for is the participants' confidence about the stimulus-pair associations. Using an orientation discrimination task, de Gee et al. (2021)[11] reported that the early post-feedback pupil response was largest on error trials and smallest on correct trials when participants were most (subjectively) confident about their choices. Although we did not ask participants to report on their confidence about choices made, both RT and the strength of the stimulus-pair priors could be taken as a proxy for confidence[13,96]. In line with this, we did obtain interactions between stimulus-pair frequency and accuracy in both RT and the early post-feedback pupil response.

In sum, since the ideal learner model does not capture participant errors, we aimed to connect these two approaches of analyzing prediction errors by fitting the information-theoretic variables to the pupil response during error and correct trials independently. Signed prediction error signals defined by the interaction between frequency and accuracy were observed in the late time window for the cue-target 2AFC task and in the early window for the letter-color 2AFC task. Shannon surprise seemed to be directly related to the later component in the cue-target 2AFC task, while both KL divergence and entropy could relate to the early interaction between frequency and accuracy observed in the letter-color 2AFC task.

Limitations and future research

This study has some limitations. First, the two associative learning paradigms differed in many ways and were not directly comparable. For instance, the shape of the mean pupil response function differed across the two tasks in accordance with a visual or auditory feedback stimulus (compare Supplementary Figure 2A with Supplementary Figure 2D), and it is unclear whether these overall response differences contributed to any differences obtained between task conditions within each task. Future work should strive to disentangle how the specific aspects of the associative learning paradigms relate to prediction errors in pupil dilation by systematically manipulating design elements within each task. Task context clearly determines the relationship between the post-feedback pupil response and the information-theoretic variables, as it determines the uncertainty conditions surrounding during decision-making. To determine exactly how the different associative learning tasks relate to different temporal components of model updating is beyond the scope of the current study, but we speculate that hybrid predictive coding models may be able to account for fast (bottom-up) and slow (top-down) prediction errors reflected in pupil dilation[66]. Second, we did not design

the associative learning paradigms to orthogonalize the information-theoretic variables, such as was done in O'Reilly et al. Indeed, multicollinearity is evident between the information-theoretic variables in each of the two tasks; therefore, we did not attempt to quantify the unique portion of explained variance on the post-feedback pupil response for each variable of interest. Cleverer associative learning paradigms may be able to overcome this limitation. Finally, we are unable to attribute the relationship between computational variables and pupil dilation to specific neural mechanisms or neuromodulatory systems with the current study. Previous work has shown how neuromodulatory systems relate to learning and decision-making under uncertainty and the ability of the pupil to reflect these underlying computational processes[14,17,28,55,57–60,94,97]. Future research should aim at identifying the neural mechanisms involved in the processes underlying associative learning as reflected in pupil dilation across all phases of a decision process, ideally through computational theory.

Understanding prediction errors through pupil dilation within an information theory framework can illuminate predictive processing mechanisms in several significant ways. Pupil dilation acts as a physiological marker of cognitive and emotional responses, allowing researchers to quantitatively assess how discrepancies between expected and actual outcomes impact cognitive processing[25,60,69]. A framework for interpreting pupil dilation in terms of information theory may also enable exploration of how prediction errors function at various levels of a hierarchical model, helping researchers examine the interaction between high-level expectations and lower-level sensory inputs across different timescales, which informs the overall predictive model. Analyzing pupil responses in relation to prediction errors can reveal the extent of new information being processed and its influence on future predictions, thereby enhancing our understanding of learning and adaptation dynamics by elucidating feedback mechanisms in predictive processing and demonstrating how the brain adjusts its predictions based on new information. Given that pupil dilation is a peripheral marker of the brain's central arousal states, understanding its relationship with prediction errors can help disentangle the cognitive and affective components of predictive processing, providing a more comprehensive view of how the brain navigates uncertainty[95,98]. By integrating such insights, researchers can gain a deeper understanding of the mechanisms underlying predictive processing and how the brain continuously updates its internal models based on new experiences.

Conclusion

To conclude, the results provide evidence for Zénon's general assumption that pupil dilation can be described by an information-theoretic perspective. Clearly, task context plays a key role in the relationship between the information-theoretic variables and the post-feedback pupil response as may be expected. The temporal dynamics of these prediction errors signals should be carefully considered, as certain components tended to emerge around the peak of the canonical impulse response function and others may be sustained over time. These subtleties highlight the importance of adopting a model-based approach for characterizing the computational processes driving prediction errors as reflected in pupil dilation. Taken together, the post-feedback pupil response is a complex and multi-faceted signal that reflects different components of information processing during associative learning. The physiological response of the pupil provides a unique window into the brain's computations involved in model updating. More work is needed to link the information-theoretic variables reflected in the post-feedback pupil response with their underlying neuromodulatory mechanisms.

References

1. Aslin, R. N. Statistical learning: a powerful mechanism that operates by mere exposure. *WIREs Cogn. Sci.* **8**, e1373 (2017).
2. Turk-Browne, N. B., Jungé, J. A. & Scholl, B. J. The Automaticity of Visual Statistical Learning. *J. Exp. Psychol. Gen.* **134**, 552–564 (2005).

