Pupil Dynamics Reflect Uncertainty-Driven Adjustments of Probability Learning

Greenhouse-Tucknott, A.[1,2], Foucault, C.[3], Buot, A.[1,2], Meyniel, F.[1,2]

1. Cognitive Neuroimaging Unit, INSERM, CEA, CNRS, Université Paris-Saclay, NeuroSpin center, 91191 Gif/Yvette, France
2. Institut de neuromodulation, GHU Paris, psychiatrie et neurosciences, centre hospitalier Sainte-Anne, pôle hospitalo-universitaire 15, Université Paris Cité, Paris, France
3. Department of Experimental Psychology, University of Oxford, Oxford, United Kingdom

ABSTRACT

Adaptive learning in dynamic environments depends on the brain's ability to represent and respond to uncertainty. Changes in arousal-mediated pupil diameter provide insight into the neural basis of this process. Prior work has focused on magnitude-based inference tasks, and revealed that learning is primarily influenced by the detection of change points and the corresponding modulation of arousal. In comparison, the role of arousal systems in learning contexts where changes are harder to detect, such as in probability learning, remains poorly understood. Here, we combined pupillometry with a probability learning task in which human participants estimated a hidden generative probability of stimuli that changed abruptly and unpredictably. We adopted a Bayesian ideal observer model to estimate, trial-by-trial, the optimal apparent learning rate and the dynamics of two uncertainty-associated factors - namely, change point probability (the probability that the generative process has abruptly changed) and prior uncertainty (about the current beliefs on this generative process). In line with normative theory, participants heavily relied on prior uncertainty to adjust apparent learning rates. Pupil analyses revealed temporally distinct profiles that separate uncertainty factors: phasic dilations tracked change point probability while tonic pupil dilations tracked prior uncertainty. Mediation analyses further indicated that both phasic and tonic pupil signals at least partially carried the effects of change point probability and prior uncertainty on apparent learning rates, respectively. Together, our results demonstrate that probability learning in a dynamic environment is underpinned by computationally rational integration of latent uncertainty factors, implemented through arousal-associated responses.

INTRODUCTION

The capacity to learn and predict future outcomes from past experiences is of paramount importance since it underpins adaptive behaviour [1]. Accurate learning requires sophisticated computations that depend critically on how sources of uncertainty are represented and managed [2]. For this reason, noise and environmental dynamics pose a fundamental challenge to learning.

Stochasticity (*i.e.,* noise) intrinsic in many real-world processes introduces an unavoidable source of uncertainty. In stable environments, the amount of variability becomes predictable over time, making it possible to form reliable beliefs that no longer fluctuate erratically following the stochasticity of observations. For instance, waiting times at a busy restaurant may vary with many unknown factors, but experience allows one to form a general expectation of what the range of waiting times will be to be seated. However, the typical environments we face are not stable but subject to changes (*i.e.,* volatility). One form of volatility corresponds to abrupt changes in the environment, which suddenly render prior beliefs obsolete and should lead learning to reset rather than to stabilize. In the same example, an unusually long wait might signal a drastic change has occurred, such as a change of ownership or staff. Adaptive learning therefore requires one to display mechanisms that discriminate between the opposite influences of stochasticity and volatility on the learning process in order to appropriately utilise new information and maintain accurate expectations about future outcomes [3,4].

The impact of new observations on learning should reflect the informativeness of these observations. Stochasticity and volatility make new observations weakly and highly informative, respectively. The informativeness ascribed by an observer to observations should be reflected in their learning rates. The learning rate controls the speed at which beliefs are updated in response to prediction errors (*i.e.* the difference between an observation and the estimate). Large learning rates prioritize new observations (deemed informative) over past experiences in the update of beliefs. Learning rates should therefore be adapted to the stochasticity and volatility of the environment. In more volatile environments, learning rates should increase, promoting flexible beliefs to respond to rapidly changing contexts. In more stochastic environments, learning rates should decrease, to stabilise beliefs in the face of less informative observations. Human behavioural experiments confirm that both average [5,6] and trial-by-trial [7–10] assessments of learning rates are adapted to environmental uncertainty.

The concept of learning rate is useful to study how learners update their estimates following new observations. In many computational accounts, the learning rate is an explicit model parameter [11,12]. Outside of this, an *apparent learning rate* can be extracted directly from agents' reports of their beliefs. Inspired by the delta rule model, the apparent learning rate is computed as the ratio of the

update (*i.e.* the difference between two consecutive estimates) to the prediction error [7,13]. We use "*apparent*" here because this metric is agnostic of the actual learning algorithm of the agent, and whether it features a parameter akin to a learning rate.

To understand how trial-by-trial adjustments in apparent learning rates are calibrated, it is useful to quantify forms of uncertainty that modulate the informativeness of new observations. To this end, normative (Bayesian) models have been used to extract latent uncertainty factors during learning. The results indicate that human learners update beliefs in response to two forms of uncertainty: the variance in their prior beliefs (*prior uncertainty*) and the probability of an abrupt environmental change (*change point probability*) [14–20].

Current understanding of how latent uncertainty factors drive trial-by-trial changes in apparent learning rates have been largely informed by the utilisation of predictive inference tasks that employ magnitude-based quantities, such as spatial positions [17,19,20] or numerical values [18]. In such cases, apparent learning rates are highly determined by dynamics in change point probability [14]. However, this does not generalise to all learned quantities. For example, magnitude learning (*i.e.* learning the size or location of an outcome) is fundamentally different to probability learning (*i.e.* learning the probability of discrete outcomes). In probability learning, each observation carries much less information about the inferred quantity compared to magnitude learning. This difference promotes a shift in the dominant driver of apparent learning rates from change point probability to prior uncertainty [14].

The neural mechanisms governing the uncertainty-dependent regulation of learning remain an area of intense interest. Neuromodulatory systems have long attracted attention as putative mechanisms, considered to play a central role in striking a balance between stable versus flexible behaviour [3,21–23]. In humans, there has been a heavy leverage of nonluminance-dependent, pupil dilations - used as a non-invasive proxy of central arousal state [24–28] - to understand how uncertainty is represented and shapes cognitive processes such as learning and decision-making [18,29,30]. During learning, consistent associations have been reported between phasic pupil dilations and the amount of prediction error and/or surprise associated with new observations [6,18,29,31–33]. Indeed, Nassar and colleagues (2012) demonstrated that phasic pupil dilations track change point probability during a magnitude inference task, while a direct manipulation of arousal state - forcing transient increases in pupil diameter - was also shown to evoke concomitant increases in apparent learning rates. Importantly, the efficacy of this causal manipulation was dependent upon the baseline, tonic pupil diameter [18]. This may reflect another form of uncertainty-weighting, as increasing evidence has associated tonic pupil response with tracking prior belief uncertainty [18,33–37]. These phasic- and tonic-associated pupil effects are independent and can be experimentally dissociated [38], suggesting

functionally separable roles for phasic and tonic arousal signalling in learning regulation. Changing tonic/phasic pupil profiles may therefore signal how arousal systems shape adaptive learning to the environmental context.

The knowledge gap that we tackle here is the role of arousal systems in probability learning where, unlike in magnitude learning, change points are harder to detect. More precisely, the goal is to test whether the encoding of different uncertainty factors in phasic and tonic pupil-linked arousal (which is supported by prior studies) regulates, on a trial-by-trial basis, the apparent learning rate in probability learning (which remains unknown). To address this outstanding question, we employed a previously developed probability learning task [14] to continuously capture participants' apparent learning rate, coupled with pupilometry to concomitantly assess changes in pupil diameter. The Bayesian optimal solution to the inference problem [33] was applied to not only provide a reference comparison for the analysis of participants' behavioural estimates, but also extract trial-by-trial latent uncertainty factors. These uncertainty factors were then used to evaluate drivers of learning regulation and the modulation of pupil dilatory responses. Overall, our results support the hypothesis that during probability learning, prior uncertainty dominates in the regulation of learning. Prior uncertainty and change point probability are separately encoded within tonic and phasic pupil signals, respectively, and these identified pupil signals mediate the effect of the associated latent uncertainty factor on the regulation of learning, providing a physiological mechanism supporting belief updating during probability inference.

RESULTS

The probability learning task required participants to continuously infer the value of a hidden state - in this instance, the generative stimulus probability - from a sequence of stimuli [14]. We used a cover story for the task to improve comprehension. The cover story described a spinning roulette wheel which was divided into two colours (*i.e.* blue and yellow). The proportional distribution of the colour blue corresponded to the hidden generative probability. Participants were tasked with trying to estimate the proportion of blue (*vs.* yellow) on the wheel as accurately as possible based on a sequence of the two colours presented to them. The sequence presented was described as the result of repeatedly spinning the wheel. Participants continuously updated their estimate of the hidden proportion after each new observation by moving the marker on a semi-circular slide tracker (*Fig 1*). Each sequence consisted of 75 observations, with observations presented as a coloured circle with a white target overlaid on top. Participants were instructed to maintain a fixed gaze on the target for the accurate recording of pupil diameter to each new observation throughout the task. The semi-circular slide tracker made fixation easier (see *Methods* for more details).

During each sequence, the generative probability (and thus the proportion of the two colours on the fictitious roulette wheel) could change at unpredictable times, which were referred to as change points. Participants were aware of the possibility that the proportional distribution of the colours could change, but were not signalled when they happened during the observation sequence. The generative probability was uniformly sampled within the range of 0.1 and 0.9, preventing long streaks of observations where no updates are required. The magnitude of the change point was constrained, to avoid changes that are too subtle to be perceived (see *Methods* for more details). As described previously, the task is engaging and participants are capable of tracking changes in the generative probability [14].

{Fig 1}

Participants continuously adapt estimates in a way that closely follows optimal inference

We employed a Bayesian ideal observer model to provide an optimal solution to the inference problem as a principled point of comparison [39]. The model returns a posterior distribution of the hidden generative probability that is updated upon each new observation (*Fig 2B*). We also extracted the dynamics of two latent, uncertainty-associated factors: prior uncertainty and change point probability (*Fig 2F and 2G*). Here, prior uncertainty is quantified as the standard deviation of the hidden probability estimated just before the current observation, while change point probability reflects the probability a change in the hidden value has occurred since the last observation - conceptually akin to the magnitude of prediction error (for further specification see *Methods*).

Evaluating participants' performance across the entirety of the task, we found stable performance across sequences (see *Supplementary Material, Fig S1*). Our primary analysis of performance focussed on how well participants' estimates compared to the optimal solution. Participants displayed a strong ability to accurately estimate the hidden generative probability given the relationship with the mean of the model's posterior distribution (mean ± s.e.m; Pearson's correlation = 0.81 ± 0.02; $t_{(30)}$ = 44.72, $p \approx 1 \times 10^{-28}$; *Fig 2C*; for more details on the relationship see *Supplement Material, Table S1*). Consistent with instructions, participants updated their estimate on most trials (mean ± s.e.m; 62.3 ± 3.49%). The frequency of updates was not associated with the strength of relationship between participant and model estimates (Pearson's correlation: 0.003, p = 0.987). The participants' apparent learning rate significantly correlated with the apparent learning rate derived from model estimates (Pearson's Correlation = 0.07 ± 0.02; $t_{(30)}$ = 4.43, p = 0.0001; see *Supplement Material, FigS2 and Table S1*). This correlation, although reliable across participants, was weak in magnitude; which may reflect noise in apparent learning rates at the trial level. This noise may, at least in part, be explained

by trial-specific "lags" in how participants' updates their beliefs, which we highlight in analysis presented below.

{Fig 2}

Prior uncertainty dominates in driving apparent learning rate adjustments

The optimal solution prescribes that learning should not be constant but dynamically adjusted from trial to trial. Around change points, participants displayed a protracted increase in their apparent learning rate, which was significantly different from baseline learning rates (i.e. before the change point) from the third to the eleventh observation after a change (*Fig 3A*). This pattern conformed with normative theory (*Fig 3A insert*) (group-level correlation between participants apparent learning rate and model-derived learning rates around change points: group-level Pearson's correlation = 0.95; participant-level Pearson's correlations: 0.48 ± 0.06; $t_{(30)} = 7.88$, $p \approx 1 \times 10^{-9}$), replicating a previous study [14].

Normative theory also prescribes that both latent uncertainty factors should play a role in determining the apparent learning rate, with prior uncertainty dominating. We initially found partial support for this prediction (*Fig 3D*). In line with normative theory, prior uncertainty was a significant driver of participants' apparent learning rate (beta = 0.09 ± 0.01; $t_{(30)} = 6.79$, $p \approx 1 \times 10^{-7}$). This could be clearly seen when examining the dynamics of prior uncertainty around change points, which paralleled the participants' apparent learning rate (*Fig 3B*). We also observe strong graded co-variation between prior uncertainty and participants' apparent learning rate across all trials (*Fig 3C*). The effect of prior uncertainty was significantly larger than the effect of change point probability ($t_{(30)} = 4.48$, $p \approx 1 \times 10^{-4}$), as predicted by normative theory.

