

# Rational regulation of learning dynamics by pupil-linked arousal systems

Matthew R Nassar<sup>1</sup>, Katherine M Rumsey<sup>1</sup>, Robert C Wilson<sup>2</sup>, Kinjan Parikh<sup>1</sup>, Benjamin Heasley<sup>1</sup> & Joshua I Gold<sup>1</sup>

The ability to make inferences about the current state of a dynamic process requires ongoing assessments of the stability and reliability of data generated by that process. We found that these assessments, as defined by a normative model, were reflected in nonluminance-mediated changes in pupil diameter of human subjects performing a predictive-inference task. Brief changes in pupil diameter reflected assessed instabilities in a process that generated noisy data. Baseline pupil diameter reflected the reliability with which recent data indicate the current state of the data-generating process and individual differences in expectations about the rate of instabilities. Together these pupil metrics predicted the influence of new data on subsequent inferences. Moreover, a task- and luminance-independent manipulation of pupil diameter predictably altered the influence of new data. Thus, pupil-linked arousal systems can help to regulate the influence of incoming data on existing beliefs in a dynamic environment.

Many decisions, from foraging to financial, depend on the ability to infer a state of the world from both historical and newly arriving information. Such inferences are particularly challenging when they must account for multiple sources of uncertainty. When uncertainty results from noise, reflecting random fluctuations in the information generated by an otherwise stable state, the average overall historical information is most predictive of future observations. In contrast, when uncertainty results from a change in the state itself, only the most recent information pertains to the new state. Thus, historical information should be discounted and beliefs should be updated rapidly to maximize their predictive power. Under certain conditions, human subjects appear to encode and respond appropriately to these different forms of uncertainty when making inferences in a dynamic environment<sup>1–3</sup>. We examined whether this ability is governed, at least in part, by arousal systems that affect pupil diameter, which are thought to include the noradrenergic brainstem nucleus locus coeruleus<sup>4–7</sup>.

Nonluminance-mediated changes in pupil diameter have long been used as indicators of clinical, cognitive and arousal states<sup>8–11</sup>. One interpretation of these pupil changes is that they reflect the amount of cognitive effort exerted at a given time, which can be related to task uncertainty<sup>11</sup>. Accordingly, changes in pupil diameter can be elicited via manipulations of the uncertainty associated with possible actions in certain choice tasks<sup>6,12</sup>. Changes in pupil diameter can also reflect perceived changes in the world, including perceptual switches during perceptual rivalry, detection of targets in oddball or near-threshold tasks, responses to low-probability go signals in a go/no-go task and perceived changes in task utility that can affect task engagement<sup>7,12–15</sup>.

These kinds of uncertainty- and change-related signals are thought to contribute to rational inference in a dynamic environment, including

helping to regulate the relative influence of historical and newly arriving information on existing beliefs<sup>2,3</sup>. Such regulation is an important feature of cognitive flexibility and can be equivalent to adjusting the learning rate in a reinforcement-learning framework<sup>1,16</sup>. We sought to determine how such learning-rate adjustments relate to pupil-linked arousal systems. We found that the arousal system, and possibly the locus coeruleus, can have important and computationally complex roles in rationally regulating the influence of incoming information on beliefs about a dynamic world.

## RESULTS

### Behavior

We measured pupil diameter in 30 human subjects while they performed an isoluminant version of a predictive-inference task<sup>2</sup>. The predictive-inference task required subjects to minimize errors in predicting the next number (outcome) in a series. The outcomes were picked from a Gaussian distribution with a mean that changed at random intervals (change points) and a s.d. (set to either 5 or 10) that was stable over each block of 200 trials (**Fig. 1**). After each prediction was recorded, the new outcome was shown using an isoluminant display for 2 s, during which the subject maintained fixation and pupil diameter was measured (**Fig. 1**). After this interval, the outcome disappeared and the previous prediction reappeared, to be updated for the subsequent trial. Payment scaled inversely with the subject's mean absolute error during the session<sup>2</sup>.

We quantified the extent to which each new outcome influenced the subsequent prediction as the learning rate in a simple delta-rule model (equation (3))<sup>2</sup>. The learning rate is equal to the magnitude of change in the prediction expressed as a fraction of the error made on the previous prediction. Thus, a learning rate of 1 indicates abandonment of the previous prediction in favor of the most recent outcome.

<sup>1</sup>Department of Neuroscience, University of Pennsylvania, Philadelphia, Pennsylvania, USA. <sup>2</sup>Department of Psychology, Princeton University, Princeton, New Jersey, USA. Correspondence should be addressed to J.I.G. (jgold@mail.med.upenn.edu).

Received 8 February; accepted 8 May; published online 3 June 2012; doi:10.1038/nn.3130

**Figure 1** Predictive-inference task sequence and pupillometry. Learning rate was computed by dividing the difference in the prediction from one trial to the next by the difference between the current outcome and the current prediction. Inset, mean  $\pm$  s.e.m. pupil diameter, averaged across z scores computed per subject, aligned to outcome presentation (time = 0). Pupil average was computed for each trial as the mean pupil diameter, z-scored by subject, across the entire 2-s fixation window (vertical dashed lines). Pupil change was computed for each trial as the difference in mean diameter, z-scored by subject, measured late (time = 1–2 s) versus early (time = 0–1 s) during fixation.

A learning rate of zero indicates maintenance of the previous prediction despite a nonzero prediction error.

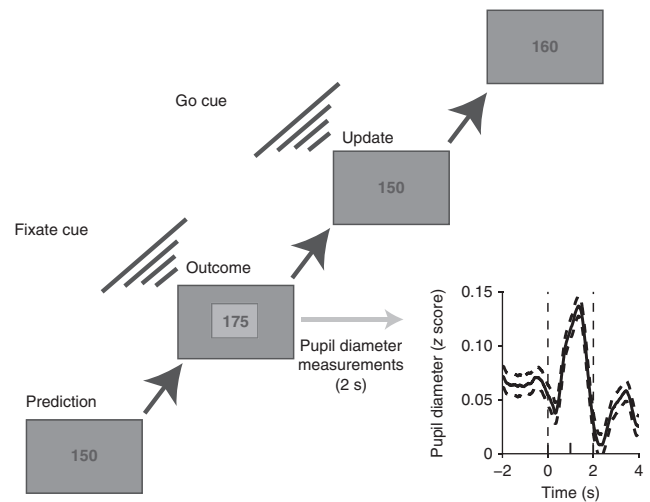
Subjects tended to use variable learning rates that spanned the entire allowed range, from zero to 1. In this range, learning rates tended to be higher for larger errors, scaled by the noise of the generative distribution (**Fig. 2a**). Learning rates also tended to be highest on the trial after a change point and then decayed for several trials thereafter (**Fig. 2b**). These basic trends were similar across subjects, although individual subjects used markedly different distributions of learning rates (**Fig. 2c**).

### Reduced Bayesian model

The learning rates used by subjects were consistent with both a full and a simplified version of the optimal (Bayesian) model<sup>2,17–19</sup>. One advantage of the reduced Bayesian model is that it updates beliefs according to a delta rule in which the learning rate depends on only two parameters computed per trial: change-point probability and relative uncertainty (**Fig. 3a**).

Change-point probability approximates the posterior probability that the mean of the generative distribution has changed since the previous trial, given all previous data. If the mean does change, then previous outcomes should be unrelated to future ones and should not contribute to an updated prediction. Accordingly, the model uses learning rates that scale linearly toward 1 (thus, discarding historical information) as the change-point probability approaches 1 (**Fig. 3a**). Change-point probability is computed by comparing the probability of each new outcome given either the current predictive distribution or the occurrence of a change point (equation (5)). Its value increases monotonically as a function of the absolute difference between predicted and actual outcome, scaled according to the s.d. of the generative distribution (equation (6); **Fig. 3b**).