3. Siegelman, N., Bogaerts, L., Armstrong, B. C. & Frost, R. What exactly is learned in visual statistical learning? Insights from Bayesian modeling. *Cognition* **192**, 104002 (2019).
4. de Lange, F. P., Heilbron, M. & Kok, P. How Do Expectations Shape Perception? *Trends Cogn. Sci.* **22**, 764–779 (2018).
5. Gau, R. & Noppeney, U. How prior expectations shape multisensory perception. *NeuroImage* **124**, 876–886 (2016).
6. Sprevak, M. & Smith, R. An Introduction to Predictive Processing Models of Perception and Decision-Making. *Top. Cogn. Sci.* **n/a**, (2023).
7. Den Ouden, H. E., Kok, P. & De Lange, F. P. How prediction errors shape perception, attention, and motivation. *Front. Psychol.* **3**, 548 (2012).
8. Montague, P. R., Hyman, S. E. & Cohen, J. D. Computational roles for dopamine in behavioural control. *Nature* **431**, 760 (2004).
9. Schultz, W. Behavioral Theories and the Neurophysiology of Reward. *Annu. Rev. Psychol.* **57**, 87–115 (2005).
10. Colizoli, O., de Gee, J. W., Urai, A. E. & Donner, T. H. Task-evoked pupil responses reflect internal belief states. *Sci. Rep.* **8**, 13702 (2018).
11. de Gee, J. W., Correa, C. M. C., Weaver, M., Donner, T. H. & van Gaal, S. Pupil Dilation and the Slow Wave ERP Reflect Surprise about Choice Outcome Resulting from Intrinsic Variability in Decision Confidence. *Cereb. Cortex* **31**, 3565–3578 (2021).
12. Ratcliff, R., Smith, P. L., Brown, S. D. & McKoon, G. Diffusion Decision Model: Current Issues and History. *Trends Cogn. Sci.* **20**, 260–281 (2016).
13. Urai, A. E., Braun, A. & Donner, T. H. Pupil-linked arousal is driven by decision uncertainty and alters serial choice bias. *Nat. Commun.* **8**, 14637 (2017).
14. de Gee, J. W. *et al.* Dynamic modulation of decision biases by brainstem arousal systems. *eLife* **6**, (2017).
15. Lloyd, B., de Voogd, L. D., Mäki-Marttunen, V. & Nieuwenhuis, S. Pupil size reflects activation of subcortical ascending arousal system nuclei during rest. *eLife* **12**, e84822 (2023).
16. McGinley, M. J., David, S. V. & McCormick, D. A. Cortical Membrane Potential Signature of Optimal States for Sensory Signal Detection. *Neuron* **87**, 179–192 (2015).
17. Murphy, P. R., O'Connell, R. G., O'Sullivan, M., Robertson, I. H. & Balsters, J. H. Pupil diameter covaries with BOLD activity in human locus coeruleus. *Hum. Brain Mapp.* **35**, 4140–4154 (2014).
18. Reimer, J. *et al.* Pupil fluctuations track rapid changes in adrenergic and cholinergic activity in cortex. *Nat. Commun.* **7**, 13289 (2016).
19. Aston-Jones, G. & Cohen, J. D. An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. *Annu.Rev.Neurosci.* **28**, 403–450 (2005).
20. Glimcher, P. W. Understanding dopamine and reinforcement learning: The dopamine reward prediction error hypothesis. *Proc. Natl. Acad. Sci.* **108**, 15647 (2011).
21. Lak, A., Nomoto, K., Keramati, M., Sakagami, M. & Kepecs, A. Midbrain Dopamine Neurons Signal Belief in Choice Accuracy during a Perceptual Decision. *Curr. Biol.* **27**, 821–832 (2017).
22. Sara, S. J. The locus coeruleus and noradrenergic modulation of cognition. *Nat. Rev. Neurosci.* **10**, 211–223 (2009).
23. Schultz, W. Getting Formal with Dopamine and Reward. *Neuron* **36**, 241–263 (2002).
24. Yu, A. J. & Dayan, P. Uncertainty, Neuromodulation, and Attention. *Neuron* **46**, 681–692 (2005).
25. Browning, M., Behrens, T. E., Jocham, G., O'reilly, J. X. & Bishop, S. J. Anxious individuals have difficulty learning the causal statistics of aversive environments. *Nat. Neurosci.* **18**, 590–596 (2015).

