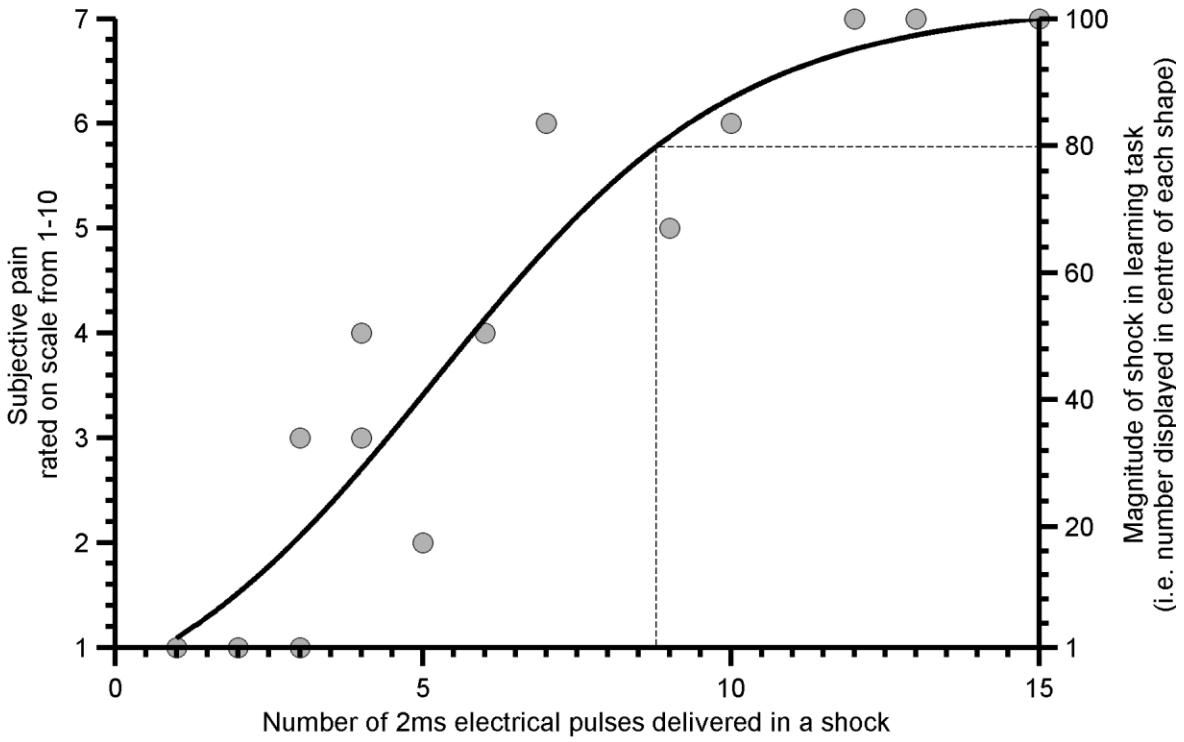
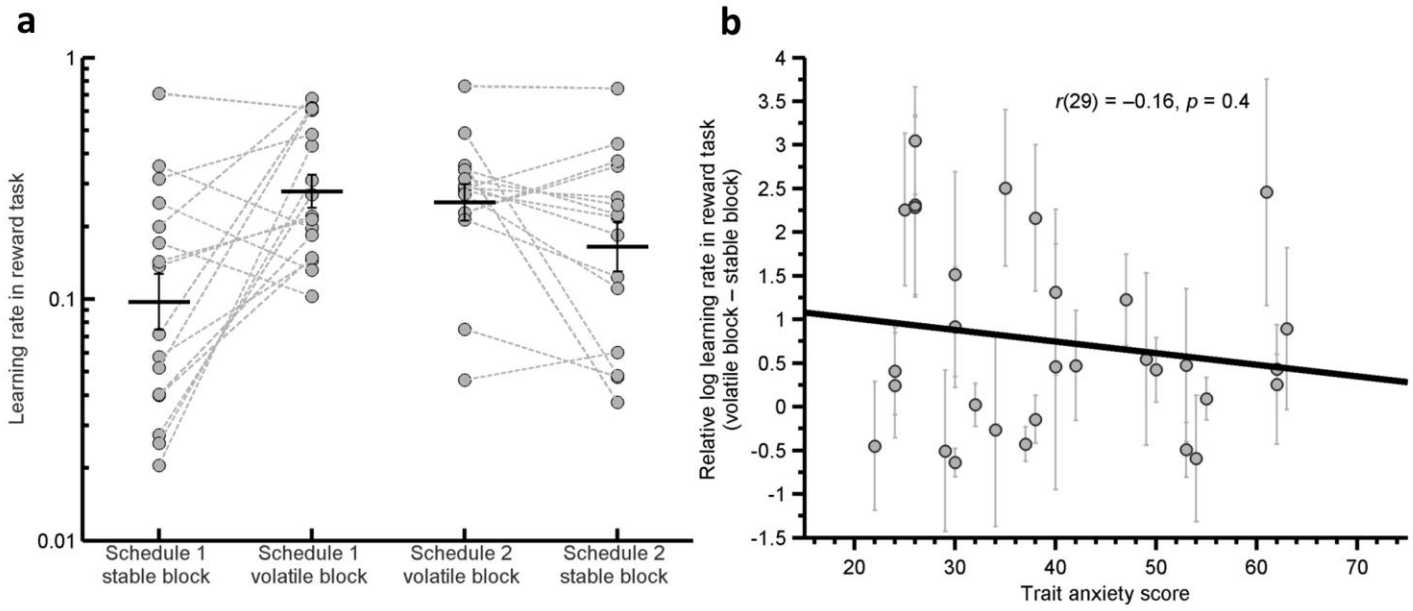