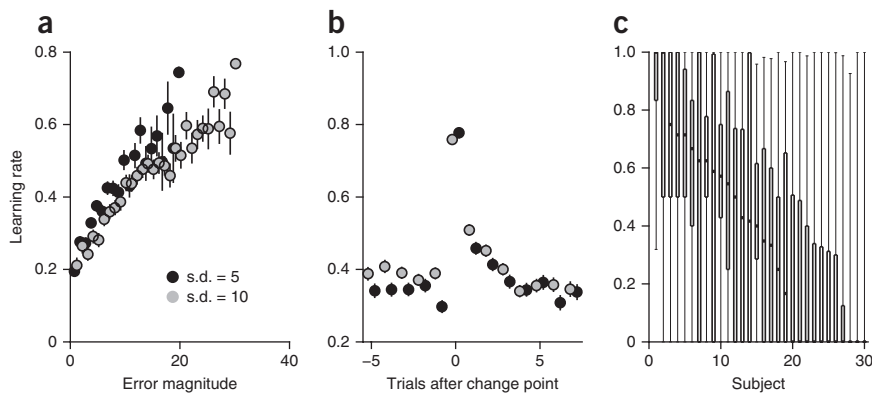
Relative uncertainty is a function of total uncertainty, which in our task arises from two sources. The first source, noise, reflects the unreliability with which a single sample can be predicted from a distribution with a known mean. The second source reflects the unreliability of the current estimate of the mean, which decreases as more data are observed from a distribution. Relative uncertainty is the magnitude of



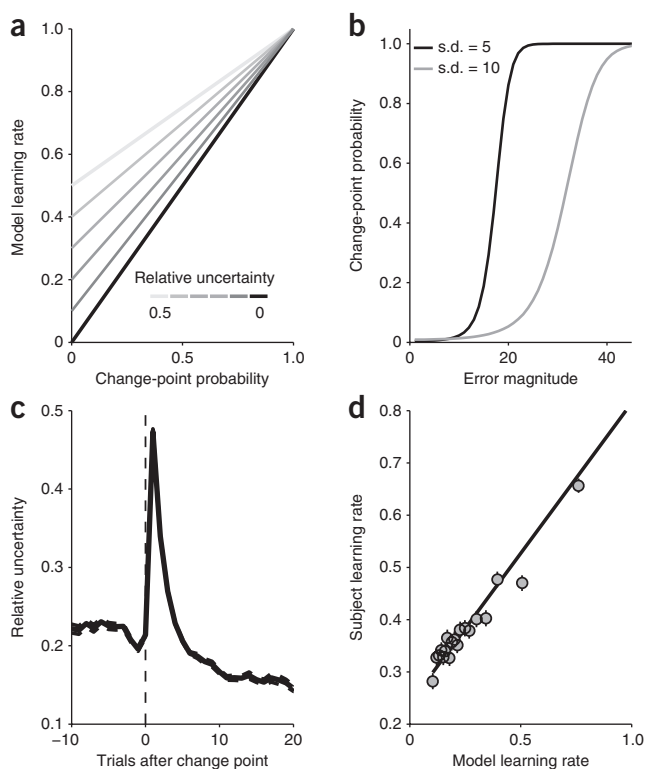
this second form of uncertainty as a fraction of total uncertainty, analogous to the gain in a Kalman filter. Relative uncertainty determines the learning rate when change-point probability is zero and sets the y intercept of the relationship between change-point probability and learning rate otherwise (**Fig. 3a**). The effects of relative uncertainty on model learning rates are greatest on the trials following a change point, when its value peaks at 0.5 and then decays over several trials (equation (7); **Fig. 3c**).

Like the human subjects, the model tended to compute learning rates that were highest just following a change point in the mean of the generative distribution and then decayed for several trials independently of noise. When applied to the exact same outcome sequences as the subjects, the model also tended to produce similar learning rates (**Fig. 3d**).

We related change-point probability and relative uncertainty computed in the model to the mean pupil diameter (pupil average) and change in pupil diameter (pupil change) measured during the 2-s outcome-viewing period (**Fig. 1**), using two linear regression models. The first, simpler model had four parameters: change-point probability and relative uncertainty computed from the reduced Bayesian model, the s.d. of generative distribution and a binary variable describing whether or not the prediction error was exactly zero. The second model included all of these parameters, as well as several potential confounding factors, such as eye position and velocity (see Online Methods). The models are complementary: the first avoids potential interactions between large numbers of parameters and therefore has coefficients that are more readily interpretable, whereas the second avoids missing out on the many factors that, in principle, could affect our pupil measurements. Both models captured a significant amount of variability in the pupil data (for pupil average/pupil change data, an *F* test rejected the null model relative to the small model for



**Figure 2** Task performance. (a) Learning rates were highest after subjects made larger errors, scaled by noise (as indicated). Points and error bars represent mean  $\pm$  s.e.m. from all subjects. (b) Learning rates were highest on change-point trials and decayed thereafter, similarly for both noise conditions. Points and error bars represent mean  $\pm$  s.e.m. from all subjects. (c) Learning-rate distributions across all trials from each of the 30 subjects (abscissa), sorted by median learning rate. Horizontal line, box and whiskers indicate median, 25<sup>th</sup>/75<sup>th</sup> percentiles and 5<sup>th</sup>/95<sup>th</sup> percentiles, respectively.



**Figure 3** Reduced Bayesian model. **(a)** Learning rate as a function of change-point probability (abscissa) and relative uncertainty (line shading), as computed by the model. **(b)** Change-point probability computed by the model as a function of error magnitude (abscissa) for the two different noise conditions, as indicated, computed for a given relative uncertainty (equal to 0.02 for this figure). **(c)** Mean  $\pm$  s.e.m. relative uncertainty computed by the model aligned to change points (vertical dashed line at zero) from all sequences experienced by the subjects for the two different noise conditions. **(d)** Trial-by-trial comparison of subject and model learning rates were computed using the same sequence of outcomes experienced by each subject. Points and error bars represent mean  $\pm$  s.e.m. data from all subjects grouped into 20 five-percentile bins according to the corresponding model learning rate. The solid line is a linear fit to the unbinned data ( $r = 0.33$ ,  $P < 0.001$ ).

One notable exception to the positive relationship between pupil change and error magnitude occurred for trials in which the error was exactly zero, which corresponded to relatively large pupil changes (Fig. 4a). Accordingly, a binary variable added to the linear model that described whether or not the subject correctly predicted the outcome was related to pupil change (the mean value of the regression coefficient was 0.180  $z$ -scored pupil change ( $z_{PC}$ ) for the four-parameter regression model and 0.156  $z_{PC}$  for the larger model;  $P < 0.05$  for  $H_0$ ; mean = 0 for each model), but not pupil average (mean regression coefficient =  $-0.076$  and  $-0.092$   $z$ -scored pupil average ( $z_{PA}$ ) for the smaller and larger regression models, respectively;  $P > 0.05$ ). Thus, pupil change reflected not only change-point probability, but also whether or not the subject correctly predicted the observed outcome.