26. Kayhan, E. *et al.* Young children integrate current observations, priors and agent information to predict others' actions. *PloS One* **14**, e0200976 (2019).
27. Koenig, S., Uengoer, M. & Lachnit, H. Pupil dilation indicates the coding of past prediction errors: Evidence for attentional learning theory. *Psychophysiology* **55**, e13020 (2018).
28. Nassar, M. R. *et al.* Rational regulation of learning dynamics by pupil-linked arousal systems. *Nat. Neurosci.* **15**, 1040 (2012).
29. O'Reilly, J. X. *et al.* Dissociable effects of surprise and model update in parietal and anterior cingulate cortex. *Proc. Natl. Acad. Sci.* **110**, E3660–E3669 (2013).
30. Preuschoff, K., 't Hart, B. & Einhauser, W. Pupil Dilation Signals Surprise: Evidence for Noradrenaline's Role in Decision Making. *Front. Neurosci.* **5**, 115 (2011).
31. Satterthwaite, T. D. *et al.* Dissociable but inter-related systems of cognitive control and reward during decision making: Evidence from pupillometry and event-related fMRI. *NeuroImage* **37**, 1017–1031 (2007).
32. Van Slooten, J. C., Jahfari, S., Knapen, T. & Theeuwes, J. How pupil responses track value-based decision-making during and after reinforcement learning. *PLOS Comput. Biol.* **14**, e1006632 (2018).
33. Rutar, D. *et al.* Differentiating between Bayesian parameter learning and structure learning based on behavioural and pupil measures. *PLOS ONE* **18**, e0270619 (2023).
34. Alamia, A., VanRullen, R., Pasqualotto, E., Mouraux, A. & Zenon, A. Pupil-Linked Arousal Responds to Unconscious Surprisal. *J. Neurosci.* **39**, 5369 (2019).
35. Bianco, R., Ptasczynski, L. E. & Omigie, D. Pupil responses to pitch deviants reflect predictability of melodic sequences. *Brain Cogn.* **138**, 103621 (2020).
36. Friedman, D., Hakerem, G., Sutton, S. & Fleiss, J. L. Effect of stimulus uncertainty on the pupillary dilation response and the vertex evoked potential. *Electroencephalogr. Clin. Neurophysiol.* **34**, 475–484 (1973).
37. Kamp, S.-M. & Donchin, E. ERP and pupil responses to deviance in an oddball paradigm. *Psychophysiology* **52**, 460–471 (2015).
38. Kloosterman, N. A. *et al.* Pupil size tracks perceptual content and surprise. *Eur. J. Neurosci.* **41**, 1068–1078 (2015).
39. Knapen, T. *et al.* Cognitive and Ocular Factors Jointly Determine Pupil Responses under Equiluminance. *PLoS ONE* **11**, e0155574 (2016).
40. Kuchinke, L., Võ, M. L.-H., Hofmann, M. & Jacobs, A. M. Pupillary responses during lexical decisions vary with word frequency but not emotional valence. *Int. J. Psychophysiol.* **65**, 132–140 (2007).
41. Lavin, C., San Martín, R. & Rosales Jubal, E. Pupil dilation signals uncertainty and surprise in a learning gambling task. *Front. Behav. Neurosci.* **7**, (2014).
42. Liao, H.-I., Yoneya, M., Kidani, S., Kashino, M. & Furukawa, S. Human pupillary dilation response to deviant auditory stimuli: Effects of stimulus properties and voluntary attention. *Front. Neurosci.* **10**, 43 (2016).
43. Qiyuan, J., Richer, F., Wagoner, B. L. & Beatty, J. The Pupil and Stimulus Probability. *Psychophysiology* **22**, 530–534 (1985).
44. Raisig, S., Welke, T., Hagendorf, H. & van der Meer, E. I spy with my little eye: Detection of temporal violations in event sequences and the pupillary response. *Int. J. Psychophysiol.* **76**, 1–8 (2010).
45. Silvestrin, F., Penny, W. D. & FitzGerald, T. H. B. Pupil dilation indexes automatic and dynamic inference about the precision of stimulus distributions. *J. Math. Psychol.* **101**, 102503 (2021).
46. Wetzel, N., Buttelmann, D., Schieler, A. & Widmann, A. Infant and adult pupil dilation in response to unexpected sounds. *Dev. Psychobiol.* **58**, 382–392 (2016).
47. Zhao, S. *et al.* Pupil-linked phasic arousal evoked by violation but not emergence of regularity within rapid sound sequences. *Nat. Commun.* **10**, 4030 (2019).

48. Ghilardi, T., Poli, F., Meyer, M., Colizoli, O. & Hunnius, S. Early roots of information-seeking: Infants predict and generalize the value of information. (2024) doi:10.7554/elife.92388.1.
49. de Berker, A. O. *et al.* Computations of uncertainty mediate acute stress responses in humans. *Nat. Commun.* **7**, 10996 (2016).
50. Findling, C., Skvortsova, V., Dromnelle, R., Palminteri, S. & Wyart, V. Computational noise in reward-guided learning drives behavioral variability in volatile environments. *Nat. Neurosci.* **22**, 2066–2077 (2019).
51. Geng, J., Blumenfeld, Z., Tyson, T. & Minzenberg, M. Pupil diameter reflects uncertainty in attentional selection during visual search. *Front. Hum. Neurosci.* **9**, (2015).
52. Fan, H. *et al.* Pupil Size Encodes Uncertainty during Exploration. *J. Cogn. Neurosci.* **35**, 1508–1520 (2023).
53. Krishnamurthy, K., Nassar, M. R., Sarode, S. & Gold, J. I. Arousal-related adjustments of perceptual biases optimize perception in dynamic environments. *Nat. Hum. Behav.* **1**, 0107 (2017).
54. Lempert, K. M., Chen, Y. L. & Fleming, S. M. Relating Pupil Dilation and Metacognitive Confidence during Auditory Decision-Making. *PLOS ONE* **10**, e0126588 (2015).
55. Murphy, P. R., Vandekerckhove, J. & Nieuwenhuis, S. Pupil-Linked Arousal Determines Variability in Perceptual Decision Making. *PLOS Comput. Biol.* **10**, e1003854 (2014).
56. Richer, F. & Beatty, J. Contrasting Effects of Response Uncertainty on the Task-Evoked Pupillary Response and Reaction Time. *Psychophysiology* **24**, 258–262 (1987).
57. Vincent, P., Parr, T., Benrimoh, D. & Friston, K. J. With an eye on uncertainty: Modelling pupillary responses to environmental volatility. *PLOS Comput. Biol.* **15**, e1007126 (2019).
58. Filipowicz, A. L., Glaze, C. M., Kable, J. W. & Gold, J. I. Pupil diameter encodes the idiosyncratic, cognitive complexity of belief updating. *eLife* **9**, e57872 (2020).
59. Harris, D. J. *et al.* Task-evoked pupillary responses track precision-weighted prediction errors and learning rate during interceptive visuomotor actions. *Sci. Rep.* **12**, 22098 (2022).
60. Pajkossy, P., Gesztesi, G. & Racsmány, M. How uncertain are you? Disentangling expected and unexpected uncertainty in pupil-linked brain arousal during reversal learning. *Cogn. Affect. Behav. Neurosci.* **23**, 578–599 (2023).
61. Zénon, A. Eye pupil signals information gain. *Proc. R. Soc. B Biol. Sci.* **286**, 20191593 (2019).
62. Modirshanechi, A., Becker, S., Brea, J. & Gerstner, W. Surprise and novelty in the brain. *Curr. Opin. Neurobiol.* **82**, 102758 (2023).
63. Friston, K. The free-energy principle: a unified brain theory? *Nat. Rev. Neurosci.* **11**, 127–138 (2010).
64. Spratling, M. W. A review of predictive coding algorithms. *Perspect. Hum. Probabilistic Inferences Bayesian Brain* **112**, 92–97 (2017).
65. Nienborg, H. & Roelfsema, P. R. Belief states as a framework to explain extra-retinal influences in visual cortex. *Large-Scale Rec. Technol. 32* **32**, 45–52 (2015).
66. Tscshantz, A., Millidge, B., Seth, A. K. & Buckley, C. L. Hybrid predictive coding: Inferring, fast and slow. *PLOS Comput. Biol.* **19**, e1011280 (2023).
67. Sales, A. C., Friston, K. J., Jones, M. W., Pickering, A. E. & Moran, R. J. Locus Coeruleus tracking of prediction errors optimises cognitive flexibility: An Active Inference model. *PLOS Comput. Biol.* **15**, e1006267 (2019).
68. Colantonio, J., Bascandziev, I., Theobald, M., Brod, G. & Bonawitz, E. Seeing the Error in My “Bayes”: A Quantified Degree of Belief Change Correlates with Children’s Pupillary Surprise Responses Following Explicit Predictions. *Entropy* **25**, (2023).