However, although normative theory prescribes that the effect of change point probability on the apparent learning rate should also be significant, although smaller than that of prior uncertainty, our initial analysis did not reveal such a significant effect (beta = -0.02 ± 0.02; $t_{(30)} = -0.80$, $p = 0.432$; *Fig 3D*). A closer inspection of participant's behaviour revealed that subjects sometimes lagged slightly behind the optimal model. Such occasional lags should not harm much the effect of prior uncertainty: being fairly smooth from trial to trial (see *Fig 2A*), a strict trial-by-trial alignment between participants and the model is not necessary. In contrast, the change point probability peaks sharply only on a few trials, making trial-by-trial alignment more critical. To assess whether trial-wise (mis)alignment influenced the observed relationship between latent uncertainty factors and learning rate, we implemented an optimization procedure that selectively shifted some subjects' updates backward by one trial to compensate for a potential delay with respect to the optimal model. For each participant,

we searched for the configuration of trial shifts that maximized the correlation between the latent uncertainty factors and the (shifted) apparent learning rate. This procedure improved the correlation with apparent learning rates for each latent uncertainty factor beyond chance level (both $p < 1 \times 10^{-308}$, see *Methods* and *Supplementary Material, Fig S3*). The absolute correlation coefficient was slightly larger for prior uncertainty compared to change point probability, in keeping with normative predictions, though we do note that relative to the null distribution there was actually a greater effect of change point probability compared to prior uncertainty (see *Supplementary Material, Fig S3*). Nevertheless, our optimisation analysis indicated that occasional lags in probability reports can obscure the true relationships between uncertainty factors and participants' apparent learning rate. The number of shifts made with respect to optimising prior uncertainty and change point probability were correlated suggesting that there were genuine lags in the participants' updates (see *Supplementary Material, Fig S3*). We note that identified lags concerned less than half of the trials (mean ± s.e.m; shifts for change point probability: 45.6 ± 0.7%; prior uncertainty: 43.3 ± 0.5%). Accordingly, all subsequent analyses of pupil dilation were performed using the raw uncertainty factors and apparent learning rates.

{Fig 3}

Uncertainty factors are encoded within distinct pupil dilatory responses

To examine the association between uncertainty factors and arousal-linked neuromodulation, we ran a multiple linear regression on pupil diameter signals for each peri-stimulus time point in a window defined around an observation (-250 to 1400 ms). The regression model included as predictors prior uncertainty and change point probability, as well as multiple covariates (see *Methods*). The analysis was performed on both phasic and tonic pupil signals, defined as changes relative to a pre-stimulus baseline (-250 to 0 ms) and the absolute (i.e. nonbaseline-corrected) signal, respectively [33]. We demonstrate separable uncertainty effects between the distinct pupil-dilatory signals: phasic changes in pupil diameter displayed only an effect of change point probability, while the tonic pupil dilatory signal was influenced only by prior uncertainty. The effect of change point probability on the phasic signal emerged 630 ms after the observation presentation until the end of the epoch (*Fig 4B*), which corresponded to a secondary dilatory phase in the baseline-corrected signal conventionally attributed to cognitive processing post orientation [40] (*Fig 4A*). The effect on the tonic signal was evident across the entire duration of the epoch, which importantly included the period prior to observation presentation signifying the presence of accumulative processing arising from past experiences (*Fig 4C*). In other words, the tonic pupil signal reflected prior uncertainty even before observation onset.

{Fig 4}

To evaluate whether these effects could be attributed to some artefact associated with increased arousal instigated through moving the cursor in response to a new observation, and establish more broadly the generalisability of the observed effects, we ran the same analysis on pupil dilatory responses obtained during a previous study conducted using auditory stimuli (n = 15). The structure of this task was extremely similar to the main study, with one important difference being that participants did not provide an explicit estimate after each observation (see *S1. Appendix* for further task details), but only occasionally on trials excluded from the analysis. The same normative model and regression analysis described in the main study were applied to the observed sequences. The results of the analysis demonstrated the same coupling between pupil signals and uncertainty factors: phasic changes in pupil diameter displayed an effect of change point probability and tonic changes were modulated by prior uncertainty (see *Supplementary Material, Fig S4*). Therefore, the observation that distinct uncertainty factors exert influence on distinct aspects of the pupil response are amodal (not specific to either vision or audition) and linked to the learning process itself, rather than behavioral reports.

Encoding of uncertainty factors in pupil size modulates participants' apparent learning rate

We next examined whether the identified pupil signatures of uncertainty factors were predictive of participants' apparent learning rate. To achieve this, we extracted the average pupil signal corresponding to the observed significant uncertainty factor effects (*Fig 5A*). For the phasic pupil response (*i.e.* baseline corrected) the average was computed from 630 to 1400 ms relative to the observation, where the effect of change point probability was identified. For the tonic pupil effect (*i.e.* non-baseline corrected) the average was taken from the period -250 ms to 0 ms before an observation, where the effect of prior uncertainty was significant. The phasic and tonic effect sizes were close in magnitude and not significantly different (phasic - beta: 0.12 ± 0.02, tonic - beta: 0.14 ± 0.02; $t_{(27)}$ = -1.19, p = 0.246; *Fig 5B insert*). However, this observation is compatible with prior uncertainty being the dominant factor in the regulation of the apparent learning rate. This is because the similar effect sizes are estimated with normalized (z-scored) changes in tonic and phasic signals, but it should be remembered that changes in tonic pupil size accounted for by prior uncertainty are three times larger than the changes in phasic pupil size accounted for by change point probability (compare effect sizes in *Fig 4B* and *4C)*.

We note the change in the tonic pupil signal around change points displayed a qualitative resemblance to the corresponding change in prior uncertainty (*Fig 5D*). Given the dominant effect of prior uncertainty on participants' apparent learning rate, coupled with the observation that the associated arousal representation of this uncertainty factor in the tonic pupil response was also predictive of apparent learning rates, we then ran a formal mediation analysis of the effect of prior uncertainty on

learning rate with the tonic pupil signal as a mediator (see *Methods* for more information). We demonstrate that the effect of prior uncertainty on participants' apparent learning rate can, at least in part, be attributed to an effect mediated by tonic pupil responses (*Fig 5F)*. The change in the phasic pupil signal appeared to track change point probability around change points, though visually this was not as apparent as the association between prior uncertainty and the tonic pupil signal (*Fig 5C*). Yet, mediation analysis indicated that change-point probability exerted its influence on participants' apparent learning rates via this pupil signal, with an indirect effect present despite no detectable total or direct effect on learning itself (*Fig 5E*). The results indicate that the activity profile of ascending arousal systems play an important role in encoding distinct latent uncertainty factors that shape the integration of new information during probability learning.

{Fig 5}

DISCUSSION

This research utilised an inferential task to understand how arousal systems, indexed through pupil dilatory responses, represent latent uncertainty factors and shape human probability learning. Similar to normative theory, we found that uncertainty (or its reciprocal, confidence) in prior beliefs is leveraged to update beliefs following new observations in an environment that makes changes difficult to identify [14]. We also add to a body of accumulating evidence that identify a separation in how latent uncertainty factors underlying statistical learning map onto arousal profiles, with tonic and phasic pupil signals covarying with prior uncertainty and change point probability, respectively [18,33,34,37]. Furthermore, we demonstrate that both of these pupil signals, at least in part, play a role in mediating how latent uncertainty factors are transformed into belief updates during human probability learning. Our results suggest that, within the context of probability learning, belief updates are driven by the representation of separate normative uncertainty factors (*i.e.* prior uncertainty and change point probability) across distinct activation modes of the ascending arousal systems.

We relied on a normative modelling framework to characterise how latent uncertainty factors shape learning in a dynamic environment. Consistent with theoretical predictions, probability learning was dominated by prior uncertainty, replicating a previous study [14]. Within this Bayesian formulation, prior uncertainty is epistemic in nature: it reflects an estimate of the lack of confidence in current beliefs about the environment [41]. Our results therefore support the view that adaptive learning is governed by a confidence-weighting mechanism, whereby belief updates are scaled according to uncertainty about prior expectations [7,13,18,33]. Several lines of evidence support the notion that the brain represents a computationally rational form of confidence. Here, "rational" means it is qualitatively similar to the Bayes optimal solution. This confidence that regulates learning seems

accessible to introspection, as model-based confidence correlates with subjective confidence reports [42,43]. Converging neuroimaging evidence has also identified similar neural correlates of this confidence signals across multiple tasks [17,43–49]. More broadly, our behavioural results may also suggest that the influence of confidence on learning is context-sensitive, becoming more prominent when new observations are weakly informative, as in probability learning, compared to contexts with higher informational gain, such as when learning about magnitudes [14].

At the mechanistic level, confidence may modulate intrinsic properties of brain states to amplify or dampen belief updates to surprising observations [33]. One mechanism by which this may occur may be arousal state, which subserves the emergence of large-scale patterns in many neural subsystems [28]. Our observed pupil dilatory responses indicate that this confidence-weighting mechanism, as well as the perception of change points, is dependent upon arousal state and neuromodulation. In line with previous research, we demonstrate that the representation of uncertainty factors during probability learning adheres to a separation of prior uncertainty and change point probability between tonic and phasic arousal responses, respectively [18,34,37]. We confirm this separation within the context of probability learning [33]. This separation was not an artefact induced through participants' reports as confirmed by similar effects observed in our supplementary analysis of an audio probability inference task requiring no overt response. Previous work has shown that both change-point probability and relative uncertainty (*i.e.* prior uncertainty) influence apparent learning rates through shared variance expressed in the P300 EEG component [15,16]. The P300 component has been proposed as the cortical index of a broader orienting response to novel and uncertain events, consistent with its covariation with phasic autonomic measures such as pupil dilation and skin conductance [50]. More precisely in computational terms, the P300 may index the updating process itself [51,52], and therefore recapitulate the multiple factors that influence learning. In contrast, our findings show that tonic and phasic arousal states do not reflect the update itself, but the dissociable uncertainty factors that regulate belief updating, by selectively tracking prior uncertainty and change point probability, respectively.

Our results align with the findings of Nassar and colleagues (2012), who showed in a magnitude inference task that outcome-evoked pupil dilation over a period of 0-2 s after feedback encodes change-point probability, whereas mean pupil size over the same period reflects relative uncertainty. In their study, outcome-evoked pupil responses predicted participants' learning rate, whereas average pupil size was a non-significant predictor. Here, we extend these findings, identifying pupil-linked-arousal as an important regulator of apparent learning rates within the context of probability learning. We demonstrate that in probability learning, tonic pupil responses are predictive of adopted learning rates. The stronger contribution of this tonic component — relative to the non-significant effect of average pupil size reported by Nassar *et al.*, (2012) — may reflect the greater

importance of prior uncertainty in probability learning compared with magnitude learning [14]. The findings together suggest effective regulation of learning, be that probabilities or magnitudes, would appear to be dependent upon fine tuning the balance between tonic versus phasic activation of arousal systems in a way that is appropriate for a given context.

Non-luminance-dependent pupil dilations, reflecting global arousal, are widely used as an indirect index of locus coeruleus (LC) activity and noradrenaline (NA) release [24,53]. The strongest evidence for this link comes from studies showing simultaneous activation of LC firing and pupil responses and demonstrations that LC microstimulation evokes pupil dilation [25,26,54,55]. In humans, pupil diameter covaries with LC BOLD activity both at rest and during cognitive tasks [56]. Although non-luminance-dependent pupil dilations are not exclusively controlled by the LC-NA system [25,26,57,58] and are state dependent [59], our results can speculatively be interpreted in relation with LC activity, which we now discuss.

Several theoretical accounts propose that the LC–NA system plays a central role in behavioural adaptation, particularly to environmental change [3,60–63]. Supporting a broad interpretation of these perspectives, LC activity tracks changes in environmental contingencies and often precedes behavioural adjustments [64–66], while causal manipulations of the NA system — including LC optogenetic activation, prefrontal NA deafferentation, and NA-targeted pharmacological interventions — modulate flexible learning and choice variability [65,67–73]. Our finding that participants' apparent learning rates were predicted by both tonic and phasic signals is compatible with the hypothesis that the brain may exploit distinct LC activity modes to adapt learning to environmental context. In the probability learning task used here, the dominance of prior uncertainty in shaping learning rates aligns with adaptive gain theory [74]. Within this framework, tonic LC activity modulates the selectivity of sensory processing: increased tonic activity is described as a "distractible" mode, defining a broad attentional state based on a global increase in neural gain where new information can more effectively challenge current beliefs. Consistent with this account, larger baseline pupil responses are predictive of individuals engaging in favouring exploration over exploitation in rewarded behavioural tasks [75–77]. Tonic pupil size also correlates with a broad attentional focus and these pupil changes are predictive of the strength of fMRI responses, as would be predicted by a global modulation of neural gain [78]. Accordingly, during probability learning, a high tonic arousal state may be preferable, especially after change points, as information is sought from new observations limited in their informativeness.

The LC also plays an important role in memory [79–81]. Phasic pupil responses have been shown to demarcate episodic memories [82]. This is compatible with the demonstration that the observed phasic pupil response is sensitive to the probability of changes in the generated sequence. Yet, higher tonic

LC activity ameliorates phasic dilation effects at event boundaries, which may hold implications for episodic memory formation [82]. High tonic arousal may change how information is processed by engaging higher-level transmodal regions, including the hippocampus [83], which may emphasise the importance of functions such as memory in the interpretation of the salience of new information. Here, a high tonic arousal state due to increased uncertainty in prior beliefs may require such engagement to track relationships between clusters of observations in order to build confidence in inferences concerning the underlying statistical structure of the environment. To this end, the activity profile of the LC may provide the gating mechanism defining precisely how the brain appropriately configures to perform predictive inference in changing environments.