#### Average pupil diameter reflected belief uncertainty

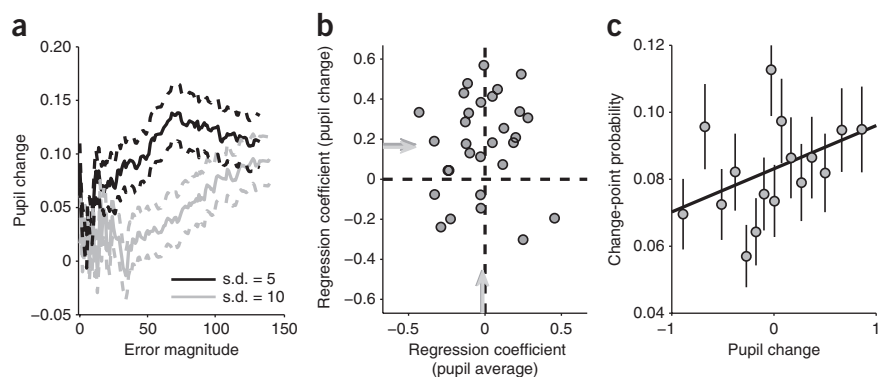
The average pupil diameter during the outcome-viewing period, as with the relative uncertainty (RU) in our model, tended to peak on the trial after a change point and then diminish in magnitude as more relevant information reinforced the existing belief (Figs. 2b, 3c and 5a). Accordingly, when computed by the model using the same sequence of outcomes experienced by each subject, relative uncertainty tended to be positively predictive of pupil average (Fig. 5b). This result did not simply reflect differences in motor output following change points (for example, longer button presses to choose a learning rate near 1), as similar results were obtained in a control experiment in which subject predictions were reset using a learning rate of 0.5 on each trial, thus requiring the same motor act to choose a learning rate of either zero or 1 (mean regression coefficient = 0.30 and 0.35  $z_{PA}/RU$  for the smaller and larger regression models, respectively;  $P < 0.05$ ).

27/15 of the 30 subjects, and a nested  $F$  test rejected the small model relative to the large model for 29/19 of the 30 subjects;  $P < 0.05$ ).

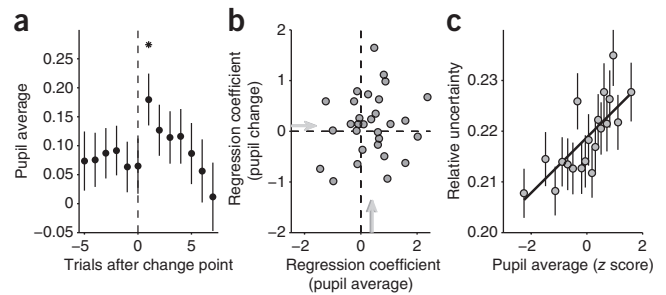
#### Pupil change reflected change-point probability

The change in pupil diameter during the outcome-viewing period, similar to the change-point probability in our model, tended to increase as a function of error magnitude, scaled as a function of noise (Figs. 3b and 4a). Accordingly, when computed by the model using the same sequence of outcomes experienced by each subject, change-point probability tended to be positively predictive of  $z$ -scored pupil change (Fig. 4b). The complement was also true: change-point probability varied systematically as a function of pupil change for data pooled across the population (Fig. 4c). In contrast, there was no consistent relationship between change-point probability and pupil average (Fig. 4b).

**Figure 4** Relationship between pupil change and change-point probability. **(a)** Mean  $\pm$  s.e.m. pupil change from all trials and all subjects for running bins of 150 trials, binned according to the absolute prediction error and sorted by noise, as indicated. **(b)** Regression coefficients describing the linear relationship between change-point probability ( $p_{CH}$ ) and  $z$ -scored pupil change ( $z_{PC}$ , ordinate) versus the regression coefficients describing the linear relationship between  $p_{CH}$  and  $z$ -scored pupil average ( $z_{PA}$ , abscissa). Points are regression coefficients computed for each subject individually, using the four-parameter regression model. Arrows indicate mean values from this model (dark, equal to 0.174  $z_{PC}/p_{CH}$ ,  $t$  test for  $H_0$ : mean = 0,  $P < 0.001$  for the ordinate,  $-0.022$   $z_{PA}/p_{CH}$ ,  $P = 0.58$  for the abscissa) or from the full model (light, equal to 0.148  $z_{PC}/p_{CH}$ ,  $P < 0.001$  for the ordinate,  $-0.014$   $z_{PA}/p_{CH}$ ,  $P = 0.70$  for the abscissa). Dark arrows are partially occluded by light ones. **(c)** Change-point probability from the reduced Bayesian model versus pupil change. Points and error bars represent mean  $\pm$  s.e.m. data from all subjects grouped into 20 five-percentile bins. The solid line is a linear fit to the unbinned data (slope = 0.012  $p_{CH}/z_{PC}$ ,  $P < 0.001$  for  $H_0$ : slope = 0).



**Figure 5** Relationship between pupil diameter and relative uncertainty. (a) Mean  $\pm$  s.e.m. pupil average from all subjects as a function of trials relative to task change points. Asterisk indicates trials differing significantly from all other trials (permutation test for  $H_0$ , equal means after correction for multiple comparisons,  $P < 0.05$ ). (b) Regression coefficients describing the relationship between relative uncertainty and z-scored pupil change ( $z_{PC}$ , ordinate) versus the regression coefficients describing the relationship between relative uncertainty and z-scored pupil average ( $z_{PA}$ , abscissa). Points are regression coefficients computed for each subject individually, using the four-parameter regression model. Arrows indicate mean values from this model (dark, equal to  $0.135 z_{PC}/RU$ ,  $t$  test for  $H_0$ : mean = 0,  $P = 0.28$  for the ordinate,  $0.35 z_{PA}/RU$ ,  $P < 0.05$  for the abscissa) or from the full model (light, equal to  $0.127 z_{PC}/RU$ ,  $P = 0.24$  for the ordinate,  $0.40 z_{PA}/RU$ ,  $P < 0.01$  for the abscissa). Dark arrows are partially occluded by light ones. (c) Relative uncertainty from the reduced Bayesian model versus pupil average. Points and error bars represent mean  $\pm$  s.e.m. data from all subjects grouped into 20 five-percentile bins. The solid line is a linear regression to unbinned data (slope =  $0.0055 RU/z_{PA}$ ,  $P < 0.001$  for  $H_0$ , slope = 0).



The complement was also true: relative uncertainty varied systematically as a function of pupil average for data pooled across the population (Fig. 5c). In contrast, there was no consistent relationship between relative uncertainty and pupil change (Fig. 5b).

Overall uncertainty in our task depends on not only relative uncertainty, but also on noise, which we manipulated by varying the s.d. of the generative distribution in blocks (s.d. = 5 or 10). Consistent with our model, in which noise is only used to compute change-point probability (equations (5) and (6)), these manipulations of noise were reflected in pupil change, but only insofar as pupil change represented change-point probability (Fig. 4a). These manipulations of noise did not have any other systematic effects on either pupil change or pupil average ( $P > 0.1$  for  $H_0$ , a mean value of zero for the regression coefficient describing the influence of noise on the given pupil measurement for both regression models). Thus, for this task, pupil average did not appear to reflect overall uncertainty about a future outcome, but rather a specific form of uncertainty that arises after change points and signals the need for rapid learning.

### Pupil metrics reflected individual learning differences

As noted above (Fig. 2c), there was a great deal of variability in the average learning rates used by individual subjects. These individual differences are thought to reflect biases that govern the extent to which subjects tend to interpret the cause of prediction errors in terms of

either noise or change points<sup>2</sup>. One advantage of our reduced model is that it can simulate these individual differences in terms of the subjective hazard rate, which is the expected rate at which change points will occur. Accordingly, fitting the model to behavioral data from individual subjects with subjective hazard rate as a single free parameter yielded fit values that varied systematically with average learning rates ( $r = 0.93$ ,  $H_0: r = 0$ ,  $P < 0.001$ ; Fig. 6a).