69. Stemerding, L. E., van Ast, V. A., Gerlicher, A. M. V. & Kindt, M. Pupil dilation and skin conductance as measures of prediction error in aversive learning. *Behav. Res. Ther.* **157**, 104164 (2022).
70. Burlingham, C. S., Mirbagheri, S. & Heeger, D. J. A unified model of the task-evoked pupil response. *Sci. Adv.* **8**, eabi9979 (2022).
71. Hoeks, B. & Levelt, W. J. M. Pupillary dilation as a measure of attention: a quantitative system analysis. *Behav. Res. Methods Instrum. Comput.* **25**, 16–26 (1993).
72. Mars, R. B. *et al.* Trial-by-Trial Fluctuations in the Event-Related Electroencephalogram Reflect Dynamic Changes in the Degree of Surprise. *J. Neurosci.* **28**, 12539 (2008).
73. Poli, F., Serino, G., Mars, R. B. & Hunnius, S. Infants tailor their attention to maximize learning. *Sci. Adv.* **6**, eabb5053 (2020).
74. Braem, S., Coenen, E., Bombeke, K., Van Bochove, M. E. & Notebaert, W. Open your eyes for prediction errors. *Cogn. Affect. Behav. Neurosci.* **15**, 374–380 (2015).
75. Critchley, H. D., Tang, J., Glaser, D., Butterworth, B. & Dolan, R. J. Anterior cingulate activity during error and autonomic response. *NeuroImage* **27**, 885–895 (2005).
76. Maier, M. E., Ernst, B. & Steinhauser, M. Error-related pupil dilation is sensitive to the evaluation of different error types. *Biol. Psychol.* **141**, 25–34 (2019).
77. Murphy, P. R., van Moort, M. L. & Nieuwenhuis, S. The Pupillary Orienting Response Predicts Adaptive Behavioral Adjustment after Errors. *PLOS ONE* **11**, e0151763 (2016).
78. Rondeel, E., Van Steenbergen, H., Holland, R. & van Knippenberg, A. A closer look at cognitive control: differences in resource allocation during updating, inhibition and switching as revealed by pupillometry. *Front. Hum. Neurosci.* **9**, (2015).
79. Wessel, J. R., Danielmeier, C. & Ullsperger, M. Error Awareness Revisited: Accumulation of Multimodal Evidence from Central and Autonomic Nervous Systems. *J. Cogn. Neurosci.* **23**, 3021–3036 (2011).
80. Berger, A. & Kiefer, M. Comparison of Different Response Time Outlier Exclusion Methods: A Simulation Study. *Front. Psychol.* **12**, (2021).
81. Mathot, S. Pupillometry: Psychology, Physiology, and Function. **1**, 16 (2018).
82. de Gee, J. W., Knapen, T. & Donner, T. H. Decision-related pupil dilation reflects upcoming choice and individual bias. *Proc. Natl. Acad. Sci. U. S. A.* **111**, E618-25 (2014).
83. Attneave, F. *Applications of Information Theory to Psychology: A Summary of Basic Concepts, Methods, and Results.* vii, 120 (Henry Holt, Oxford, England, 1959).
84. C. E. Shannon. A mathematical theory of communication. *Bell Syst. Tech. J.* **27**, 379–423 (1948).
85. Strange, B. A., Duggins, A., Penny, W., Dolan, R. J. & Friston, K. J. Information theory, novelty and hippocampal responses: unpredicted or unpredictable? *Neural Netw.* **18**, 225–230 (2005).
86. Itti, L. & Baldi, P. Bayesian surprise attracts human attention. *Vis. Atten. Psychophys. Electrophysiol. Neuroimaging* **49**, 1295–1306 (2009).
87. Gramfort, A. *et al.* MEG and EEG data analysis with MNE-Python. *Front. Neurosci.* 267 (2013).
88. JASP Team. JASP. (2022).
89. Lakens, D. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. *Front. Psychol.* **4**, 863 (2013).
90. Olejnik, S. & Algina, J. Generalized Eta and Omega Squared Statistics: Measures of Effect Size for Some Common Research Designs. *Psychol. Methods* **8**, 434–447 (2003).
91. Morrissey, M. B. & Ruxton, G. D. Multiple Regression Is Not Multiple Regressions: The Meaning of Multiple Regression and the Non-Problem of Collinearity. *Philos. Theory Pract. Biol.* **10**, (2018).
92. Fisher, R. A. Frequency Distribution of the Values of the Correlation Coefficient in Samples from an Indefinitely Large Population. *Biometrika* **10**, 507–521 (1915).