We now discuss the choice of model used in our analysis. We used a hidden Markov implementation, also known as Bayesian filtering, of an ideal Bayesian observer [33] which assumes the brain computes the full posterior over possible outcomes. In a previous study [84], this model was reported to best fit behavior in another dataset using the present probability learning task. The benchmark included a range of alternative models — both approximately Bayesian (*e.g.* hierarchical Gaussian filter, volatile Kalman filter, adaptive mixture of delta rules, reduced Bayesian, and change point models) and clearly non-Bayesian (*e.g.* proportional-integral-derivative controllers, adaptive delta rules). More precisely, this previous study used a parameterized ideal observer model, in which the volatility assumed by the model was subject-specific, rather than matching the generative value. We explored this possibility here but it did not improve the results substantially. This computationally demanding form of learning required by the Bayesian model poses the question of whether its accuracy is worth the cost. Across tasks, recent evidence suggests that higher-level systems may arbitrate over multiple algorithmic implementations of increasing complexity, dictating which controller drives behaviour given the complexity of the task [85]. There appears a small window across stochastic, low volatility environments, where the computational cost of the Bayesian model is worth the added accuracy that past observations offer in terms of future predictions [86]. In the present task, the generative volatility was 0.05 (one change point every 20 observations, on average), which resides within a space that could benefit from using a Bayesian model [86]. But, this in turn raises the longstanding question of the biological feasibility of Bayesian inference [87–89]. Our lab has recently provided evidence that an artificial neural network displays the capacity to solve probability learning tasks quasi-optimally with a very small number of units (about 3 to 10), arguing in favor of the biological plausibility of such computations, but only when the network had a gated, recurrent neural architecture [90]. LC activity may provide a gating mechanism in brain circuits, akin to the one of artificial recurrent gated neural networks, that enables confidence-weighted learning.

We note that the tonic pupil response only partially mediated the effect of prior uncertainty on participants' apparent learning rates, indicating that additional factors contributed to either pupil

dynamics or learning behavior. Our mediation analysis may not capture the nonlinear coupling of arousal to brain states dynamics [28] which may underlie learning. Moreover, areas such as the anterior cingulate cortex (ACC) have been implicated in the updating of internal models and control of learning rate [5,31]. Top-down processing, including from the ACC, influences pupil-linked-arousal [25,91], but may also have other effects on learning, which would result in the tonic pupil size mediating (in the statistical sense) only partly the effect of prior uncertainty onto the apparent learning rate. Regarding pupil measurements, several limitations should be considered. Pupil diameter can be affected by interactions with other autonomic processes. Recent evidence shows that pupil constriction and dilation correspond to inspiratory and expiratory phases of breathing, respectively [92], which may confound interpretations of pupil responses to latent uncertainty. Additionally, our protocol does not allow the identification of any potential strategies participants may have employed to perform the task [38], which may have required more or less mental effort that may also be captured within pupil dilatory responses [93]. A further limitation may concern our definition of the tonic pupil response. In the present study, tonic and phasic pupil signals were operationalized as non-baseline and baseline corrected dilatory responses, respectively. With short interstimulus intervals (here, 1.5 s), this approach may inflate tonic estimates due to cumulative phasic responses [94]. Emerging algorithms may offer improved decomposition of tonic and phasic components, but require further validation against traditional approaches across more complex cognitive tasks.

In summary, our results support human probability learning to be based, at the computational level, on a qualitatively optimal integration of latent uncertainty factors, with confidence in prior beliefs serving as the primary driver of belief updating. Importantly, we show that this uncertainty-guided regulation of probability learning is implemented via arousal-related responses in dynamically changing environments.

METHODS

Code and data availability

The collected data, code for the experiment, and the analysis code are all available on the following GitHub repository:
https://github.com/TheComputationalBrain/Adalearn/tree/main/Project/adalearn_pupil

Participants

Thirty-two participants were recruited via *the Paris Brain Institute PRISME Core Facility (RRID:SCR_026394), Paris, France.* They provided written informed consent prior to participation. The study was approved by the Ethics Committee of the Paris-Saclay University (Comité d'Ethique de la Recherche; reference: CER-Paris-Saclay-2023-010) and took place in the PRISME behavioural facility. To evaluate participants' proficiency with the task, we used the correlation coefficient between participants and ideal observer model estimates as a performance criterion. A negative relationship was taken as evidence that the participant failed to fully comprehend the task instructions. One participant displayed a negative correlation (Pearson's correlation = -0.16) and was therefore removed from all further analysis. For the remaining 31 participants (17 female), the average age was (mean ± standard deviation) 38 ± 11 years, ranging between 19 and 63 years. Participants spent 4 ± 3 years in higher education, and 8 reported additional mathematical training as part of their education or occupation. All participants reported normal, non-corrected vision. To ensure the quality of the analysed pupil signal, checks on the degree of missing data led to the exclusion of 2 participants from analyses involving pupil diameter (further explained in *Pupil diameter recording and preprocessing*). A further participant was removed from all pupil analysis due to a failure in the recording of the pupil data. Accordingly, 28 participants were included in the pupil-based analyses, while all other analyses were performed on 31 participants. Participants received 25€ for participation, in addition to a performance-based bonus up to 5€.

Task

The probability learning task has previously been described [14]. One hundred and fifty sequences were generated for the task [14], which were randomly sampled without replacement from these sets. The generation of the sequences was based on a Bernoulli process, whose parameter (*i.e.* the hidden generative probability) changed with a probability of zero for the first six observations following a change point and 1/20 for all trials thereafter. As previously described, the generative probability was sampled uniformly between 0.1 and 0.9 both initially and after a change, with the latter subject to the constraint that the resulting change in the odds-ratio be no less than fourfold.

Task sessions required participants to continuously estimate the hidden generative probability (*i.e.* proportion of blue on the fictitious roulette wheel) from a sequence of 75 observations, by moving the slider along the track in response to each observation. The interval between the onset of each new observation was 1.5 s, with stimulus offset occurring at 1.3 s. in the 1.3-1.5 s inter-stimulus interval. The slider was highlighted to indicate that the participants' estimate was recorded, reflecting its position on the slide tracker at 1.5 s post-stimulus. At the end of each sequence participants received feedback, used to maximise engagement with the task. The feedback displayed participants' performance on that sequence (quantified as a percentage score), performance-based monetary gain

(which was proportional to their score), the number of change points within the sequence, the real generative probabilities and participants' own estimate in relation to those changes. Participants' score for a sequence was calculated based on their mean absolute error, evaluated in relation to two reference points: 1) the ideal observer model estimates (high score), and 2) a 'stay still' strategy (low score). These raw scores were then transformed with a softplus function to place them on a smooth, positive scale, making it easier to compare participants' performance relative to these benchmarks.

The task comprised two steps: 1) instructions and training, 2) task sessions. In step one, participants read through the provided instructions before completing a single sequence of the task. Verbal assessment of understanding identified if a further sequence was required for full comprehension. In step two, participants completed 22 task sessions of the probability learning task (*Fig 1*), split into two equal blocks of 11 sequences. The task sessions took between 45-55 mins to complete.

Several modifications were made to the original design to improve its compatibility with pupil diameter recordings. The visual angle of the slider was set to 3° from the centre of the monitor based on a typical sitting distance (approximately 60 cm). This enabled the whole of the slider to be seen by the participant while they maintained a fixed gaze on the presented stimuli, limiting large saccades during the task and any potential distortion of the pupil diameter recording [95]. A semi-circular slider track was preferred to the original horizontal track to facilitate the smaller visual angle and maintain the same distance to all points along the edge of the track from the middle of the monitor in the eventuality of saccades (*Fig 1*). The horizontal position of the slider was mapped onto the semi-circular track, moving from '0' (100% yellow) to '1' (100% blue) from left to right, respectively. The segments of the slider track between "0–10%" and "90–100%" were filled in order to indicate that the hidden probability never lies within these intervals. Each observation was presented in the middle of the slider track with a fixation target overlaid. Blue and yellow colours were maintained in the task, but hues were adjusted in order to match luminosity, calibrated using a photometer and correcting for the individual systems' gamma curve (colours presented in *Fig 1* do not accurately reflect the colours presented in the task and are displayed as such simply to emphasise the two different outcomes). The task was presented on a homogenous grey background.

Behavioural analysis

The apparent learning rate in response to each new observation (a_t) was quantified using a delta rule [7,13] based on the degree estimate beliefs (p) were updated (p_t - p_{t-1}) in proportion to level of prediction error (x_t - p_{t-1}):

$$a_t = (p_t - p_{t-1}) / (x_t - p_{t-1})$$

(1)

In keeping with previous studies [14,18,20], outliers arising occasionally due to prediction error being very close to 0 were removed from the analysis. Cut-off thresholds of 1.3 and -0.6 were adopted following a previous study [14].

Normative model

The ideal Bayesian observer model applied to each sequence is described previously [14,33] and is openly available at the following repository: https://github.com/TheComputationalBrain/TransitionProbModel.

Briefly, the model infers the posterior distribution of the hidden generative probability (h_t) from past observations ($x_{1:t}$) using a Markovian update rule governed by volatility (v). It assumes knowledge of the true volatility level and that, when a change occurs, the new generative probability is sampled uniformly from a proposal distribution, $p_0(\theta)$, within designated bounds [0.1, 0.9]:

$$p(h_t \mid x_{1:t}, v) \propto p(x_t \mid h_t)(vp_0(h_t) + (1 - v)p(h_{t-1} \mid h_{t-1}=h_t, x_{1:t-1}, v)) \tag{2}$$

To compute this posterior numerically, the model discretises the probability space over a grid with a resolution of 20.

This model provides a normative account of the optimal solution to the inference problem, and enables the extraction of observation-by-observation changes in the model-derived apparent learning rate and latent uncertainty factors. Two factors were specifically examined: 1) prior uncertainty, the uncertainty of participants about the latent probability before seeing the new observation (also referred to as epistemic uncertainty, or inferential uncertainty Meyniel, Schluneger & Dehaene 2015), which we quantified as the standard deviation of prior probability distribution:

$$u_t = SD[h_t \mid x_{1:t-1}] \tag{3}$$

2) change point probability, the probability a change in the hidden state has occurred since the last observation, which is associated with the degree of divergence between the likelihood of new evidence and previous beliefs:

$$\Omega_t = p(h_t \neq h_{t-1} \mid x_{1:t}) \tag{4}$$

Lag optimisation analysis

To assess whether belief updating lagged behind the influence of latent uncertainty factors, we implemented an optimisation procedure that evaluated whether shifting trial-wise behavioural estimates improved their correlation. Within each sequence, estimates at each trial were permitted to take the value of the immediately following trial (excluding the final trial of the sequence), thereby modelling a one-trial delay in the behavioural response. To prevent double-counting of information, when a trial's estimate was reassigned to a preceding trial, the original estimate was removed unless it was itself replaced by an estimate from a later trial. The optimisation algorithm proceeded through ten iterations, during which the shift status of each movable trial was tested in random order. A trial's update was temporarily replaced by that of the next trial, and the change was retained only if it increased the overall correlation between the participant learning rate and a given uncertainty factor; otherwise, the shift was reverted. The procedure converged upon a stable trial-shift pattern that maximised this correlation. To evaluate whether the observed improvement could arise by chance, a null distribution of optimised correlations was generated for each participant. For this, the participant's update values were held constant, while the latent uncertainty factor and error terms were drawn from other participants (n = 30). Each null pairing underwent the same ten-iteration optimisation procedure, providing an empirical distribution of chance-level correlations. The difference between the participant's optimised correlation and the mean of the null distribution quantified the strength of the lag-adjusted association beyond chance. This approach ensured that any apparent temporal alignment between the behavioural and latent computational variables could not be attributed to random correlations or to artefacts of the optimisation process itself.

Pupil diameter recording and preprocessing

Pupil diameter signals were recorded from both eyes using a Tobii Pro Nano (n = 24) or Tobii Pro Fusion (n = 4) at a sampling rate of 60 Hz. The eyetracker was positioned approximately 10-15 cm under the monitor. The onset of observations were integrated into pupil signal recordings through the use of a websocket connection, facilitating time-locked pupil diameter analysis of each new observation. Our analysis was performed on a nominated eye (*i.e.* the right eye). We checked whether the results held when using the eye with the least missing data for each participant and confirmed that it did, so proceeded with analysis performed on a nominated eye. Preprocessing of the pupil diameter signal was performed using a custom script. Blinks were delineated (adding a margin of 100 ms

before and after) with the data linearly interpolated in between. The signal was filtered using a low-pass, bi-directional, 4th-order Butterworth filter (5 Hz), before being epoched within -250 to 1400 ms around each new observation onset. The pupil signal was then z-scored per participant.

Epochs with more than 20% interpolated data were excluded from the analysis. For further quality control, participants with >50% of their total epochs removed due to excessive interpolation were also excluded from the subsequent pupil analysis [96]. This threshold initially flagged four participants. Upon review, we identified that two of the four participants only marginally exceeded this imposed threshold (50.4% and 52.4%, respectively). Given the large number of epochs still available for these participants and in order to prioritise sample retention, it was decided to keep these participants in the group-level pupil analysis. The two participants removed from the analysis recorded a much larger percentage of epochs removed (61.2% and 75.6%, respectively). Following this procedure, the total proportion of epochs removed from the dataset used in the pupil analysis was 20.8%.

Regression Analyses

All analyses were performed using Python, unless stated otherwise. All regression based models included constants and z-scored predictors and were implemented using ordinary least squares models within the **statsmodels** package.

Prediction of participants' apparent learning rate based on latent uncertainty factors was performed using a model that included change point probability and prior uncertainty as primary predictors and the interaction term between these uncertainty factors as a covariate.
For the epoch-based analysis of the pupil diameter signal, several covariates were included within the regression model. These covariates were used to control for any remaining luminosity discrepancies, linear and non-linear learning effects, fatigue-associated effects, historic biases as well as other (outcome) uncertainty factors. The covariates included the observation (*i.e.* blue vs. yellow), the position of the observation in the current sequence (as well as the square and square root of this position), the task session number, the previous prior belief, and the estimated level of entropy of the prediction. In addition, pupil position was entered as a covariate, computed as the euclidean distance from the fixation point was computed, to account for any positional effects on diameter assessments (*equation 5*).

$$\text{Pupil diameter}_{i,t} = \beta_{0,t} + \beta_{1,t}\text{prior uncertainty}_i + \beta_{2,t}\text{CPP}_i + \beta_{3,t}\text{observation}_i + \beta_{4,t}\text{obs. position}_i + \beta_{5,t}\text{obs. position}^2_i + \beta_{6,t}\sqrt{\text{obs. position}_i} + \beta_{7,t}\text{task session}_i + \beta_{8,t}\text{prior}_i + \beta_{9,t}\text{entropy}_i + \beta_{10,t}\text{pupil position}_t + \varepsilon_{i,t} \quad (5)$$

To control for multiple comparisons, the epoch-based analysis evaluated the statistical significance of each latent uncertainty regression weight using a cluster-based permutation analysis (**MNE** statistics package). A one-sample permutation cluster test was performed on the participant × time data matrix with 1,000 permutations, testing against zero under a two-tailed null hypothesis, with alpha set at 0.05. Of note, the same correction procedure was used for the analysis of differences between uncertainty factors within epochs, as well as learning rate dynamics around change point, testing apparent learning rates against baseline means.