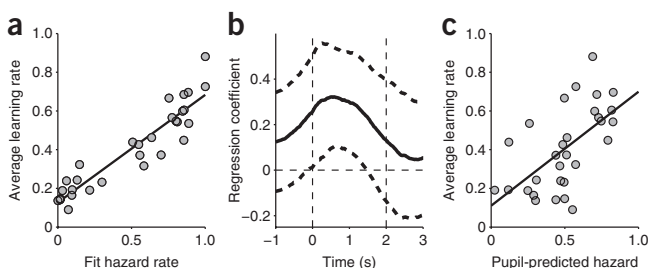
These individual differences in the inferred (fit) subjective hazard rates correspond to individual differences in both the temporal dynamics and magnitude of outcome-locked pupil responses. We quantified the temporal dynamics using an index that related the pupil response on a given trial to a mean-subtracted version of the template shown in Figure 6b. This template describes the strength of the across-subject, linear relationship between pupil diameter and hazard rate in a sliding time window. This relationship was strongest soon after outcome onset, likely reflecting prior expectations about the newly arriving outcome. There was a positive relationship between the mean value of this index and fit hazard rate for individual subjects ( $r = 0.51$ ,  $P < 0.01$ ). In addition, there was a positive relationship between pupil average and fit hazard rate for individual subjects ( $r = 0.40$ ,  $P < 0.05$ ).

On the basis of these relationships, we constructed a linear regression model using the temporal-dynamics index and pupil average to explain individual differences in task performance. The model yielded strong, pupil-based predictions of per-subject values of both fit hazard rate ( $r = 0.59$ ,  $P < 0.001$ ) and average learning rate ( $r = 0.59$ ,  $P < 0.001$ ; Fig. 6c). Thus, individual differences in average learning rate, which can be described computationally as differing expectations about the rate of change points, can be predicted from the temporal dynamics and average magnitude of pupil diameter measured during outcome viewing.

### Pupil metrics predicted trial-by-trial learning rates

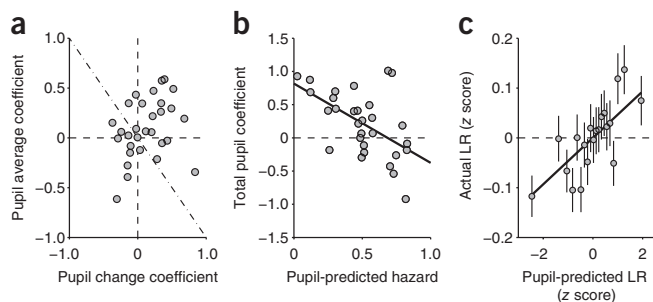
The relationships between pupil metrics and parameters of the reduced Bayesian model suggest that measurements of pupil diameter during the outcome-viewing period can be used to predict the subsequent learning rate. For example, we found positive relationships between pupil change and change-point probability (Fig. 4) and between pupil average and relative uncertainty (Fig. 5). Thus, observing relatively high values of either pupil metric on a given trial should indicate that the subject will use a larger-than-average learning rate when adjusting beliefs according to the outcome observed on that trial. We tested this idea directly, as follows.

First, we examined the relationship between pupil change, pupil average and learning rate for individual subjects. We used a regression model to describe learning rate (z-scored per subject,  $z_{LR}$ ) in terms of pupil change and pupil average. On average, this linear regression computed per subject yielded a positive coefficient for pupil change



**Figure 6** Individual differences in learning rate, hazard rate and pupil diameter. (a) Mean learning rate per subject versus the hazard rate of the reduced Bayesian model that best fit that subject's performance (points). The solid line is a linear fit ( $r = 0.93$ ,  $P < 0.001$ ). (b) Regression coefficients describing the relationship between fit hazard rates and bin-by-bin pupil measurements across subjects, computed in sliding 8.3-ms bins and aligned to outcome presentation (time = 0). Dotted lines indicate 95% confidence intervals. (c) Relationship between pupil-predicted hazard rate and average learning rate for each subject (points). Pupil-predicted hazard rates were computed using a linear regression model that included both shape and magnitude of the average pupil response for each subject (see Online Methods). The solid line is a linear fit ( $r = 0.59$ ,  $P < 0.001$ ).





**Figure 7** Pupil metrics predict learning rate. **(a)** Regression coefficients describing the linear, trial-by-trial relationships between pupil change and the subsequent learning rate (ordinate) and between pupil average and the subsequent learning rate (abscissa). Points are regression coefficients computed for each subject individually, using a four-parameter regression model that also included trial number and block number as covariates. **(b)** The relationship between learning rate and pupil parameters depended on the subject's baseline pupil response. For each subject, the sum of the regression coefficients from **a** are plotted as a function of the pupil-predicted hazard rate from **Figure 6c**. The line is a linear fit ( $r = -0.059$ ,  $P < 0.001$ ). **(c)** Predicted versus actual learning rate (LR). Both values are z-scored per subject. Points and error bars represent mean  $\pm$  s.e.m. values from 20 equally spaced bins of predicted learning rate across all subjects. The line is a linear fit to the unbinned data (slope =  $0.052 z_{\text{actual}}/z_{\text{predicted}}$ ,  $P < 0.001$  for  $H_0$ , slope = 0).

(mean =  $0.108 z_{\text{LR}}/z_{\text{PC}}$ ,  $P < 0.05$  for  $H_0$ ; mean = 0) and a smaller, not statistically significant, positive coefficient for pupil average (mean =  $0.085 z_{\text{LR}}/z_{\text{PA}}$ ,  $P = 0.13$ ; **Fig. 7a**).

Second, we used a simple, weighted sum of pupil change and pupil average to assess their combined predictive power across subjects. Using weights equal to the mean value of the per-subject regression coefficients from the previous analysis (**Fig. 7a**), the weighted sum was moderately predictive of learning rate across all subjects ( $r = 0.067$ ,  $P < 0.001$ ). However, this analysis did not take into account a systematic, negative dependence of the sum of these per-subject coefficients (which is related to the overall ability of the weighted sum to account for learning rate) on subjective hazard rate predicted by pupil dynamics (**Fig. 7b**). Subjects with low pupil-predicted hazard rates had pupil responses that were good predictors of learning rate. Subjects with increasingly high pupil-predicted hazard rates had pupil responses that were increasingly less predictive, and in some cases negatively predictive, of learning rate.

Third, we used a more complicated linear model that also included across-subject differences in pupil dynamics that related to subjective hazard rates, which markedly improved our overall ability to use pupil metrics to predict learning rates. This model had three terms: the sum

of pupil change and pupil average computed per trial, weighted according to average regression coefficients (**Fig. 7a**), the pupil-predicted hazard rate, computed per subject (**Fig. 6c**), and the multiplicative interaction between these two variables. Using this model, pupil measurements could effectively predict learning rates for all data from all subjects ( $r = 0.38$ ,  $P < 0.001$ ). These predictions accounted for variations in learning rates both across (**Fig. 6b**) and within (**Fig. 7c**) subjects.