Supplementary Figure 1

Calibration of Electrical Shocks.

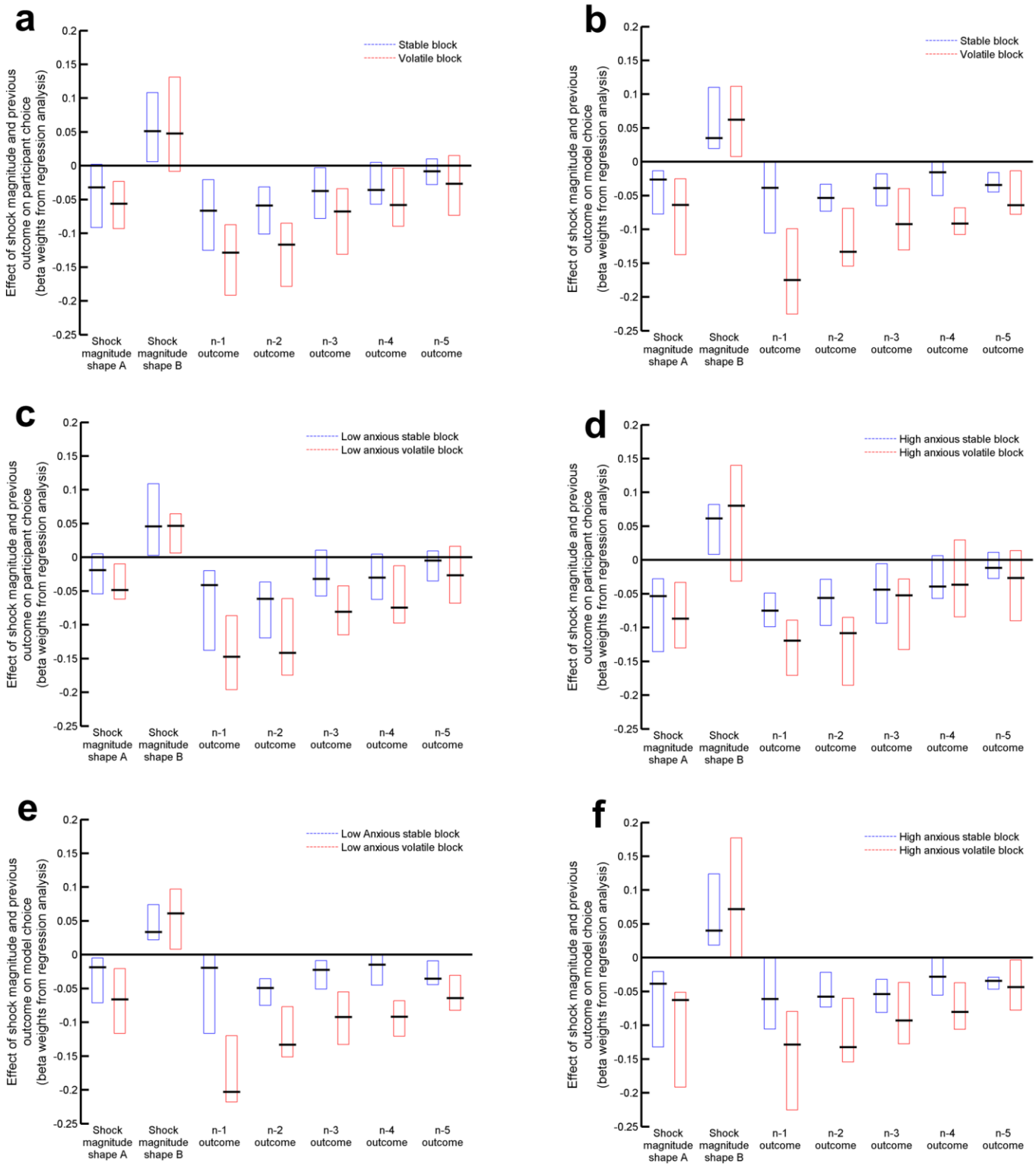
Before each participant completed the aversive learning task, subjective pain scores (left hand axis) were collected for different objective strengths of electrical shock (x axis), see online methods. A 10 point scale was used for subjective pain where 1=minimal pain, 10=worst possible pain. No shock that exceeded a subjective level of 7 was applied. A sigmoid curve fitted to this data was used to equate the subjective level of pain experienced during the learning task across participants. The graph shows data points (circles) and the fitted sigmoid (black line) from one participant. In the task, the shock magnitude associated with each shape was varied on a trial-to-trial basis from 1-99, where 1 equated to a subjective pain level of 1, and 99 to a subjective pain level of 7. Magnitude values between these extremes were calculated using the sigmoidal fit. This is illustrated here by the dashed line, which shows how the objective strength of the electrical stimulus (x axis) corresponding to a desired 'magnitude' of 80 (right hand axis) would be calculated for this participant.