93. Sheraz, M., Dedu, S. & Preda, V. Entropy Measures for Assessing Volatile Markets. *2nd Int. Conf. Econ. Sci. Res. - Theor. Empir. Pract. Approaches ESPERA 2014 13-14 Novemb. 2014 Buchar. Rom.* **22**, 655–662 (2015).
94. Murphy, P. R., Wilming, N., Hernandez-Bocanegra, D. C., Prat-Ortega, G. & Donner, T. H. Adaptive circuit dynamics across human cortex during evidence accumulation in changing environments. *Nat. Neurosci.* (2021) doi:10.1038/s41593-021-00839-z.
95. Nassar, M. R. Toward a computational role for locus coeruleus/norepinephrine arousal systems. *Curr. Opin. Behav. Sci.* **59**, 101407 (2024).
96. Sanders, J. I., Hangya, B. & Kepecs, A. Signatures of a Statistical Computation in the Human Sense of Confidence. *Neuron* **90**, 499–506 (2016).
97. Meyniel, F. Brain dynamics for confidence-weighted learning. *PLOS Comput. Biol.* **16**, e1007935 (2020).
98. Pulcu, E. & Browning, M. The Misestimation of Uncertainty in Affective Disorders. *Trends Cogn. Sci.* **23**, 865–875 (2019).
99. Barbur, J. L., Harlow, A. J. & Sahraie, A. Pupillary responses to stimulus structure, colour and movement. *Ophthalmic Physiol. Opt.* **12**, 137–141 (1992).

Supplementary Materials: Pupil dilation offers a time-window on prediction error

Authors: Olympia Colizoli[1]*, Tessa van Leeuwen[1,2], Danaja Rutar[1,2], Harold Bekkering[1]

Affiliations:
[1] Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, The Netherlands
[2] Department of Communication and Cognition, Tilburg University, Tilburg, The Netherlands
[3] Department of Psychology, Sigmund Freud University, Ljubljana, Slovenia

Corresponding Author:
*Olympia Colizoli
Radboud University Nijmegen
Donders Centre for Cognition
Thomas van Aquinostraat 4
6525 GD Nijmegen

Materials & Correspondence: olympia.colizoli@donders.ru.nl (O.C.)

Supplementary section 1. Control tasks for data set #2

Different colors and tones could influence the pupil response due to inherent properties of the stimuli, and thereby confound true feedback-related signals. Therefore, complementary to the main analysis, we administered two control tasks in one independent sample of participants to directly assess whether confounding effects on the pupil's response to the colors and tones presented in the letter-color 2AFC task should be expected (see section 2.2).

Participants and informed consent

An independent sample of fifteen participants (*M* = 24.27 y, *SD* = 5.79, gender data was not acquired) completed both control tasks in a single session that lasted 30 minutes in total. All stimuli and equipment were identical to the setup reported for the letter-color 2AFC task (data set #2, see section 2.2). All participants gave written informed consent before participating and were financially compensated for participation with the standard amount.

Supplementary section 1.1 The pupil's response to colored squares returns to baseline before feedback onset

Procedure and Methods

Color was a stimulus dimension in the letter-color 2AFC task, and colors are known to result in impulse responses driven by the light reflex of the pupil[99]. This light reflex is uninteresting in the context of the current experiment and furthermore could potentially affect the feedback-related pupil response of interest differently for different colors. Participants were presented with the six colored squares used in the letter-color 2AFC tasks and instructed to react with a button press as soon as the color appeared. In the main letter-color 2AFC task, the colored square was replaced with a fixation cross upon the participant's response, thus stimulus duration varied across trials and participants. To account for part of this variation in stimulus duration, each of the 15 control participants was yoked to one participant in the letter-color 2AFC task. In the control task for colors, the stimulus duration of the colored squares was equal to the mean reaction time (across all trials) of the yoked counterpart of each participant. On each trial, a fixation cross was presented for 0.5 s, before the colored square was presented for a variable duration (see above), followed by the fixation cross again for the inter-trial interval (3.5 – 5.5 s; uniform distribution). Each color was presented 20 times for a total of 120 trials while pupil dilation was simultaneously recorded. Behavioral data on the control task were not analyzed.

Results

From the results, on average, the colors resulted in a constriction of the pupil dilation (Supplementary Figure 1A). Crucially, when inspecting each of the six colors individually, all responses returned to baseline well before the delay period (3.5 – 5.5 s; uniform distribution) was terminated and the auditory feedback was presented (Supplementary Figure 1B).

In sum, the control task for colors showed that the pupil's impulse response to the six different colors used in the letter-color 2AFC task would not have affected the upcoming feedback stimulus.