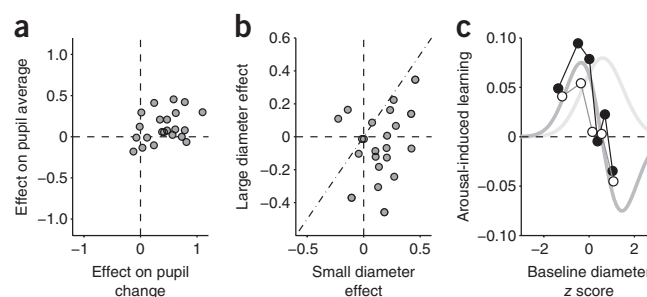
### Task-independent pupil manipulation altered behavior

To examine whether the correlations between pupil measures and learning behavior might reflect an underlying causal process, we used an arousal manipulation that affected pupil diameter and measured its effects on learning behavior. In particular, we occasionally and without warning switched the auditory cue that preceded fixation. Subjects were told that these auditory-cue switches were unrelated to the task and that they should therefore ignore the specific sounds. Nevertheless, this manipulation led to increases in both pupil average and pupil change on trials in which the fixation cue was switched ( $t$  test for  $H_0$ : mean effect size = 0,  $P < 0.001$  for both pupil average and pupil change; **Fig. 8a**). Thus, we caused consistent changes in the pupil measures that were correlated with the computational variables needed to solve the task.

This manipulation resulted in systematic changes in task performance that depended on baseline pupil diameter (**Fig. 8b**). For trials with relatively small baseline diameter (that is, less than its per-subject median value), individual subjects tended to use larger learning rates on auditory-switch trials than otherwise (mean across subjects =  $0.113$ ,  $t$  test for  $H_0$ : mean = 0,  $P < 0.01$ ; **Fig. 8b**). For trials with relatively large baseline diameter, subjects used slightly smaller learning rates on auditory-switch trials than otherwise, although this trend was not statistically significant (mean =  $-0.037$ ,  $P = 0.35$ ; **Fig. 8b**). The average difference in the size of these effects from small- versus large-diameter trials was greater than zero, implying that the effects of this manipulation depended on baseline pupil diameter (paired  $t$  test,  $P < 0.001$ ; **Fig. 8b**). These effects were not the results of systematic differences in task conditions for switch versus nonswitch trials, as the same three analyses yielded no effects when applied to learning rates computed by our reduced Bayesian model ( $P > 0.5$ ).

This dependence on baseline pupil diameter is suggestive of the Yerkes-Dodson 'inverted U' relationship between arousal and learning. According to this idea, learning is highest for moderate levels of arousal and lowest for either overly high or overly low levels of arousal<sup>20</sup>. Our subjects appeared to be consistently engaged during task performance, implying that we were probably not sampling overly low or high arousal states. Nevertheless, in a narrower range and

**Figure 8** Effects of the pupil manipulation. **(a)** Evoked changes in pupil diameter. For each subject, pupil average (ordinate) and pupil change (abscissa) were z-scored across all trials. Each point represents the difference in the mean z scores for auditory switch versus nonswitch trials for an individual subject. Positive values indicate larger values on switch trials. **(b)** Evoked changes in learning behavior. For each subject, learning rate was z-scored across all trials and fit to a cumulative Weibull as a function of error magnitude for each noise condition, to account for the relationship shown in **Figure 4a**. Each point represents the difference in the mean value of the residuals from these fits for auditory switch versus nonswitch trials for an individual subject, separated by trials in which the initial pupil diameter was smaller (ordinate) or larger (abscissa) than its median value. Positive values indicate larger learning rates on auditory switch trials. **(c)** A possible relationship between learning and arousal based on an inverted U (light gray, modeled as Gaussian). A given change in learning for a given a change in pupil metrics (ordinate), plotted as a function of baseline pupil diameter (abscissa), is shown for the hypothesized Gaussian (its derivative is shown in dark gray), the measured effects of the auditory manipulation (open points), and the measured relationship between pupil metrics and learning rate during non-manipulation sessions.



assuming a correspondence between arousal state and baseline pupil diameter, we found that the relationships between learning behavior and our arousal manipulation were qualitatively consistent with an inverted U. In particular, auditory-switch trials tended to correspond to the largest increases in learning rate when baseline pupil diameter was relatively low (steepest ascent in the inverted U) and the largest decreases in learning rate when baseline pupil diameter was relatively high (steepest descent in the inverted U; **Fig. 8c**).

This inverted U relationship was also apparent in our previous pupil measurements in two ways. First, across subjects, those with larger-than-average pupil diameters during outcome viewing tended to use learning rates that were less, or even negatively, predicted by fluctuations in pupil metrics relative to other subjects (**Fig. 7b**). Second, subjects that had lower pupil-predicted hazard rates used learning rates that were positively correlated with pupil metrics when their baseline pupil diameter was lowly, but negatively, correlated when their baseline pupil diameter was high (**Fig. 8c**). Thus, results from both our pupil-manipulation and pupil-measurement experiments were consistent with the arousal system being important for the rational regulation of learning.

## DISCUSSION

We examined the relationship between pupil diameter, which is related to arousal and autonomic state, and learning rate, which describes the extent to which new information is used to adjust existing cognitive beliefs. Consistent with previous work<sup>2,21,22</sup>, we found that human subjects performing a predictive-inference task were most heavily influenced by outcomes that occurred shortly after a change point in the outcome-generating process. One possible mechanism for this effect is a dynamic regulation of the relative influence of incoming information on cortical processing<sup>3</sup>. Insights into the computations required for such a regulator are provided by a reduced model that approximates the ideal observer for the task, describes subject behavior and bases learning rates on two parameters that we found to be represented in pupil measurements: change-point probability and relative uncertainty.

In our model, change-point probability depends on the absolute value of the most recent prediction error and drives increased learning after surprisingly large errors. We found that change-point probability was positively correlated with changes in pupil diameter. This relationship is consistent with early pupillometry studies that found an inverse relationship between stimulus-evoked pupil responses and stimulus probability, as well as with more recent work interpreting outcome-locked pupil responses in terms of the surprise associated with errors in judging uncertainty, referred to as the risk prediction error<sup>22–24</sup>. We also found that pupil change was not always directly related to change-point probability, with particularly large pupil changes on trials with exactly zero error that might have been surprisingly rewarding and/or reflected an association with an atypical consequence (that is, no possibility of updating the next prediction).

Relative uncertainty, the second parameter in our model, represents uncertainty about the true underlying mean and drives learning from outcomes that occur after a change point. We found that relative uncertainty was correlated with average pupil diameter. We also found that changes in another form of uncertainty that should not drive learning (that is, changes in the s.d. of the generative process in our task) did not lead to similar effects on pupil diameter. These results are consistent with the recent finding that pupil diameter tends to increase during exploratory decisions that occur during periods of uncertainty about the best available option<sup>6</sup>. These findings suggest

that pupil-linked arousal systems encode an uncertainty signal that facilitates both learning and information-seeking behaviors.

We also found strong individual differences in task behavior that could be captured by fitting a prior expectation about the rate of change points (hazard rate) to behavioral data. We found that subjects who were fit by higher hazard-rate models tended to have larger pupil dilations during the outcome-viewing period. This physiological difference arose early in the viewing period, consistent with the idea that these individual differences reflected a prior expectation about the source of the upcoming error.

We used these relationships between pupil metrics and change-point probability, relative uncertainty and the hazard-rate before predicting the extent to which subjects were influenced by each new outcome. We also manipulated pupil diameter using a task-irrelevant auditory manipulation that resulted in changes in task performance that were consistent with our measured relationships between pupil metrics and key task variables. These results provide new insights into the specific computations that are reflected in pupil diameter and establish their causal role in belief updating.