For the model assessing participants' apparent learning rate as a function of identified pupil effects, averaged phasic and tonic pupil responses within time-windows extracted from the epoch analysis were the primary predictors. Covariates included the observation (*i.e.* blue vs. yellow), the position of the observation in the current sequence (as well as the square and square root of this position), the task session number and linear and non-linear learning effects described previously. The moderating effect of the average divergence of the pupil from the fixation point on the pupil diameter response was also entered as a covariate.

The absence of effects between latent uncertainty factors (*e.g.* change point probability) and apparent learning rates did not preclude mediation analysis [97,98], with mediation defined as the presence of an indirect effect. For the mediation analyses, we ran participant-level models to examine whether the effect of latent uncertainty factors (X) on the apparent learning rate (Y) was mediated by associated pupil signals (M). The pupil signals reflected the average signal over a window where the respective uncertainty factor exerted influence (*Fig 5A*). The indirect effect was assessed at the group level using bootstrap procedures. Mediation was implemented using linear regression equations via the **Pingouin** package:

Path a ($X \rightarrow M$):

$$M_i = aX_i + \varepsilon_{Mi} \tag{6}$$

Path b and direct effect, c' ($M, X \rightarrow Y$):

$$Y_i = c'X_i + bM_i + \varepsilon_{Yi} \tag{7}$$

Total effect (Path c):

$$Y_i = cX_i + \varepsilon_{Yi} \tag{8}$$

Indirect effect (Path ab):

$$ab = a \times b$$

(9)

Here, the coefficients represent:

a: effect of prior uncertainty on the tonic pupil response;
b: effect of the tonic pupil response on learning rate, controlling for prior uncertainty;
c′: direct effect of prior uncertainty on learning rate, controlling for the tonic pupil response;
c: total effect of prior uncertainty on learning rate;
ab: indirect effect of prior uncertainty on learning rate through the tonic pupil response;

At the group level, the significance of the pathways were assessed by constructing bootstrapped 95% confidence intervals across 5000 samples.

Funding statement

This study was funded by an ERC Starting grant (#948105-NEURAL-PROB) to F.M.

Acknowledgements

The authors would like to extend their thanks to Hamza Sabek and Thias Marquez for their help in the data collection and piloting the study, respectively.

REFERENCES

1. O'Reilly JX. Making predictions in a changing world—inference, uncertainty, and learning. Front Neurosci. 2013;7. doi:10.3389/fnins.2013.00105

2. Bach DR, Dolan RJ. Knowing how much you don't know: a neural organization of uncertainty estimates. Nat Rev Neurosci. 2012;13: 572–586. doi:10.1038/nrn3289

3. Yu AJ, Dayan P. Uncertainty, neuromodulation, and attention. Neuron. 2005;46: 681–692. doi:10.1016/j.neuron.2005.04.026

4. Soltani A, Izquierdo A. Adaptive learning under expected and unexpected uncertainty. Nat Rev Neurosci. 2019;20: 635–644. doi:10.1038/s41583-019-0180-y

5. Behrens TEJ, Woolrich MW, Walton ME, Rushworth MFS. Learning the value of information in an uncertain world. Nat Neurosci. 2007;10: 1214–1221. doi:10.1038/nn1954

6. Browning M, Behrens TE, Jocham G, O'Reilly JX, Bishop SJ. Anxious individuals have difficulty learning the causal statistics of aversive environments. Nat Neurosci. 2015;18: 590–596. doi:10.1038/nn.3961

7. Nassar MR, Wilson RC, Heasly B, Gold JI. An approximately Bayesian delta-rule model explains the dynamics of belief updating in a changing environment. J Neurosci. 2010;30: 12366–12378. doi:10.1523/jneurosci.0822-10.2010

8. Piray P, Daw ND. A model for learning based on the joint estimation of stochasticity and volatility. Nat Commun. 2021;12: 6587. doi:10.1038/s41467-021-26731-9

9. Piray P, Daw ND. Computational processes of simultaneous learning of stochasticity and volatility in humans. Nat Commun. 2024;15: 9073. doi:10.1038/s41467-024-53459-z

10. Foucault C, Weber LA, Hunt L. Environmental dynamics shape human learning: change points versus random walks. bioRxiv; 2025. p. 2025.11.26.690700. doi:10.1101/2025.11.26.690700

11. Rescorla RA, Wagner AR. A theory of Pavlovian conditioning variations in the effectiveness of reinforcement and nonreinforcement. A. H. Black, W. F. Prokasy (Eds.). Classical conditioning II: current research and theory. A. H. Black, W. F. Prokasy (Eds.). New York Appleton: Century-Crofts; 1972. pp. 64–99. Available: https://www.scirp.org/reference/referencespapers?referenceid=1871303

12. Sutton RS, Barto AG. Introduction to reinforcement learning. MIT Press; 1998.

13. Heilbron M, Meyniel F. Confidence resets reveal hierarchical adaptive learning in humans. PLOS Comput Biol. 2019;15: e1006972. doi:10.1371/journal.pcbi.1006972

14. Foucault C, Meyniel F. Two determinants of dynamic adaptive learning for magnitudes and probabilities. Open Mind Discov Cogn Sci. 2024;8: 615–638. doi:10.1162/opmi_a_00139

15. Jepma M, Murphy PR, Nassar MR, Rangel-Gomez M, Meeter M, Nieuwenhuis S. Catecholaminergic regulation of learning rate in a dynamic environment. O'Reilly JX, editor. PLOS Comput Biol. 2016;12: e1005171. doi:10.1371/journal.pcbi.1005171

16. Jepma M, Brown SBRE, Murphy PR, Koelewijn SC, De Vries B, Van Den Maagdenberg AM, et al. Noradrenergic and cholinergic modulation of belief updating. J Cogn Neurosci. 2018;30: 1803–1820. doi:10.1162/jocn_a_01317

17. McGuire JT, Nassar MR, Gold JI, Kable JW. Functionally dissociable influences on learning rate in a dynamic environment. Neuron. 2014;84: 870–881. doi:10.1016/j.neuron.2014.10.013

18. Nassar MR, Rumsey KM, Wilson RC, Parikh K, Heasly B, Gold JI. Rational regulation of learning dynamics by pupil-linked arousal systems. Nat Neurosci. 2012;15: 1040–1046. doi:10.1038/nn.3130

19. Prat-Carrabin A, Wilson RC, Cohen JD, Azeredo da Silveira R. Human inference in changing environments with temporal structure. Psychol Rev. 2021;128: 879–912. doi:10.1037/rev0000276

20. Vaghi MM, Luyckx F, Sule A, Fineberg NA, Robbins TW, De Martino B. Compulsivity reveals a novel dissociation between action and confidence. Neuron. 2017;96: 348-354.e4. doi:10.1016/j.neuron.2017.09.006

21. Doya K. Metalearning and neuromodulation. Neural Netw. 2002;15: 495–506. doi:10.1016/S0893-6080(02)00044-8

22. Gershman SJ. Dopamine, inference, and uncertainty. Neural Comput. 2017;29: 3311–3326. doi:10.1162/neco_a_01023

23. Nassar MR. Toward a computational role for locus coeruleus/norepinephrine arousal systems. Curr Opin Behav Sci. 2024;59: 101407. doi:10.1016/j.cobeha.2024.101407

24. Joshi S, Gold JI. Pupil size as a window on neural substrates of cognition. Trends Cogn Sci. 2020;24: 466–480. doi:10.1016/j.tics.2020.03.005

25. Joshi S, Li Y, Kalwani RM, Gold JI. Relationships between pupil diameter and neuronal activity in the locus coeruleus, colliculi, and cingulate cortex. Neuron. 2016;89: 221–234. doi:10.1016/j.neuron.2015.11.028

26. Reimer J, McGinley MJ, Liu Y, Rodenkirch C, Wang Q, McCormick DA, et al. Pupil fluctuations track rapid changes in adrenergic and cholinergic activity in cortex. Nat Commun. 2016;7: 13289. doi:10.1038/ncomms13289

27. Lloyd B, de Voogd LD, Mäki-Marttunen V, Nieuwenhuis S. Pupil size reflects activation of subcortical ascending arousal system nuclei during rest. Mather M, Büchel C, Mather M, Jacobs HI, editors. eLife. 2023;12: e84822. doi:10.7554/eLife.84822

28. Raut RV, Rosenthal ZP, Wang X, Miao H, Zhang Z, Lee J-M, et al. Arousal as a universal embedding for spatiotemporal brain dynamics. Nature. 2025; 1–8. doi:10.1038/s41586-025-09544-4

29. Preuschoff K, ’t Hart BM, Einhauser W. Pupil dilation signals surprise: evidence for noradrenaline’s role in decision making. Front Neurosci. 2011;5. doi:10.3389/fnins.2011.00115

30. Urai AE, Braun A, Donner TH. Pupil-linked arousal is driven by decision uncertainty and alters serial choice bias. Nat Commun. 2017;8: 14637. doi:10.1038/ncomms14637

31. O'Reilly JX, Schüffelgen U, Cuell SF, Behrens TEJ, Mars RB, Rushworth MFS. Dissociable effects of surprise and model update in parietal and anterior cingulate cortex. Proc Natl Acad Sci. 2013;110: E3660–E3669. doi:10.1073/pnas.1305373110

32. de Berker AO, Rutledge RB, Mathys C, Marshall L, Cross GF, Dolan RJ, et al. Computations of uncertainty mediate acute stress responses in humans. Nat Commun. 2016;7: 10996. doi:10.1038/ncomms10996

33. Meyniel F. Brain dynamics for confidence-weighted learning. PLOS Comput Biol. 2020;16: e1007935. doi:10.1371/journal.pcbi.1007935

34. Krishnamurthy K, Nassar MR, Sarode S, Gold JI. Arousal-related adjustments of perceptual biases optimize perception in dynamic environments. Nat Hum Behav. 2017;1: 0107. doi:10.1038/s41562-017-0107

35. Slooten JCV, Jahfari S, Knapen T, Theeuwes J. Individual differences in eye blink rate predict both transient and tonic pupil responses during reversal learning. PLOS ONE. 2017;12: e0185665. doi:10.1371/journal.pone.0185665

36. Vincent P, Parr T, Benrimoh D, Friston KJ. With an eye on uncertainty: modelling pupillary responses to environmental volatility. PLOS Comput Biol. 2019;15: e1007126. doi:10.1371/journal.pcbi.1007126

37. Filipowicz AL, Glaze CM, Kable JW, Gold JI. Pupil diameter encodes the idiosyncratic, cognitive complexity of belief updating. Donner TH, Behrens TE, Donner TH, editors. eLife. 2020;9: e57872. doi:10.7554/eLife.57872

38. Gesztesi G, Pajkossy P. Wink or blush? Pupil-linked phasic arousal signals both change and uncertainty during assessment of changing environmental regularities. Cognition. 2025;264: 106256. doi:10.1016/j.cognition.2025.106256

39. Ma WJ, Kording KP, Goldreich D. Bayesian Models of Perception and Action: An Introduction. Cambridge, MA, USA: MIT Press; 2023. Available: https://mitpress.mit.edu/9780262047593/bayesian-models-of-perception-and-action/

40. Mathot S. Pupillometry: psychology, physiology, and function. J Cogn. 2018;1. doi:10.5334/joc.18

41. Meyniel F, Sigman M, Mainen ZF. Confidence as Bayesian probability: from neural origins to behavior. Neuron. 2015;88: 78–92. doi:10.1016/j.neuron.2015.09.039

42. Meyniel F, Schlunegger D, Dehaene S. The sense of confidence during probabilistic learning: a normative account. PLOS Comput Biol. 2015;11: e1004305. doi:10.1371/journal.pcbi.1004305

43. Foucault C, Bounmy T, Demortain S, Thirion B, Eger E, Meyniel F. A non-monotonic code for event probability in the human brain. Nat Commun. 2025;16: 10695. doi:10.1038/s41467-025-65727-7

44. Iglesias S, Mathys C, Brodersen KH, Kasper L, Piccirelli M, den Ouden HEM, et al. Hierarchical prediction errors in midbrain and basal forebrain during sensory learning. Neuron. 2013;80: 519–530. doi:10.1016/j.neuron.2013.09.009

45. Payzan-LeNestour E, Dunne S, Bossaerts P, O'Doherty JP. The neural representation of unexpected uncertainty during value-based decision making. Neuron. 2013;79: 191–201. doi:10.1016/j.neuron.2013.04.037

46. Meyniel F, Dehaene S. Brain networks for confidence weighting and hierarchical inference during probabilistic learning. Proc Natl Acad Sci. 2017;114: E3859–E3868. doi:10.1073/pnas.1615773114

47. Tomov MS, Truong VQ, Hundia RA, Gershman SJ. Dissociable neural correlates of uncertainty underlie different exploration strategies. Nat Commun. 2020;11: 2371. doi:10.1038/s41467-020-15766-z