Supplementary section 1.2 Mean pupil response is similar for the two 'feedback' tones

Procedure and Methods

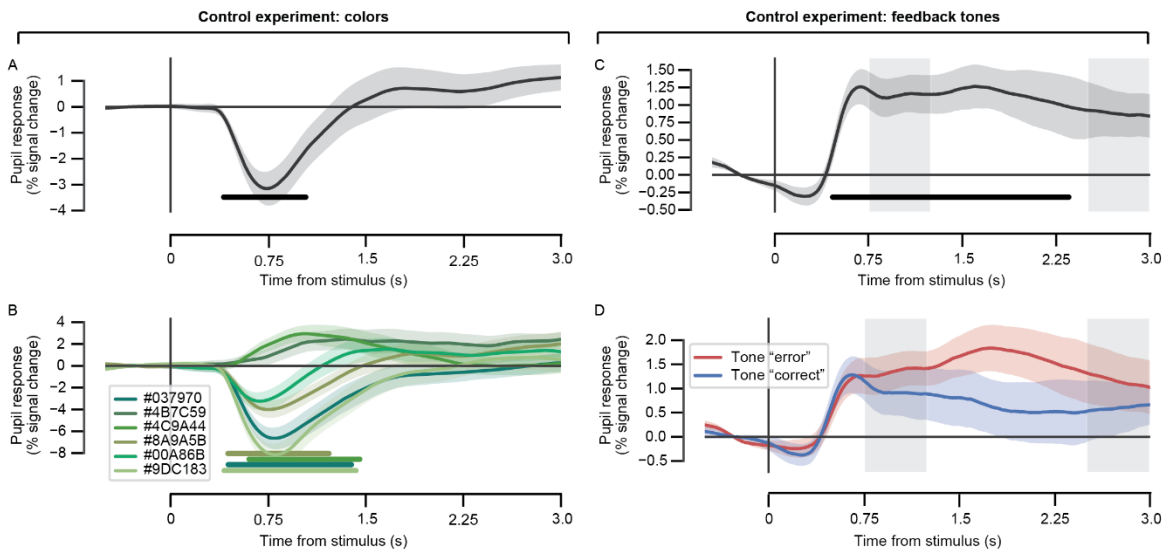
The mapping of feedback tones to indicate accuracy on each trial was as follows: a high tone (4th octave "B") for correct trials and a low tone (3rd octave "D") for error trials. Thus, these tones were not counterbalanced across participants and could potentially confound a main effect of accuracy obtained in the pupil response (although we note that it could not account

for any interaction obtained between accuracy and frequency conditions). Participants were presented with the two tones used as trial-wise feedback on accuracy in the letter-color 2AFC task and instructed to maintain their gaze on the fixation cross in the center of the screen for the duration of the task (no responses were required from the participants). On each trial, a fixation cross was presented for 0.5 s, before the tone was presented for 0.3 s (see also section 2.2), followed by the fixation cross again for the inter-trial interval (3.5 – 5.5 s; uniform distribution). Each tone was presented 50 times for a total of 100 trials while pupil dilation was simultaneously recorded.

Results

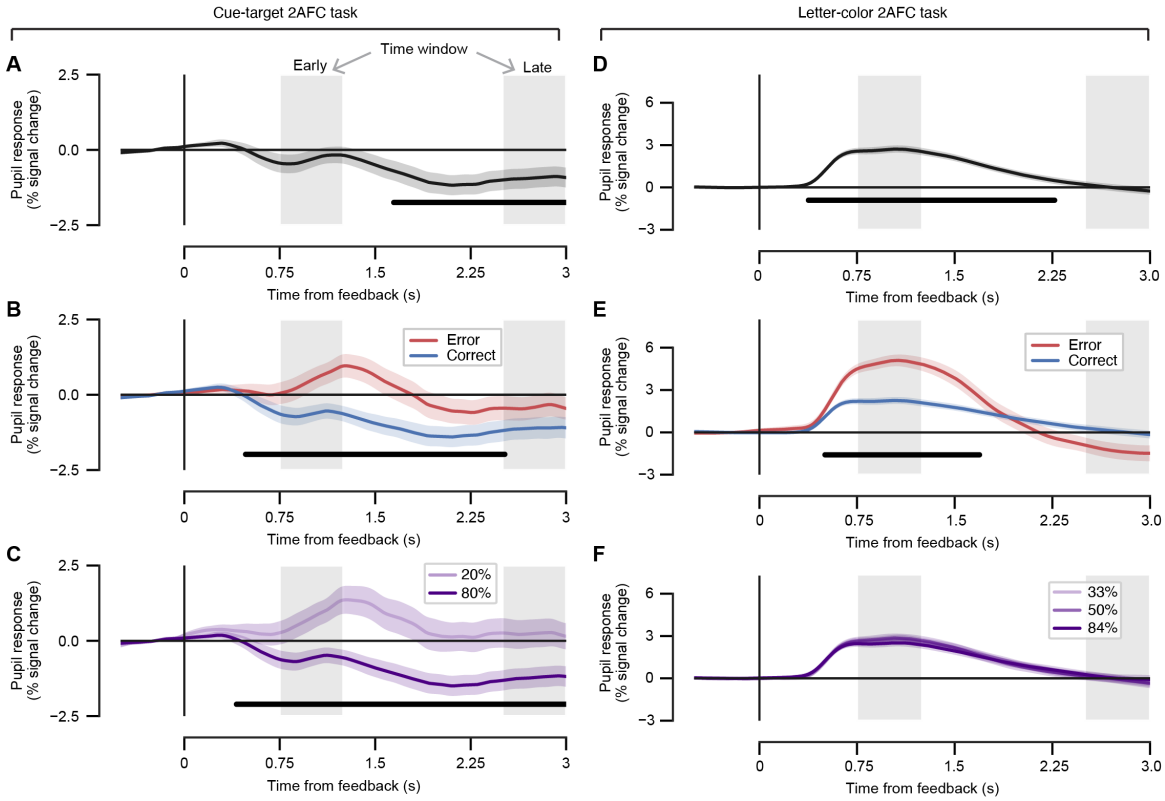
Results showed that the auditory tone dilated pupils on average (Supplementary Figure 1C). Crucially, however, the two tones did not differ from one another in either of the time windows of interest (Supplementary Figure 1D; no significant time points were obtained).