These computations likely involve, at least in part, neural activity in the locus coeruleus. One possibility is that the two important variables from our model are encoded by two distinct modes of locus coeruleus activation<sup>5</sup>: change-point probability, reflected in pupil change, is encoded by phasic activation of the locus coeruleus, whereas relative uncertainty, reflected in pupil average, is encoded by tonic activation of the locus coeruleus. Although direct confirmation is still needed, this idea is supported by several lines of evidence, including a compelling example of simultaneous measurements of locus coeruleus activity and pupil diameter in a monkey that are closely correlated<sup>5</sup>, similar modulations of pupil diameter and locus coeruleus activity under certain task conditions, such as changes in utility that affect behavioral engagement<sup>6,7</sup>, and a proposed anatomical substrate involving common activation from the nucleus paragigantocellularis, which contributes to both locus coeruleus and sympathetic nervous system function<sup>4,25</sup>. The consequence of locus coeruleus involvement would be the task-related release of norepinephrine throughout the nervous system. Consistent with our results, norepinephrine release is thought to permit or facilitate changes in behavior that follow unexpected changes in the environment and learning in general, possibly by modulating experience-dependent neural plasticity<sup>3,26–31</sup>.

More generally, our results are consistent with the idea that brain areas that regulate the influence of newly arriving information on existing beliefs are also strongly linked to arousal and autonomic function<sup>1,3,6,7,24,32,33</sup>. These areas likely include not just the locus coeruleus, but also the anterior cingulate cortex, which has strong reciprocal connections with the locus coeruleus and whose activity encodes several signals closely related to change-point probability, including unsigned prediction errors and learning rates<sup>1,5,34,35</sup>. This arousal system appears to govern not simply overall alertness or other nonspecific factors that might affect overall task performance, but rather a computationally sophisticated process that rationally regulates the influence of new sensory information in a dynamic environment. These computations take into account both ongoing processing of task-relevant variables, such as change-point probability and relative uncertainty, and state variables, such as prior expectations about the rate of change. These factors are combined in a manner that is consistent with the Yerkes-Dodson inverted U relationship between arousal level and learning rate (**Fig. 8c**)<sup>20</sup>.

In summary, our findings suggest a relationship between arousal state and learning rate that is likely a result of a coordinated

learning-arousal network, including the locus coeruleus and anterior cingulate cortex. The representation of normative learning variables in this network suggests that subtle changes in arousal might reflect rational regulation of the influence of new information on ongoing inferences about a dynamic world.

## METHODS

Methods and any associated references are available in the online version of the paper.

## ACKNOWLEDGMENTS

We thank J. Cohen, S. du Lac, L. Ding, Y. Li and J. Nassar for helpful comments. This work was supported by the US National Institutes of Health (R01 EY015260, F31 MH093099 and T90 DA22763), the McKnight Endowment Fund for Neuroscience, the Burroughs-Wellcome Fund and the Sloan Foundation.

## AUTHOR CONTRIBUTIONS

M.R.N., J.I.G. and B.H. designed the experiment and tasks. M.R.N., K.M.R. and K.P. collected and analyzed data. M.R.N. and R.C.W. developed and applied the reduced Bayesian model. M.R.N. and J.I.G. wrote the manuscript.

## COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at <http://www.nature.com/doi/10.1038/nn.3130>.

Reprints and permissions information is available online at <http://www.nature.com/reprints/index.html>.

- Behrens, T.E., Woolrich, M.W., Walton, M.E. & Rushworth, M.F. Learning the value of information in an uncertain world. *Nat. Neurosci.* **10**, 1214–1221 (2007).
- Nassar, M.R., Wilson, R.C., Heasly, B. & Gold, J.I. An approximately Bayesian delta-rule model explains the dynamics of belief updating in a changing environment. *J. Neurosci.* **30**, 12366–12378 (2010).
- Yu, A.J. & Dayan, P. Uncertainty, neuromodulation, and attention. *Neuron* **46**, 681–692 (2005).
- Nieuwenhuis, S., De Geus, E.J. & Aston-Jones, G. The anatomical and functional relationship between the P3 and autonomic components of the orienting response. *Psychophysiology* **2**, 162–175 (2011).
- Aston-Jones, G. & Cohen, J.D. An integrative theory of locus coeruleus–norepinephrine function: adaptive gain and optimal performance. *Annu. Rev. Neurosci.* **28**, 403–450 (2005).
- Jepma, M. & Nieuwenhuis, S. Pupil diameter predicts changes in the exploration-exploitation trade-off: evidence for the adaptive gain theory. *J. Cogn. Neurosci.* **23**, 1587–1596 (2011).
- Gilzenrat, M.S., Nieuwenhuis, S., Jepma, M. & Cohen, J.D. Pupil diameter tracks changes in control state predicted by the adaptive gain theory of locus coeruleus function. *Cogn. Affect. Behav. Neurosci.* **10**, 252–269 (2010).
- Krugman, H.E. Some applications of pupil measurement. *J. Mark. Res.* **1**, 15–19 (1964).
- Granholm, E. & Steinhauer, S.R. Pupillometric measures of cognitive and emotional processes. *Int. J. Psychophysiol.* **52**, 1–6 (2004).
- Schmidt, H.S. & Fortin, L.D. Electronic pupillography in disorders of arousal. in *Sleeping and Waking Disorders: Indication and Technique* (ed. Guilleminault, C.) 127–143 (Addison-Wesley, 1982).
- Kahneman, D. & Beatty, J. Pupil diameter and load on memory. *Science* **154**, 1583–1585 (1966).
- Richer, F. & Beatty, J. Contrasting effects of response uncertainty on the task-evoked pupillary response and reaction time. *Psychophysiology* **24**, 258–262 (1987).
- Hakerem, G., Sutton, S. & Zubin, J. Pupillary reactions to light in schizophrenic patients and normals. *Ann. NY Acad. Sci.* **105**, 820–831 (1964).
- Einhäuser, W., Stout, J., Koch, C. & Carter, O. Pupil dilation reflects perceptual selection and predicts subsequent stability in perceptual rivalry. *Proc. Natl. Acad. Sci. USA* **105**, 1704–1709 (2008).
- Van Olst, E.H., Heemstra, M.L. & Ten Kortenaar, T. Stimulus significance and the orienting reaction. in *The Orienting Reflex in Humans: International Conference Proceedings* (eds. Kimmel, H., Olst, E.H. & van Orlebeke, J.F.) 521–547 (Erlbaum, 1979).
- Sutton, R.S. & Barto, A.G. *Reinforcement Learning: An Introduction* (MIT Press, Cambridge, Massachusetts, 1998).
- Adams, R.P. & MacKay, D.J.C. Bayesian online changepoint detection. Preprint at <<http://arxiv.org/abs/0710.3742v1>> (2007).
- Fearnhead, P. & Liu, Z. On-line inference for multiple changepoint problems. *J. R. Stat. Soc. Series B Stat. Methodol.* **69**, 589–605 (2007).
- Wilson, R.C., Nassar, M.R. & Gold, J.I. Bayesian online learning of the hazard rate in change-point problems. *Neural Comput.* **22**, 2452–2476 (2010).
- Yerkes, R.M. & Dodson, J.D. The relation of strength of stimulus to rapidity of habit-formation. *J. Comp. Neurol. Psychol.* **18**, 459–482 (1908).
- Behrens, T.E., Woolrich, M.W., Walton, M.E. & Rushworth, M.F. Learning the value of information in an uncertain world. *Nat. Neurosci.* **10**, 1214–1221 (2007).
- Raisig, S., Welke, T., Hagendorf, H. & van der Meer, E. I spy with my little eye: detection of temporal violations in event sequences and the pupillary response. *Int. J. Psychophysiol.* **76**, 1–8 (2010).
- Friedman, D., Hakerem, G., Sutton, S. & Fleiss, J.L. Effect of stimulus uncertainty on the pupillary dilation response and the vertex evoked potential. *Electroencephalogr. Clin. Neurophysiol.* **34**, 475–484 (1973).
- Preuschhoff, K., 't Hart, B.M. & Einhäuser, W. Pupil dilation signals surprise: evidence for noradrenaline's role in decision making. *Front. Neurosci.* **5**, 115 (2011).
- Aston-Jones, G., Ennis, M., Pieribone, V.A., Nickell, W.T. & Shipley, M.T. The brain nucleus locus coeruleus: restricted afferent control of a broad efferent network. *Science* **234**, 734–737 (1986).
- Sara, S.J., Vankov, A. & Hervé, A. Locus coeruleus-evoked responses in behaving rats: a clue to the role of noradrenaline in memory. *Brain Res. Bull.* **35**, 457–465 (1994).
- Aston-Jones, G., Rajkowski, J. & Kubiak, P. Conditioned responses of monkey locus coeruleus neurons anticipate acquisition of discriminative behavior in a vigilance task. *Neuroscience* **80**, 697–715 (1997).
- Tully, K. & Bolshakov, V.Y. Emotional enhancement of memory: how norepinephrine enables synaptic plasticity. *Mol. Brain* **3**, 15 (2010).
- Harley, C.W. A role for norepinephrine in arousal, emotion and learning? Limbic modulation by norepinephrine and the Kety hypothesis. *Prog. Neuropsychopharmacol. Biol. Psychiatry* **11**, 419–458 (1987).
- Corbetta, M., Patel, G. & Shulman, G.L. The reorienting system of the human brain: from environment to theory of mind. *Neuron* **58**, 306–324 (2008).
- Bouret, S. & Sara, S.J. Network reset: a simplified overarching theory of locus coeruleus noradrenaline function. *Trends Neurosci.* **28**, 574–582 (2005).
- Critchley, H.D., Mathias, C.J. & Dolan, R.J. Neural activity in the human brain relating to uncertainty and arousal during anticipation. *Neuron* **29**, 537–545 (2001).
- Critchley, H.D. Neural mechanisms of autonomic, affective, and cognitive integration. *J. Comp. Neurol.* **493**, 154–166 (2005).
- Krugel, L.K., Biele, G., Mohr, P.N., Li, S.C. & Heekeren, H.R. Genetic variation in dopaminergic neuromodulation influences the ability to rapidly and flexibly adapt decisions. *Proc. Natl. Acad. Sci. USA* **106**, 17951–17956 (2009).
- Matsumoto, M., Matsumoto, K., Abe, H. & Tanaka, K. Medial prefrontal cell activity signaling prediction errors of action values. *Nat. Neurosci.* **10**, 647–656 (2007).