48. Bounmy T, Eger E, Meyniel F. A characterization of the neural representation of confidence during probabilistic learning. NeuroImage. 2023;268: 119849. doi:10.1016/j.neuroimage.2022.119849

49. Skowron A, Kosciessa JQ, Lorenz RC, Hertwig R, van den Bos W, Garrett DD. Neural variability in the default mode network compresses with increasing belief precision during Bayesian inference. Cereb Cortex. 2025;35: bhaf219. doi:10.1093/cercor/bhaf219

50. Nieuwenhuis S, de Geus EJ, Aston-Jones G. The anatomical and functional relationship between the P3 and autonomic components of the orienting response. Psychophysiology. 2011;48: 162–175. doi:10.1111/j.1469-8986.2010.01057.x

51. Donchin E. Surprise!… Surprise? Psychophysiology. 1981;18: 493–513. doi:10.1111/j.1469-8986.1981.tb01815.x

52. Kolossa A, Kopp B, Fingscheidt T. A computational analysis of the neural bases of Bayesian inference. NeuroImage. 2015;106: 222–237. doi:10.1016/j.neuroimage.2014.11.007

53. Grujic N, Polania R, Burdakov D. Neurobehavioral meaning of pupil size. Neuron. 2024;112: 3381–3395. doi:10.1016/j.neuron.2024.05.029

54. Varazzani C, San-Galli A, Gilardeau S, Bouret S. Noradrenaline and dopamine neurons in the reward/effort trade-off: a direct electrophysiological comparison in behaving monkeys. J Neurosci. 2015;35: 7866–7877. doi:10.1523/JNEUROSCI.0454-15.2015

55. Liu Y, Rodenkirch C, Moskowitz N, Schriver B, Wang Q. Dynamic lateralization of pupil dilation evoked by locus coeruleus activation results from sympathetic, not parasympathetic, contributions. Cell Rep. 2017;20: 3099–3112. doi:10.1016/j.celrep.2017.08.094

56. Murphy PR, O'Connell RG, O'Sullivan M, Robertson IH, Balsters JH. Pupil diameter covaries with BOLD activity in human locus coeruleus. Hum Brain Mapp. 2014;35: 4140–4154. doi:10.1002/hbm.22466

57. Cazettes F, Reato D, Morais JP, Renart A, Mainen ZF. Phasic activation of dorsal raphe serotonergic neurons Increases pupil size. Curr Biol. 2021;31: 192-197.e4. doi:10.1016/j.cub.2020.09.090

58. Grujic N, Tesmer A, Bracey E, Peleg-Raibstein D, Burdakov D. Control and coding of pupil size by hypothalamic orexin neurons. Nat Neurosci. 2023;26: 1160–1164. doi:10.1038/s41593-023-01365-w

59. Megemont M, McBurney-Lin J, Yang H. Pupil diameter is not an accurate real-time readout of locus coeruleus activity. Reimer J, Gold JI, editors. eLife. 2022;11: e70510. doi:10.7554/eLife.70510

60. Bouret S, Sara SJ. Network reset: a simplified overarching theory of locus coeruleus noradrenaline function. Trends Neurosci. 2005;28: 574–582. doi:10.1016/j.tins.2005.09.002

61. Aston-Jones G, Cohen JD. An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. Annu Rev Neurosci. 2005;28: 403–450. doi:10.1146/annurev.neuro.28.061604.135709

62. Jordan R. The locus coeruleus as a global model failure system. Trends Neurosci. 2024;47: 92–105. doi:10.1016/j.tins.2023.11.006

63. Sales AC, Friston KJ, Jones MW, Pickering AE, Moran RJ. Locus coeruleus tracking of prediction errors optimises cognitive flexibility: an active inference model. PLOS Comput Biol. 2019;15: e1006267. doi:10.1371/journal.pcbi.1006267

64. Bouret S, Sara SJ. Reward expectation, orientation of attention and locus coeruleus-medial frontal cortex interplay during learning. Eur J Neurosci. 2004;20: 791–802. doi:10.1111/j.1460-9568.2004.03526.x

65. McBurney-Lin J, Vargova G, Garad M, Zagha E, Yang H. The locus coeruleus mediates behavioral flexibility. Cell Rep. 2022;41. doi:10.1016/j.celrep.2022.111534

66. Xiang L, Harel A, Gao H, Pickering AE, Sara SJ, Wiener SI. Behavioral correlates of activity of optogenetically identified locus coeruleus noradrenergic neurons in rats performing T-maze tasks. Sci Rep. 2019;9: 1361. doi:10.1038/s41598-018-37227-w

67. Lapiz MDS, Morilak DA. Noradrenergic modulation of cognitive function in rat medial prefrontal cortex as measured by attentional set shifting capability. Neuroscience. 2006;137: 1039–1049. doi:10.1016/j.neuroscience.2005.09.031

68. McGaughy J, Ross RS, Eichenbaum H. Noradrenergic, but not cholinergic, deafferentation of prefrontal cortex impairs attentional set-shifting. Neuroscience. 2008;153: 63–71. doi:10.1016/j.neuroscience.2008.01.064

69. Seu E, Lang A, Rivera RJ, Jentsch JD. Inhibition of the norepinephrine transporter improves behavioral flexibility in rats and monkeys. Psychopharmacology (Berl). 2009;202: 505–519. doi:10.1007/s00213-008-1250-4

70. Janitzky K, Lippert MT, Engelhorn A, Tegtmeier J, Goldschmidt J, Heinze H-J, et al. Optogenetic silencing of locus coeruleus activity in mice impairs cognitive flexibility in an attentional set-shifting task. Front Behav Neurosci. 2015;9. doi:10.3389/fnbeh.2015.00286

71. Jahn CI, Gilardeau S, Varazzani C, Blain B, Sallet J, Walton ME, et al. Dual contributions of noradrenaline to behavioural flexibility and motivation. Psychopharmacology (Berl). 2018;235: 2687–2702. doi:10.1007/s00213-018-4963-z

72. Glennon E, Carcea I, Martins ARO, Multani J, Shehu I, Svirsky MA, et al. Locus coeruleus activation accelerates perceptual learning. Brain Res. 2019;1709: 39–49. doi:10.1016/j.brainres.2018.05.048

73. Chen CS, Mueller D, Knep E, Ebitz RB, Grissom NM. Dopamine and norepinephrine differentially mediate the exploration–exploitation tradeoff. J Neurosci. 2024;44. doi:10.1523/JNEUROSCI.1194-23.2024

74. Aston‐Jones G, Cohen JD. Adaptive gain and the role of the locus coeruleus–norepinephrine system in optimal performance. J Comp Neurol. 2005;493: 99–110. doi:10.1002/cne.20723

75. Gilzenrat MS, Nieuwenhuis S, Jepma M, Cohen JD. Pupil diameter tracks changes in control state predicted by the adaptive gain theory of locus coeruleus function. Cogn Affect Behav Neurosci. 2010;10: 252–269. doi:10.3758/CABN.10.2.252

76. Jepma M, Nieuwenhuis S. Pupil diameter predicts changes in the exploration–exploitation trade-off: evidence for the adaptive gain theory. J Cogn Neurosci. 2011;23: 1587–1596. doi:10.1162/jocn.2010.21548

77. Pajkossy P, Szőllősi Á, Demeter G, Racsmány M. Tonic noradrenergic activity modulates explorative behavior and attentional set shifting: evidence from pupillometry and gaze pattern analysis. Psychophysiology. 2017;54: 1839–1854. doi:10.1111/psyp.12964

78. Eldar E, Cohen JD, Niv Y. The effects of neural gain on attention and learning. Nat Neurosci. 2013;16: 1146–1153. doi:10.1038/nn.3428

79. Hoffing RC, Seitz AR. Pupillometry as a glimpse into the neurochemical basis of human memory encoding. J Cogn Neurosci. 2015;27: 765–774. doi:10.1162/jocn_a_00749

80. Takeuchi T, Duszkiewicz AJ, Sonneborn A, Spooner PA, Yamasaki M, Watanabe M, et al. Locus coeruleus and dopaminergic consolidation of everyday memory. Nature. 2016;537: 357–362. doi:10.1038/nature19325

81. Wagatsuma A, Okuyama T, Sun C, Smith LM, Abe K, Tonegawa S. Locus coeruleus input to hippocampal CA3 drives single-trial learning of a novel context. Proc Natl Acad Sci. 2018;115: E310–E316. doi:10.1073/pnas.1714082115

82. Clewett D, Huang R, Davachi L. Locus coeruleus activation “resets” hippocampal event representations and separates adjacent memories. Neuron. 2025;113: 2521-2535.e8. doi:10.1016/j.neuron.2025.05.013

83. Grimm C, Duss SN, Privitera M, Munn BR, Karalis N, Frässle S, et al. Tonic and burst-like locus coeruleus stimulation distinctly shift network activity across the cortical hierarchy. Nat Neurosci. 2024;27: 2167–2177. doi:10.1038/s41593-024-01755-8

84. Chung WY, Meyniel F. Sophistication of human adaptive probability learning in dynamic environments. 2025. Available: https://openreview.net/forum?id=61WWwSW2wr&noteId=4QZ9de5VN6

85. Verbeke P, Verguts T. Humans adaptively select different computational strategies in different learning environments. Psychol Rev. 2025;132: 581–602. doi:10.1037/rev0000474

86. Tavoni G, Doi T, Pizzica C, Balasubramanian V, Gold JI. Human inference reflects a normative balance of complexity and accuracy. Nat Hum Behav. 2022;6: 1153–1168. doi:10.1038/s41562-022-01357-z

87. Fiser J, Berkes P, Orbán G, Lengyel M. Statistically optimal perception and learning: from behavior to neural representations. Trends Cogn Sci. 2010;14: 119–130. doi:10.1016/j.tics.2010.01.003

88. Bowers JS, Davis CJ. Bayesian just-so stories in psychology and neuroscience. Psychol Bull. 2012;138: 389–414. doi:10.1037/a0026450

89. Findling C, Chopin N, Koechlin E. Imprecise neural computations as a source of adaptive behaviour in volatile environments. Nat Hum Behav. 2021;5: 99–112. doi:10.1038/s41562-020-00971-z

90. Foucault C, Meyniel F. Gated recurrence enables simple and accurate sequence prediction in stochastic, changing, and structured environments. Ostojic S, Frank MJ, Ostojic S, Jazayeri M, editors. eLife. 2021;10: e71801. doi:10.7554/eLife.71801

91. Muller TH, Mars RB, Behrens TE, O'Reilly JX. Control of entropy in neural models of environmental state. Donner TH, Frank MJ, Gershman SJ, Meyniel F, editors. eLife. 2019;8: e39404. doi:10.7554/eLife.39404

92. Schaefer M, Mathôt S, Lundqvist M, Lundström JN, Arshamian A. The pupillary respiratory-phase response: pupil size is smallest around inhalation onset and largest during exhalation. J Physiol. 2025;603: 1607–1625. doi:10.1113/JP287205

93. Zénon A. Eye pupil signals information gain. Proc R Soc B Biol Sci. 2019;286: 20191593. doi:10.1098/rspb.2019.1593

94. Mittner M, Groot JM. A novel method for modeling tonic and phasic pupil dynamics in humans. Behav Res Methods. 2025;57: 233. doi:10.3758/s13428-025-02755-7

95. Brisson J, Mainville M, Mailloux D, Beaulieu C, Serres J, Sirois S. Pupil diameter measurement errors as a function of gaze direction in corneal reflection eyetrackers. Behav Res Methods. 2013;45: 1322–1331. doi:10.3758/s13428-013-0327-0

96. Ikwunne B, Parham J, Pulcu E. A nonlinear relationship between prediction errors and learning rates in human reinforcement-learning. PLOS Comput Biol. 2025;21: e1013445. doi:10.1371/journal.pcbi.1013445

97. Zhao X, Lynch JG, Chen Q, article. JD served as editor and GF served as associate editor for this. Reconsidering Baron and Kenny: myths and truths about mediation analysis. J Consum Res. 2010;37: 197–206. doi:10.1086/651257

98. O'Rourke HP, MacKinnon DP. When the test of mediation is more powerful than the test of the total effect. Behav Res Methods. 2015;47: 424–442. doi:10.3758/s13428-014-0481-z

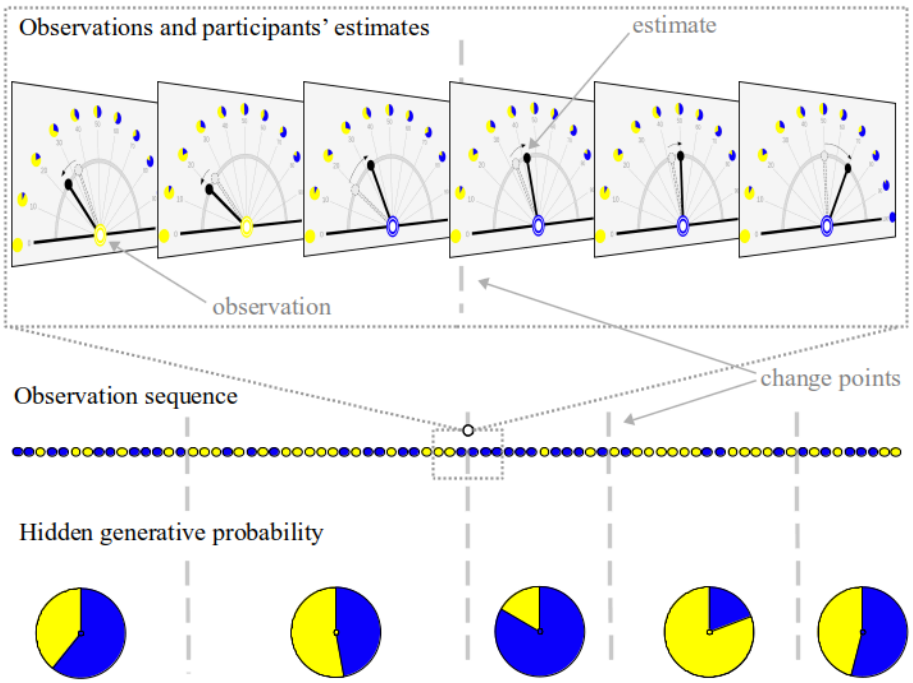

Fig 1. The dynamic probability learning task. Participants were shown a sequence of 75 blue and yellow balls at the fixation dot (center of the screen). They were required to continuously estimate the hidden generative probability of the colour blue (vs. yellow). To make the task more intuitive, participants were told that this probability corresponded to the proportion of blue (vs. yellow) on a fictitious spinning roulette wheel. During the sequence the hidden generative probability could change at undeclared points, referred to as change points (depicted by the dotted lines and changing proportions on the hidden generative probabilities). Change points within a sequence were sampled with constant probability and the stimulus generative probability was sampled within the range of 0.1 to 0.9. Participants reported their estimated probability by moving a marker along a semi-circle. Ticks marked every 10% of the probability with the corresponding roulette wheel depicted. This display maintains a constant distance between the cursor and the fixation dot to prevent saccades and in order to record pupil size throughout. The zoom depicts the behavior of a mock participant across 6 trials containing a change point half through. Participants completed 22 sequences (task sessions). The colours presented in this figure do not accurately reflect the colours presented in the task and are displayed as such simply to differentiate and emphasise the two different outcomes.