In sum, the control task for feedback tones showed that the pupil responded similarly to the two different tones used in the letter-color 2AFC task. Thus, the different tone stimuli used to indicate error or correct trials would not have accounted for any differences obtained in the pupil's response on error as compared with correct trials.



Supplementary Figure 1. Control tasks for data set #2: letter-color 2AFC task. Left column, results from the control task for colors. Right column, results from the control task for feedback tones. (A) Mean tone-locked pupil response across all trials. (B) Feedback-locked pupil response time course plotted as a function of color used in the main letter-color 2AFC task (hexadecimal codes are given in the legend). (C) as *A,* for the control task for feedback tones. (D) Pupil response time courses plotted as a function of feedback tone used for error and correct trials in the main letter-color 2AFC task. In *C, D*, Gray boxes, time windows of interest; early time window, [0.75, 1.25]; late time window, [2.5, 3.0]. Shading represents the standard error of the mean. The black and green horizontal bars indicate a significant effect of interest (cluster-corrected for multiple comparisons).

Supplementary section 2. Frequency and accuracy effects in pupil time course



Supplementary Figure 2. Main effects of frequency and accuracy in the feedback-locked pupil time courses. Left column, results from the cue-target 2AFC task. Right column, results from the letter-color 2AFC task. (A) Mean feedback-locked pupil response across all trials. (B) Feedback-locked pupil response time course plotted as a function of accuracy. (C) Feedback-locked pupil response time course plotted as a function of stimulus-pair frequency. (D, E, F) as *A, B, C* for the letter-color 2AFC task. Gray boxes, time windows of interest; early time window, [0.75, 1.25]; late time window, [2.5, 3.0]. Shading represents the standard error of the mean. The black horizontal bars indicate a significant effect of interest (panels A-E were cluster-corrected for multiple comparisons; panel F, a one-way repeated-measures ANOVA was conducted on each time point and corrected for multiple comparisons with the false discovery rate).

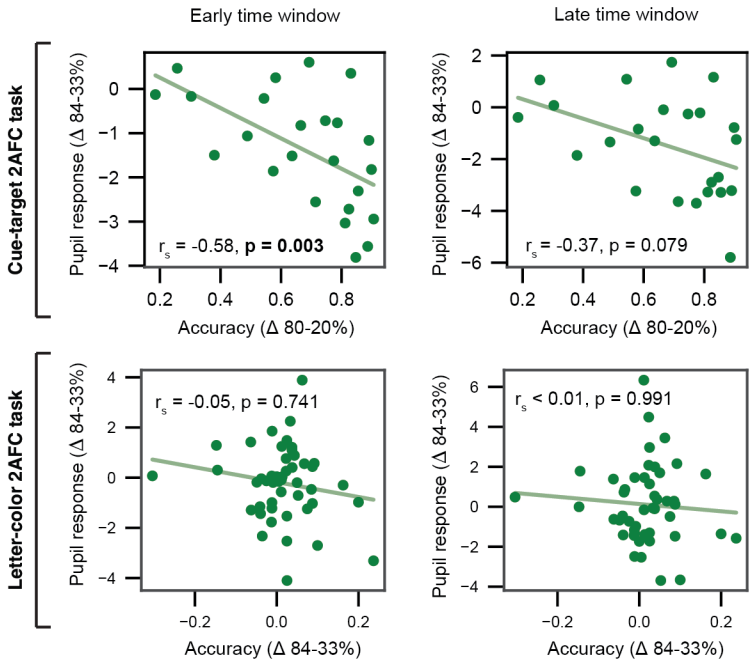
Supplementary section 3. Individual differences analysis between accuracy and pupil responses

Data set #1: cue-target 2AFC task

In data set #1, it was previously reported that pupil responses decreased for the high-frequency (expected) trials as compared with low-frequency (unexpected) over the course of the experiment as a result of learning the cue-target contingences[33] (see their Supplementary Figure 2). In the current analysis, we explored whether a similar relationship between the feedback-locked pupil response and behavioral accuracy was also evident at the level of individual differences across the participant sample. We similarly expected that the feedback-locked pupil response and accuracy would 'mirror' one another in such a way that those individuals who showed a larger difference between the 80% and 20% frequency conditions in accuracy, would also show the larger (absolute) difference between frequency conditions in the feedback-locked pupil response. Within the early time window, a significant negative Spearman correlation was obtained indicating that individuals who showed larger differences between the 80% as compared with the 20% frequency conditions in accuracy also showed smaller differences between frequency conditions in the feedback-locked pupil response (Supplementary Figure 3, top left). The negative direction of the correlation can be explained because the pupil responses are larger on average for the 20% frequency condition as compared with the 80% frequency condition (see Supplementary Figure 2C). A trend toward the same negative scaling could be seen in the late time window (Supplementary Figure 3, top right).