## ONLINE METHODS

**Predictive-inference task.** Human subject protocols were approved by the University of Pennsylvania Internal Review Board. Thirty subjects (19 female, 11 male; age range = 19–29 years) participated in the primary study and an independent sample of 29 subjects (17 female, 12 male; age range = 19–25 years) participated in the arousal manipulation study after providing informed consent. Both studies used a predictive-inference task that required subjects to predict each subsequent number to be presented in a series<sup>2</sup>. For each trial  $t$ , a single integer ( $X_t$ ) was presented that was a rounded pick sampled independently and identically from a Gaussian distribution whose mean ( $\mu_t$ ) changed at unsignaled change points and whose s.d. ( $\sigma_t$ ) was fixed to either 5 or 10 in each block of 200 trials. Change points occurred with a probability of zero for the first three trials following a change point and 0.1 for all trials thereafter.

To facilitate measurements of nonluminance-mediated effects on pupil diameter, we used a different visual display and task timing than in our previous study<sup>2</sup>. Subjects were shown a numeric representation of their current prediction at a central location on a CRT monitor. Background screen pixels were a checkerboard of light and dark pixels (mean  $\pm$  s.d. luminance in a circle with radius of 6.5 cm =  $0.457 \pm 0.010$  cd m<sup>-2</sup>). Numbers were drawn in an intermediate gray color ( $0.445 \pm 0.005$  cd m<sup>-2</sup>). When viewed passively by a control group of four subjects outside of the context of the predictive-inference task, no individual stimulus (number) had a significant effect on average pupil diameter or evoked changes in pupil diameter ( $t$  test for  $H_0$ : equal means between each stimulus and all others,  $P > 0.3$  for all stimuli after correcting for multiple comparisons), nor did the number of digits contained in the stimulus affect either pupil variable ( $P > 0.4$ ).

For each trial, the subject indicated his or her updated prediction using a video gamepad. Each prediction was constrained to be between the previous prediction and the most recent outcome, limiting learning rates to between zero and 1. After the new prediction was chosen, the numeric representation of this prediction disappeared, an auditory cue was played and a numeric representation of the new outcome was shown. Subjects were instructed to fixate centrally for 2 s at this point; failure to do so (in a square window, 9° per side) resulted in a tone indicating a fixation error. After 2 s, the new outcome disappeared, the prediction reappeared and an auditory cue was played to indicate that the prediction should be updated. 14 subjects also participated in a control version of the task in which the prediction was reset after viewing the new outcome to reflect an update equivalent to a learning rate of 0.5. For this task, the same motor output (in terms of number or duration of button presses) was required to use a learning rate of either zero or 1 on each trial.

Subjects were told that the numbers were generated from a noisy process and that several discreet change points would occur over the course of the task. They were instructed to make a prediction on each trial ( $B_t$ ) such that the average error made on all predictions,  $\langle |B_t - X_t| \rangle$ , would be minimized. Payout depended on how well they achieved this goal, as described previously<sup>2</sup>.

The pupil-manipulation task was identical to primary version of the task, except that the auditory cue played at the beginning of fixation was occasionally switched to another sound from a library of 31 sound effects downloaded from an online library. Sounds were 0.09–1.4 s in duration (mean  $\pm$  s.d. =  $0.72 \pm 0.42$  s) and played at 56–70 dB ( $A$ -weighted; mean  $\pm$  s.d. =  $62.5 \pm 3.9$  dB). Switch trials occurred at random, with a probability of 0.1 on the nine trials following a switch, 0.8 thereafter. On switch trials, the given sound was played, on average, 7 dB louder than otherwise. We excluded 7 of 29 subjects completing the pupil-manipulation task from further analyses because of an excessive number of fixation errors (blinks or lost fixation on >40% of trials).

**Pupil-diameter measurements.** Pupil diameter was sampled at 120 Hz and recorded throughout the task using an infrared video eye-tracker (ASL). Blinks were identified using a custom blink filter on the basis of pupil diameter and vertical and horizontal eye position, then removed by linear interpolation of values measured just before and after each identified blink. Blink-filtered diameter was low-pass filtered using a Butterworth filter with a cutoff frequency of 3.75 Hz. These filtered measurements were then  $z$ -scored in each session.

All analyses excluded trials in which blinks or fixation errors during outcome viewing were detected online (these events were followed by a beep to remind the subject to minimize their occurrence). The first 20 trials from each block were also excluded to avoid possible changes in average luminance at block boundaries.

Pupil average was computed for each trial by taking the mean of all 240  $z$ -scored pupil measurements from the 2-s-long outcome-viewing period of the trial. Pupil change was computed for each trial by subtracting the average pupil measurement from early in the outcome-viewing period (0–1 s after outcome presentation) from the average pupil measurement from late in the outcome-viewing period (1–2 s after outcome presentation). Trials that included blinks that were detected offline (but not online) were used to compute pupil average by interpolating values from just before and just after the blink. These trials were not used to compute pupil change, which was much more sensitive to the timing of blinks.