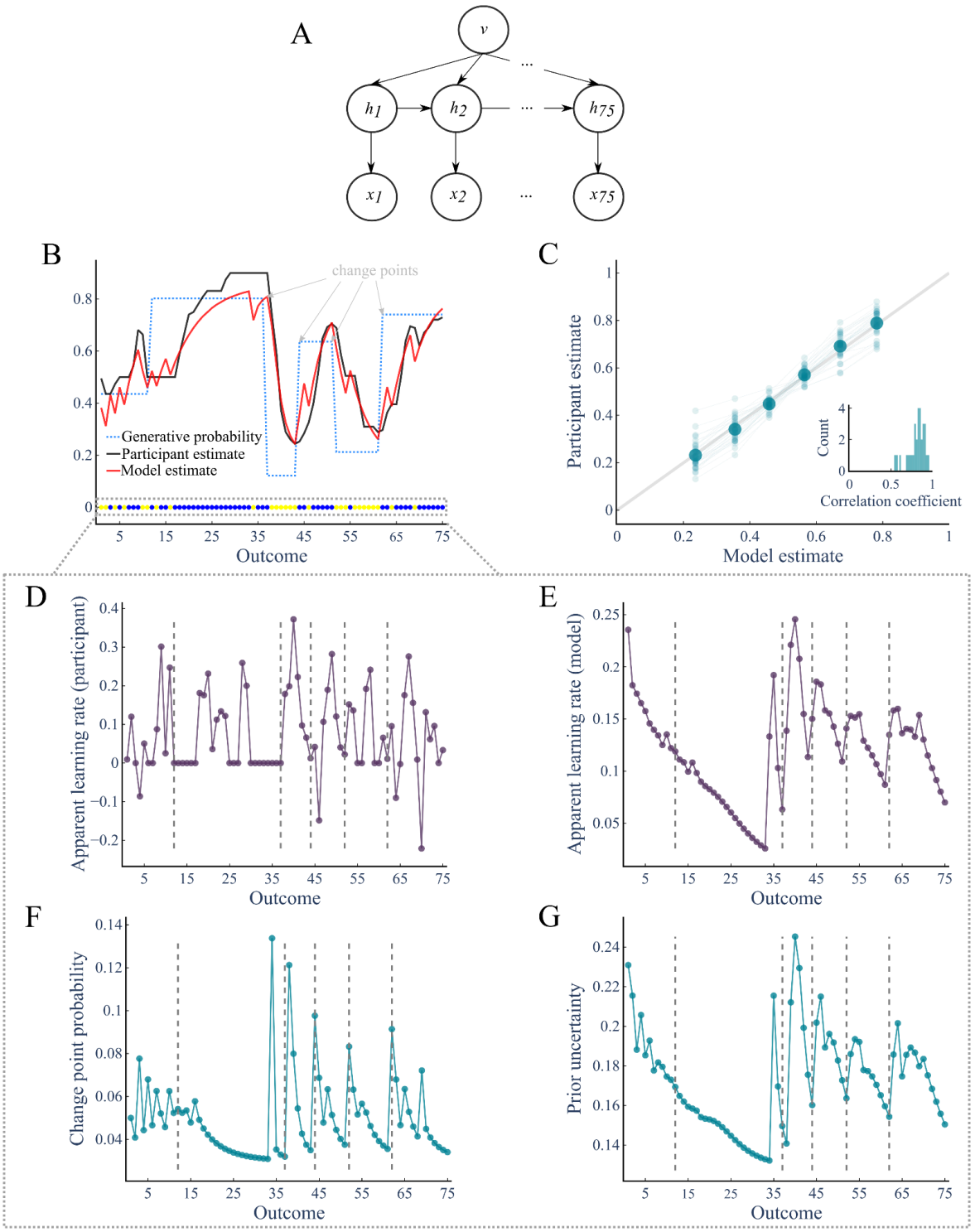

Fig 2. Subjective probability estimates and the ideal observer model. (A) A Bayesian ideal observer model was adopted to provide an optimal solution to the inferential problem. Panel A, describes the dependencies within the model. Based on a Markovian process, the model infers the posterior distribution of the (hidden) generative probability (h_t) given previous observations ($x_{1:t}$), assuming the true level of volatility (v) and that the new generative probability is sampled uniformly from a proposal distribution. **(B)** An example stimulus sequence is represented by the coloured circles at the bottom of the top plot. The dotted blue line represents the (hidden) generative probability across

the sequence, demonstrating a number of change points over the sequence. The black and red lines represent the estimates from the example participant and model, respectively. **(C)** Accuracy of participants' estimates for each observation compared to model estimates (see text). For illustration purposes, the data were binned into 6 equal quantiles of model estimate and averaged within-participants. Small points and lines are the average for each participant, with the larger points reflecting the group mean. The dotted line represents the identity line. Inset displays a histogram of the subject-level coefficients (estimated without binning data). **(D and E)** Apparent learning rates for both participants and the model were derived on each observation from the degree beliefs were updated in proportion to the level of prediction error based on a simple delta rule. **(F and G)** From the model, latent uncertainty-associated factors (*i.e.* prior uncertainty and change point probability) were also extracted for each observation. Thin dotted vertical lines in panels **D-G** represent the first observation after a change point.

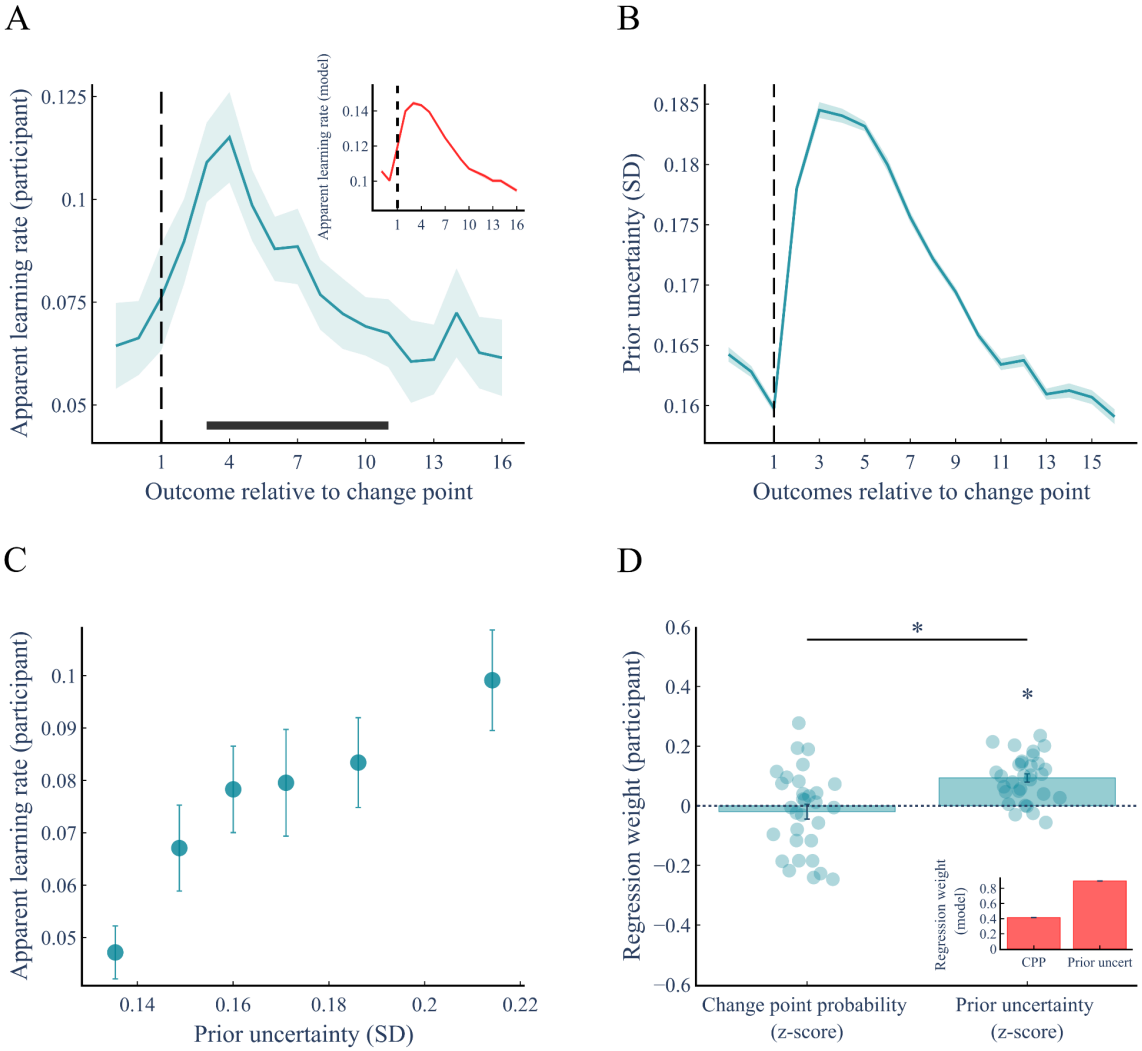

Fig 3. Uncertainty factors guide the dynamic adjustment of participants' apparent learning rate. (A) The dynamics of participants' apparent learning rate were examined around change points. Participants dynamically increase their apparent learning rate after a change point, with learning increased for a protracted period as new stable beliefs are formed. The horizontal line denotes the statistically significant differences compared to the pre-change point learning rate ($p < 0.05$, two-tailed, FWE cluster-corrected for multiple comparisons across time). The average response was similar to normative theory, shown in the inset. **(B)** Prior uncertainty parallels the change in learning rate, dramatically increasing after change points and remaining elevated for multiple trials after as new evidence is accumulated to form stable beliefs. **(C)** Across all trials, participants' apparent learning rate increases as prior uncertainty increases (illustrated here with 6 quantiles of prior uncertainty). **(D)** Weights of prior uncertainty and change point probability effects on participants' learning rates, calculated per participant by multiple linear regression. As predicted by normative theory (inset), prior uncertainty was the dominant driver of participants apparent learning rate (* represents $p < 0.05$). The effect of prior uncertainty was greater than that observed for change point probability, which was not identified as a predictor of participant' apparent learning rate. All plots display the mean ± s.e.m.

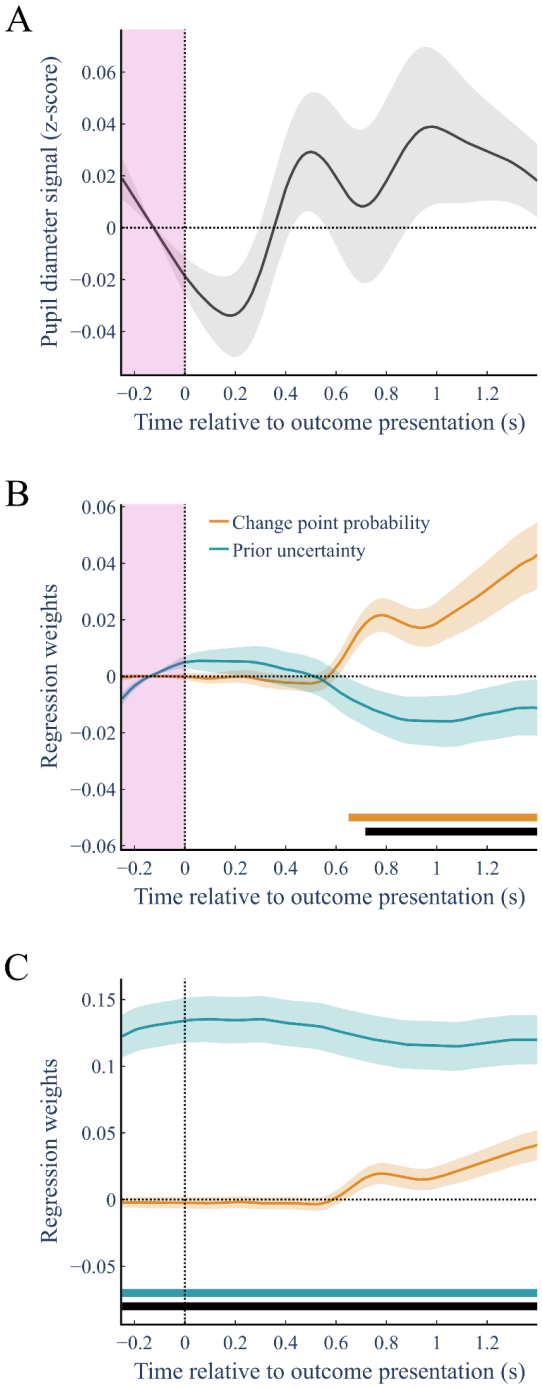

Fig 4. Phasic- and tonic-associated pupil diameter dynamics reflect distinct uncertainty factors. **(A)** Two aspects of pupil responses were evaluated: a phasic response, which was defined by the baseline-corrected (-250 to 0 ms) pupil diameter signal (highlighted on plots), and a tonic response which was defined by the non-corrected, absolute signal. Participants' displayed transient pupil dilations to outcomes (*i.e.* stimuli within the sequences), as revealed by the average, z-scored baseline-corrected pupil signal. **(B and C)** The effect of uncertainty factors on the change in pupil signals were evaluated using multiple linear regression, along with covariates (see *Methods*). This regression model was estimated across trials separately for each peristimulus time point (-250 to 1400 ms) for both the phasic **(B)** and tonic **(C)** pupil diameter responses. We found an exclusive coupling between pupil signals and uncertainty factors effects: phasic-associated responses displayed an effect

on change point probability from 630 ms after observation presentation until the end of the epoch, while tonic-associated responses displayed an effect of prior uncertainty across the entirety of the epoch. Horizontal lines display significant time points for the corresponding (color-coded) uncertainty parameter based on cluster-based permutation analysis ($p < 0.05$, two-tailed, FWE cluster-corrected for multiple comparisons across time). Black horizontal line represents significant time points for a difference between uncertainty parameters based on cluster-based permutation analysis ($p < 0.05$, two-tailed, FWE cluster-corrected for multiple comparisons across time). All plots display the mean ± s.e.m.