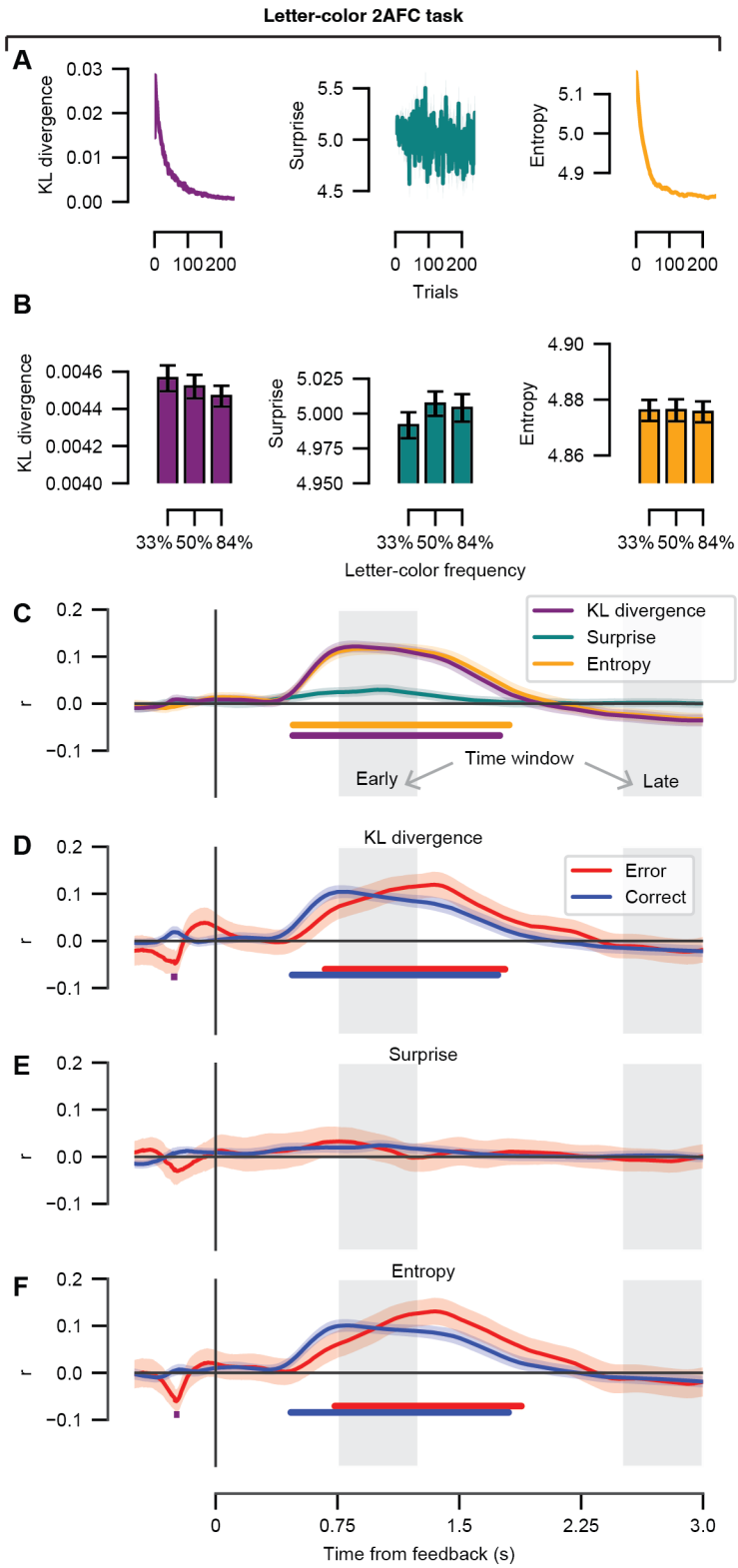
Data set #2: letter-color 2AFC task

We explored whether this relationship between the feedback-locked pupil response and behavioral accuracy was also evident at the level of individual differences during the letter-color 2AFC task. However, in contrast to the cue-target 2AFC task, we did not find any evidence that the frequency effect in task accuracy scaled with the frequency effect in the feedback-locked pupil response in either the early or late time window (see Supplementary Figure 3, bottom row).



Supplementary Figure 3. Individual differences analysis between accuracy and pupil responses. *Top row*, cue-target 2AFC task. *Bottom row*, letter-color 2AFC task. *Left column*, early time window. Right column, late time window. The average feedback-locked pupil response difference (80-20% frequency conditions) are plotted against the difference in accuracy (80-20%). Data points, individual participants.

Supplementary section 4. Ideal learner model fits using a prior distribution in the letter-color 2AFC task



Supplementary Figure 4. Correlations between the feedback-locked pupil response time course and the information-theoretic variables using a uniform prior distribution in the letter-color 2AFC task. (A) The KL divergence, surprise, and entropy parameters are shown as a function of task trial. Model parameter units are in bits. (B) The mean KL divergence, surprise, and entropy parameters are shown as a function of frequency condition. (C) Trial-by-trial correlations between the ideal learner model parameters (KL divergence, surprise, and entropy) at each time point in the feedback-locked pupil response. Shading represents the standard error of the

mean. Gray boxes, time windows of interest; early time window, [0.75, 1.25]; late time window, [2.5, 3.0]. (D) Trial-by-trial correlations between the KL divergence parameter and the feedback-locked pupil response separately for the error and correct trials. (E) As *C*, for the surprise parameter. (F) As *D*, for the entropy parameter. All panels, the colored horizontal bars indicate time periods of significant correlation coefficients tested against zero for each model parameter or condition of interest (cluster-corrected, permutation test). No differences between error and correct trials were obtained (cluster-corrected, permutation test).