**Reduced Bayesian model.** Optimal performance on the predictive-inference task requires inferring the probability distribution over possible outcomes on the next time step, given all previous data and the process by which those data were generated,  $p(X_{t+1}|X_{1:t})$ . Because the relationship between the data on the next time step is independent of all previous data conditioned on the mean of the current distribution ( $\mu$ ), the solution can be formulated in terms of  $\mu$

$$p(X_{t+1}|X_{1:t}) = \sum_{\mu_t} p(X_{t+1}|\mu_t) p(\mu_t|X_{1:t}) \quad (1)$$

and the probability distribution over possible means given previous data can be inverted according to Bayes' rule

$$p(\mu_t|X_{1:t}) = \frac{p(X_{1:t}|\mu_t)p(\mu_t)}{p(X_{1:t})} \quad (2)$$

Although computationally tractable solutions to this problem exist, these solutions specify learning rates that are complicated functions of either the probability distribution over all possible means<sup>1</sup> or over all possible 'runs' of nonchange-point trials<sup>19</sup>. To simplify the algorithm, the reduced model computes the posterior probability distribution over possible means as described above, but maintains only the first two moments of this distribution. This assumption massively reduces the number of required computations, but has minimal effects on performance<sup>2</sup>. An added advantage of this model is that it can be formulated as a delta rule

$$\begin{aligned} B_{t+1} &= B_t + \alpha_t \times \delta_t \\ \delta_t &= X_t - B_t \end{aligned} \quad (3)$$

where  $B$  is the belief about the mean of the underlying distribution,  $\alpha$  is the learning rate and  $\delta$  is the prediction error, which is the difference between the actual and predicted outcome. The learning rate depends on two variables that are updated on each trial

$$\alpha_t = \tau_t + (1 - \tau_t)\Omega_t \quad (4)$$

where change-point probability ( $\Omega$ ) reflects the probability that  $\mu_t$  is not equal to  $\mu_{t-1}$ , and relative uncertainty ( $\tau$ ) reflects the variance on the predictive distribution in  $\mu$  (that is, uncertainty about the location of the mean) divided by the variance on the predictive distribution in  $X$  (that is, total uncertainty about the location of the next outcome).

Performance of the reduced Bayesian model also depends on an expectation about the prior probability on change points, or the hazard rate. Specifically, hazard rate directly influences the computation of change-point probability on each trial

$$\Omega_t = \frac{U(X_t|0, 300)H}{U(X_t|0, 300)H + N(X_t|B_t, \sigma_t^2)(1-H)} \quad (5)$$

where  $U$  and  $N$  represent uniform and normal distributions, respectively;  $H$  is the hazard rate,  $B_t$  is the model's prediction on trial  $t$  and  $\sigma^2$  is the total variance of the predictive distribution. We incorporated hazard rate into the model in two ways: using the true generative hazard rate for trials in which a change point did not recently occur (0.1), or by fitting the model to behavior by minimizing the total squared difference between subject and model predictions using a constrained search algorithm (fmincon in MATLAB) with hazard rate as a free parameter.



The total variance on the predictive distribution in the model comes from two sources

$$\sigma_t^2 = N^2 + \frac{\tau_t N^2}{1 - \tau_t} \quad (6)$$

The first source is the s.d. on the outcome-generating distribution ( $N$ ). The second source is uncertainty about the mean of that distribution and depends on both  $N$  and relative uncertainty ( $\tau$ ). Here we set  $N$  to be the actual experimental s.d., but we updated  $\tau$  after each outcome according to the variance on the predictive distribution over possible means

$$\tau_{t+1} = \frac{N^2 \Omega_t + (1 - \Omega_t)(\tau_t N^2) + \Omega_t(1 - \Omega_t)(\tau_t + B_t(1 - \tau_t) - X_t)}{N^2 \Omega_t + (1 - \Omega_t)(\tau_t N^2) + \Omega_t(1 - \Omega_t)(\tau_t + B_t(1 - \tau_t) - X_t) + N^2} \quad (7)$$

such that if a change point occurs, relative uncertainty is reset to 0.5 (first term in numerator); if a change point does not occur, relative uncertainty is reduced (second term in numerator), and if the model is uncertain about whether a change point occurred, relative uncertainty is increased to reflect this uncertainty (third term in numerator).

**Statistical analyses.** Trial-by-trial values of pupil average and pupil change were each  $z$ -scored for the full session ( $z_{PA}$  and  $z_{PC}$ , respectively) and then fit with a linear regression model using four parameters: change-point probability (computed by the reduced Bayesian model for each trial), relative uncertainty (computed by the reduced Bayesian model for each trial), noise (the s.d. of the outcome-generating distribution) and a binary vector specifying whether or not the subject correctly predicted the outcome on that trial. We also used a larger model that, in addition to the above four parameters, included the average horizontal and vertical eye position, and the change in horizontal and vertical eye position measured during the outcome-viewing period, the subject's prediction and the computer-generated outcome from the current trial, the pupil change measured on the previous trial, and the trial number within the block and within the session.

Pupil-predicted hazard rates were derived from pupil measurements and the reduced Bayesian model as follows. First, we inferred the subjective hazard rate used by each subject by fitting his or her behavioral data to the reduced

Bayesian model with hazard rate ( $H$ ) as the only free parameter. Next, we fit a linear regression model explaining  $H$  in terms of pupil measurements. That model had two terms, computed per subject: the mean value of pupil average and an index of pupil dynamics. The index was computed as the mean value of the dot product of trial-by-trial pupil measurements and the mean-subtracted curve (Fig. 6b). Finally, we used the coefficients from a linear fit that excluded the data from an individual subject to combine the mean pupil average and pupil-dynamics index (from the excluded subject) into a pupil-predicted hazard rate for that subject.

Pupil-predicted learning rates were computed using a regression model that incorporated both within- and across-subject differences in pupil responses to predict trial-by-trial learning rates. The terms in the regression model were a weighted average of pupil average and pupil change for each trial (with weights proportional to average regression coefficients across subjects; see Fig. 7a), the pupil-predicted hazard rate assigned to each subject (Fig. 6c) and the multiplicative interaction between these two variables. This regression model was used to compute a pupil-predicted learning rate for each trial for each subject.

Arousal-induced learning effects for the inverted U analyses were computed separately for sound-manipulation and nonmanipulation sessions. For sound-manipulation sessions, learning rates were fit to a cumulative Weibull as a function of error magnitude for each subject and noise condition, to account for the relationship shown in Figure 4a. Residuals from this fit, which reflected error-independent variability in learning rate, were  $z$ -scored per subject. Initial pupil diameter, as measured by the average diameter during the first 100 ms of the outcome phase, was also  $z$ -scored per subject. Data were binned across subjects according to the initial diameter  $z$  score. The effect of the sound manipulation was computed as a signed  $d'$  describing the difference in the  $z$ -scored residual learning rates used on auditory shift versus nonauditory shift trials. For nonmanipulation sessions, the relationship between pupil metrics and learning rate was characterized only for subjects with low pupil-predicted hazard rates ( $<0.6$ ). Subjects with high pupil-predicted hazard rates tended to have small or negative relationships between pupil metrics and learning rate, and thus were omitted from this analysis. Arousal effect size was computed as the correlation coefficient between the weighted sum of pupil metrics and learning rate, each  $z$ -scored per subject (positive/negative values indicate that learning rates tended to increase/decrease as pupil effects increased) for equally sized bins of baseline pupil diameter ( $z$ -scored per subject).