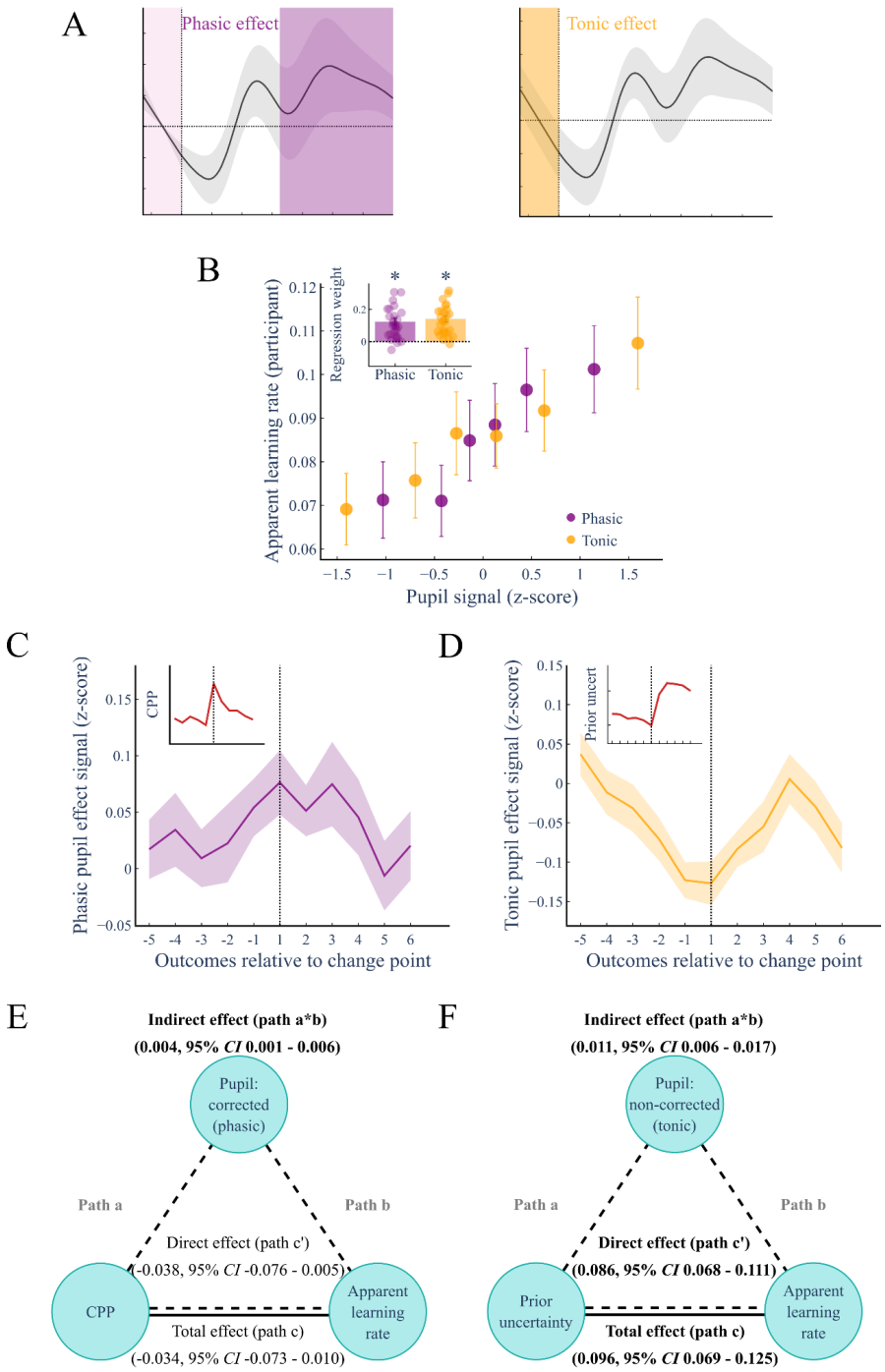

Fig 5. Tonic pupil diameter changes mediate the effect of prior uncertainty on participants' apparent learning rate. (A) The peristimulus regression analysis identified separable effects of change point probability and prior uncertainty within the baseline-corrected and nonbaseline-corrected pupil signals, respectively (see *Fig 4B and 4C*). The traces show the time windows of the observed effects for both signals within which we average the respective pupil signal to obtain one value per trial. **(B)** Tonic and phasic pupil signal effects exert positive influences on participants' apparent learning rate. To illustrate these effects, the data were binned into 6 equal quantiles based on each pupil signal and averaged within-participants. Analysis was performed using multiple linear regression, the group-average and individual regression weights are displayed in the inset (Phasic: $p \approx 1 \times 10^{-5}$; Tonic: $p \approx 1 \times 10^{-7}$). There was no difference between the effect of the two latent uncertainty-associated variables ($p = 0.246$). **(C)** The tonic pupil signal decreases prior to a change

point, before increasing after it; this pattern parallels the prior uncertainty around change points (see inset). **(D)** The phasic pupil signal did not visually display as strong a similarity to change point probability around change points (see inset). **(E)** Given the effect of prior uncertainty on participants' apparent learning rate, a mediation analysis was performed to examine whether the associated tonic pupil diameter response could, at least in part, explain this effect. The indirect path was significant indicating the tonic-associated pupil responses, at least in part, mediated the effect of prior uncertainty on participants' apparent learning rate. **(F)** We also demonstrated that an effect of change point probability on participants' apparent learning rate could be identified when considering the mediating effect of the phasic pupil signal, as indicated by the significant indirect effect. Mediation analyses were performed for each participant and then averaged across-participants, with regression weights and confidence intervals (95% *CI*) based on bootstrapping with 5000 samples. Significance was determined based on the 95% *CI* not spanning across zero, with all instances displayed visually as path texts highlighted in bold. All other traces display the mean ± s.e.m.

SUPPORTING INFORMATION

Fig S1. Absolute error and variability in error across task sessions.

Fig S2. Comparison of participants' apparent learning rates to model-derived estimates.

Fig S3. Outcome of "lag" optimization of the latent uncertainty-associated factors relationship with participants' apparent learning rate.

S1 Appendix. Auditory probability learning task.

Fig S4. Phasic and tonic pupil diameter dynamics reflect distinct uncertainty factors within an auditory probability learning task.

Table S1. Slope of the linear regression between participants' versus model estimates and learning rates

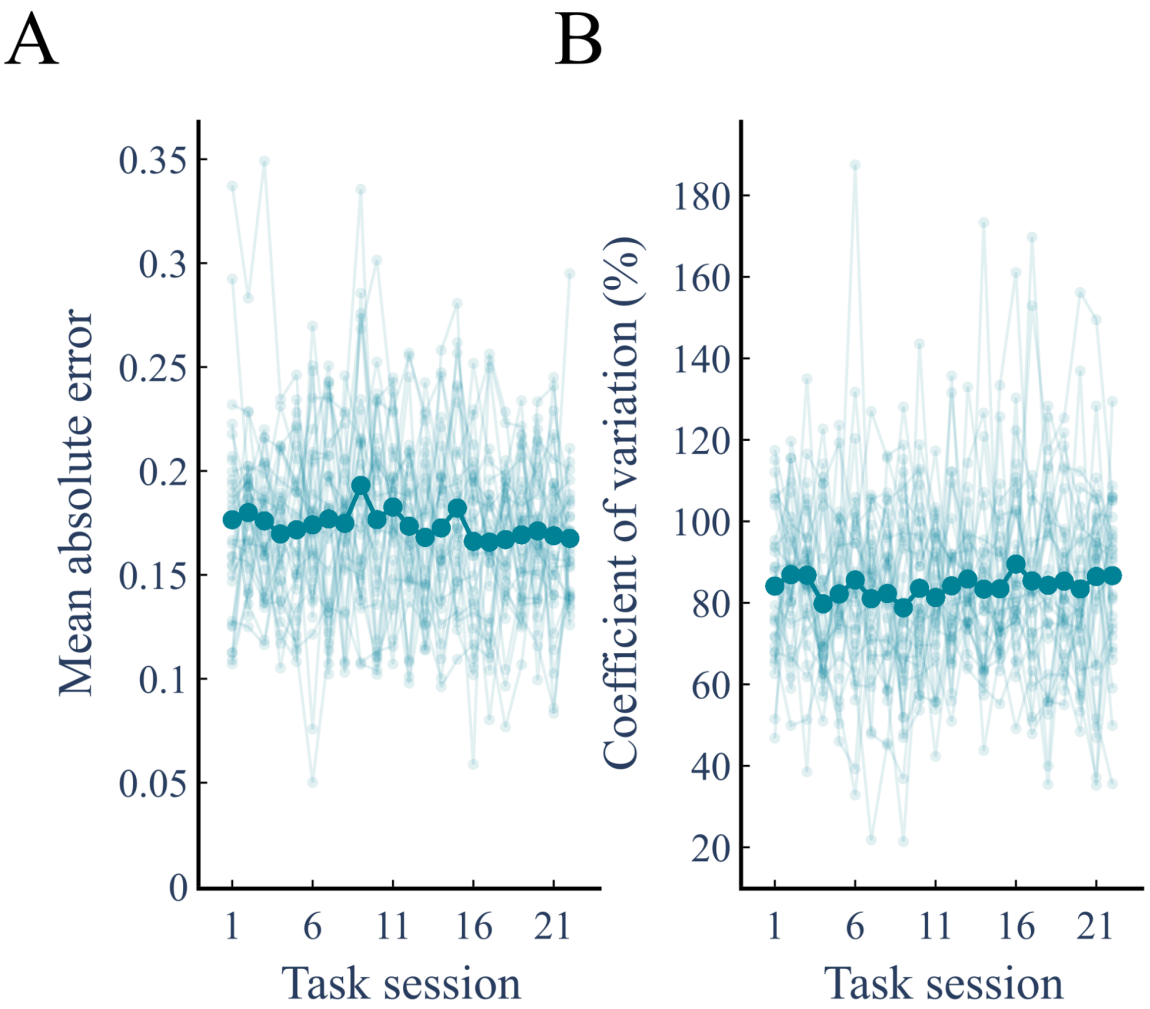

Fig S1. Absolute error and variability in error across task sessions. Participants' performance was stable over the performed task sessions, as quantified through the **(A)** absolute error (F = 0.950, p = 0.520) and **(B)** the coefficient of variation (F = 0.506, p = 0.969). Absolute error was defined by the mean absolute difference between the participant's estimate and the true value of the hidden generative probability. The variability of this error within each task session was quantified by the coefficient of variation, reflecting the standard deviation in the absolute error for a given task session divided by the mean of that error. Thin dots and lines connecting them each denote one participant, with the larger circles representing the mean across participants.

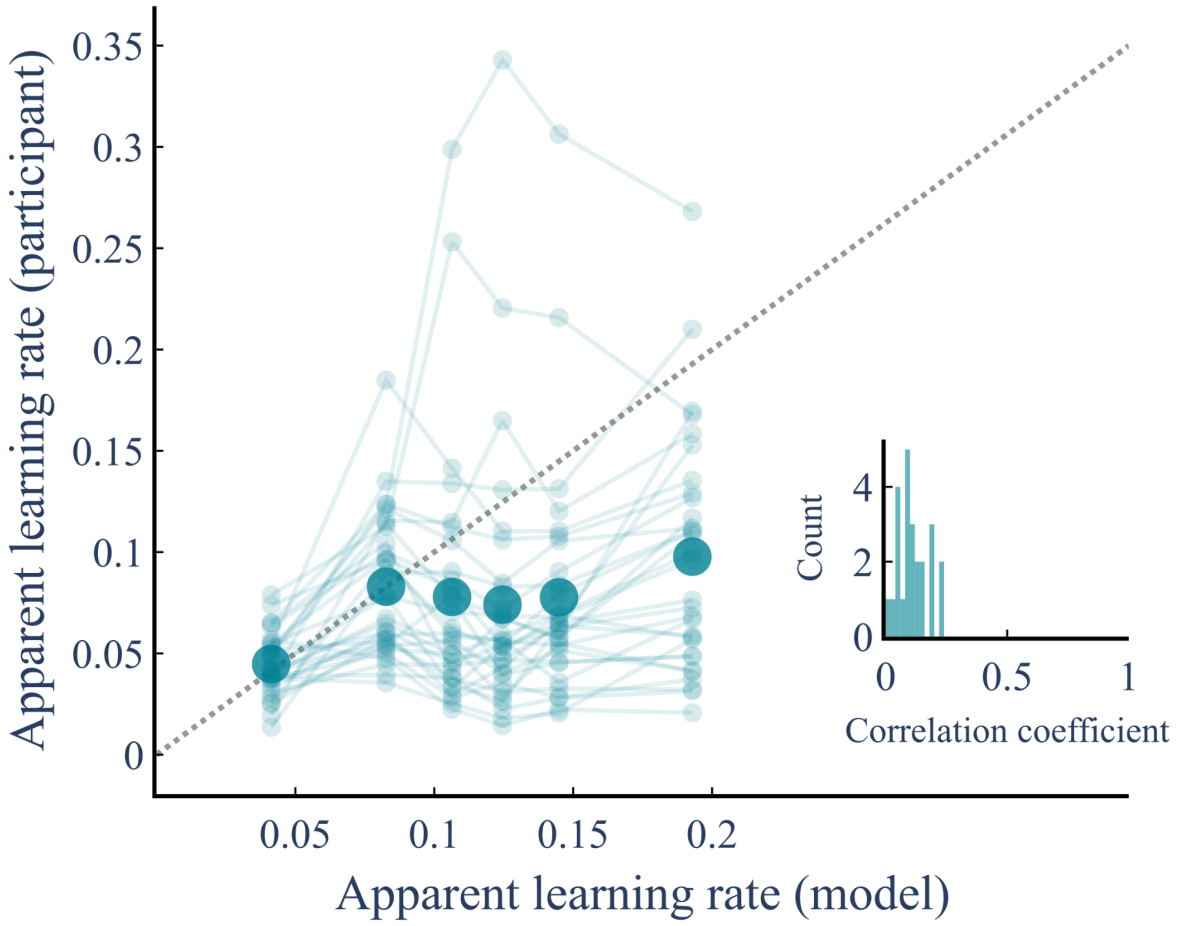

Fig S2. Comparison of participants' apparent learning rates to model-derived estimates. The data were binned into 6 equal quantiles of model learning rate estimates and averaged within-participants. Small points and lines are the average for each participant, with the larger points reflecting the mean. The dotted line represents the identity line. The insert displays a histogram of the observed correlation coefficients across all participants.

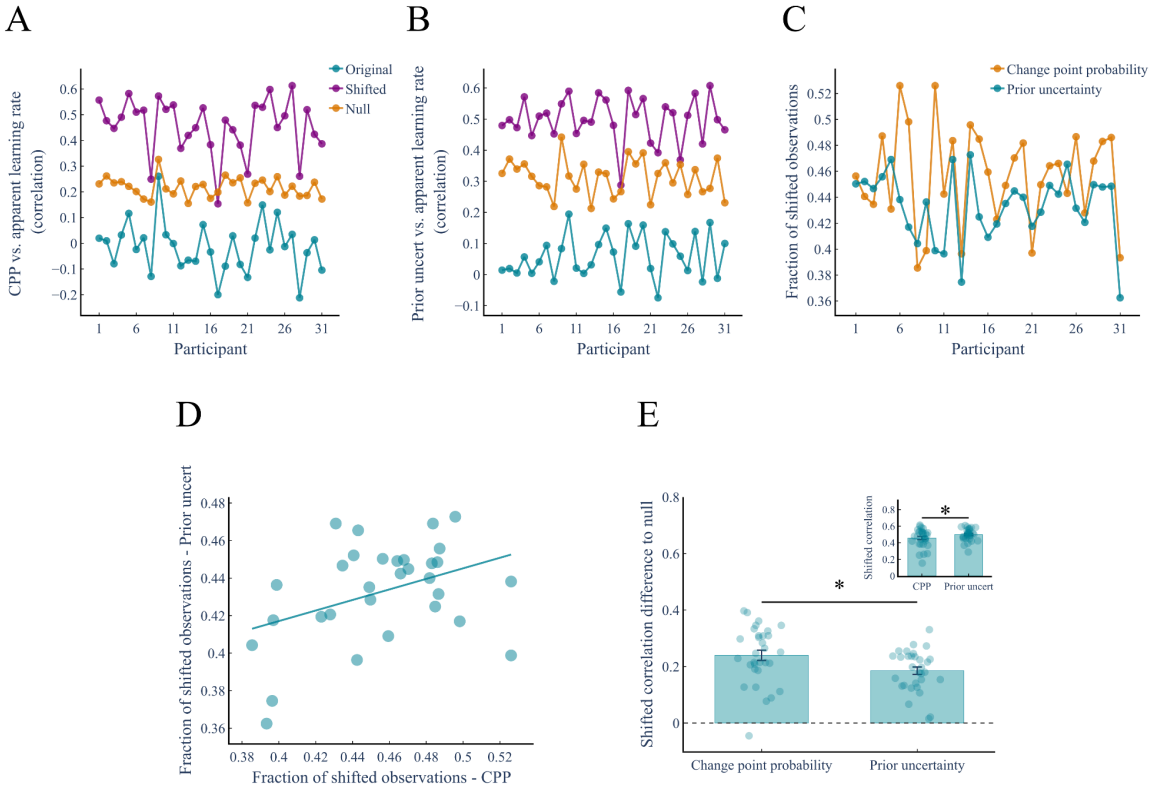

Fig S3. Outcome of "lag" optimization of the latent uncertainty-associated factors relationship with participants' apparent learning rate. **(A and B)** The optimisation procedure improved the correlation compared to the original correlation and that which may be expected by chance for both change point probability (CPP) and prior uncertainty (mean ± s.e.m; CPP - original correlation: -0.02 ± 0.02, shifted correlation: 0.46 ± 0.02, null correlation: 0.22 ± 0.07, shift *vs*. null: $t_{(30)} = 13.19$, $p < 1 \times 10^{-308}$; Prior uncertainty - original correlation: 0.06 ± 0.01, shifted correlation: 0.50 ± 0.01, null correlation: 0.32 ± 0.01, shift *vs*. null: $t_{(30)} = 14.05$, $p < 1 \times 10^{-308}$). **(C)** Between participants, there was a positive correlation between the fraction of trials shifted when optimising the correlation for change point probability and prior uncertainty (Pearson's correlation = 0.394, $p = 0.028$). **(D)** For each participant, there was a high degree of agreement in the fraction of observations shifted to optimize the correlations. **(E)** There was a slightly larger (and statistically significant) effect seen for prior uncertainty compared to change point probability when comparing the absolute correlation coefficients after the optimisation procedure ($t_{(30)} = -2.54$, $p = 0.017$), in keeping with normative theory (insert). However, when the correlation coefficients were expressed relative to the generated null distribution, there was a significant difference between uncertainty factors, with a larger effect seen for change point probability compared to prior uncertainty (mean ± s.e.m; CPP - 0.24 ± 0.02, Prior uncertainty - 0.19 ± 0.01; $t_{(30)} = 3.17$, $p = 0.003$).

S1 Appendix. Auditory probability learning task

To evaluate whether movement artefacts contributed to the observed uncertainty-associated pupil dynamics, we checked if the same effects were present through reanalysis of data previously collected in our lab that used an auditory probability learning task very similar to the main task. In brief, the auditory probability learning task consisted of auditory sequences of 420 stimuli, with stimuli composed of 2 clearly distinguishable tones. The separable tones refer to sets of chords (350, 700 and 1400 Hz vs. 500, 1000 and 2000 Hz) played into both ears for 50 ms (including 7 ms rise and fall times). Auditory stimuli were presented with an inter-stimuli interval of ~1.3 s. On each trial, the probability of a tone was based on a changing Bernoull process. The probability of a change in the underlying generative probability was constant and slightly smaller than the main task (1/75). When a change occurred the generative probability was sampled from the same interval as the main task (0.1-0.9), with similar constraints employed as the main task (*i.e.* the change in odds ratio should be at least 4-fold). This meant that the task did not have overly long stable sequences and that changes were not so small as to be imperceivable. Importantly, movements during the auditory probability learning task were minimal. Every 22 tones (± 3 tones), participants were asked to rate the generative probability as well as their confidence in their estimate using a mouse-tracker. This was the only movement involved in the task, therefore reducing motor involvement considerably compared to the main task. Participants were presented with at least 4 sequences each, resulting in a total of 1680 observations. Here, we only analysed participants with pupil data which exceeded a quality control threshold related to the amount of missing data within the pupil recording. In keeping with the main task, epoch windows were created around each tone presentation (-250 to 1200 ms). Blinks were accounted for within an epoch through linear interpolation. If the amount of interpolated data exceeded 20% of all data points within that epoch, the epoch was removed from the analysis. If a participant had more than 50% of their total epochs removed they were removed from the analysis (see *Methods*). In total, 15 participants were included in the analysis. The effects of latent uncertainty-factors (*i.e.* prior uncertainty and change point probability) were assessed using a multiple linear regression across each peri-stimulus time-point, in line with the cluster-based permutation analysis performed in the main study. Of note, the two tasks differed in the eyetracker used to record pupil size (in the auditory probability learning task, EyeLink 1000), with the sampling rate of the pupil signal higher in the auditory probability task (500 Hz) compared to the main study (60 Hz). Results of the analysis are presented in *Fig S4*.

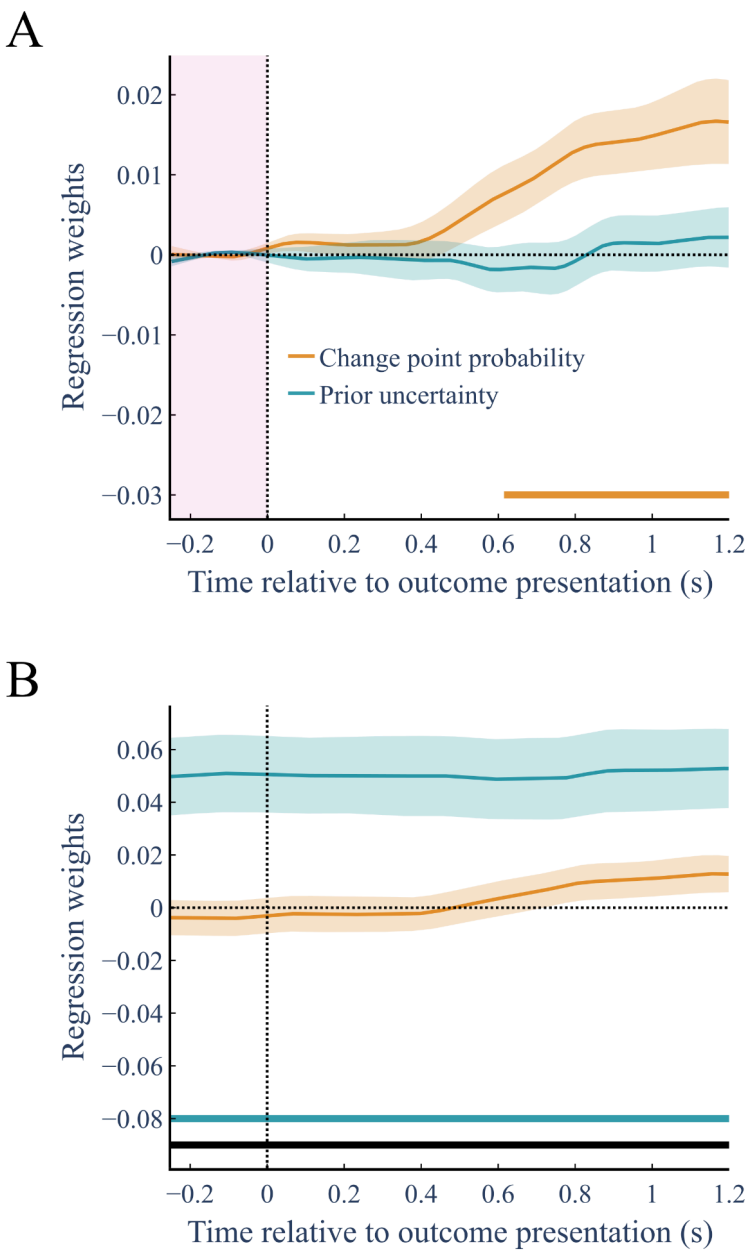

Fig S4. Phasic and tonic pupil diameter dynamics reflect distinct uncertainty factors within an auditory probability learning task. **(A and B)** Peristimulus regression analysis was performed on both phasic (**A**) and tonic **(B)** pupil diameter responses, as defined in the main text. The baseline for the phasic response (-250 to 0 ms) is represented on the plot (**A**). We demonstrate that the coupling between pupil signals and uncertainty factors effects holds in an auditory probability learning task of similar structure. Horizontal lines display significant time points for the corresponding (color-coded) uncertainty parameter based on cluster-based permutation analysis ($p < 0.05$, two-tailed, FWE cluster-corrected for multiple comparisons across time). Black horizontal line represents significant time points for a difference between uncertainty parameters based on cluster-based permutation analysis ($p < 0.05$, two-tailed, FWE cluster-corrected for multiple comparisons across time). All plots display the mean ± s.e.m.

Table S1. Slope of the linear regression between participants' versus model estimates and learning rates

Parameter	Mean	s.e.m	Standard deviation	T-test against 0		T-test against 1	
				t	*p*	*t*	*p*
Probability Estimate	1.04	0.02	0.12	25.96	<0.0001	1.09	0.285
Learning rate	0.31	0.004	0.03	4.48	0.0001	N/A	N/A