Supplementary Figure 2

Effects of volatility upon participants' learning rates in a structurally equivalent reward task.

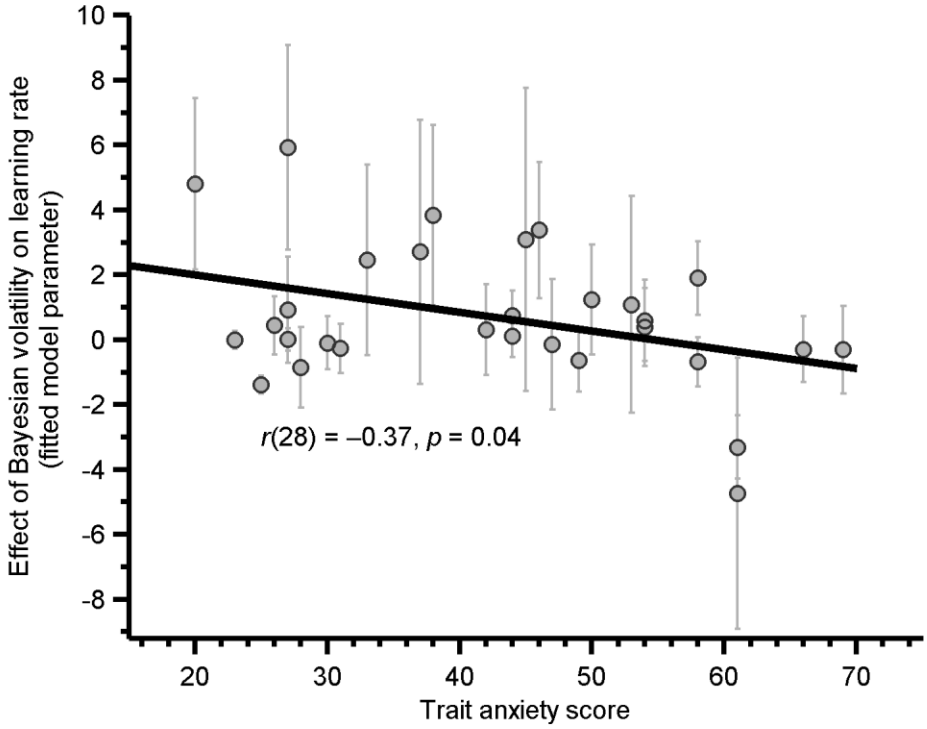
In addition to the aversive learning task, reported within the main manuscript, participants also completed a structurally equivalent reward learning task. Here the two shapes chosen between had different probabilities of leading to winning points, with the magnitude of reward points being unique for each shape and varying randomly between 1 and 99 across trials. Following the task the points participants had won were converted into a monetary reward (either £5 or £10). **(a)** Participants' choices during the stable and volatile blocks of the reward task were fitted with the same Rescorla Wagner learning model used in the aversive task. Estimates of individual participants' learning rates are displayed (circles), using a logarithmic scale, separately for the stable and volatile blocks for the two task schedules (Schedule 1 = stable task block first, n = 16, Schedule 2 = volatile task block first, n = 15). Black lines display mean (+–SEM) of participants' estimated learning rates, grey dotted lines link the learning rates in volatile and stable blocks for each participant. Participants showed higher learning rates in the volatile versus stable blocks regardless of the order in which they were completed, $F(1,29) = 15.3$, $p = 0.001$, replicating previous results[8]. **(b)** The relative log learning rate for the volatile versus the stable blocks (i.e. log(LR in volatile block) – log(LR in stable block)) was not significantly correlated with participant trait anxiety, $r(29) = –0.16$, $p = 0.4$. It should be noted that the difference in the effect of anxiety on learning rate for this reward learning task and the aversive learning task reported in the main manuscript was also not significant, $F(1,28) = 0.57$, $p = 0.46$. This hence limits the conclusions that can be drawn regarding the specificity of the anxiety-related deficit in adjusting learning rate to cases where outcomes are aversive as opposed to reward-related. Error bars represent the standard deviation of the estimated parameters from the behavioral model for each participant.



Supplementary Figure 3

Influence of potential shock magnitude and prior outcome history on actual and simulated behavior choice as a function of block volatility (**a,b**) and participant anxiety **(c-f).**

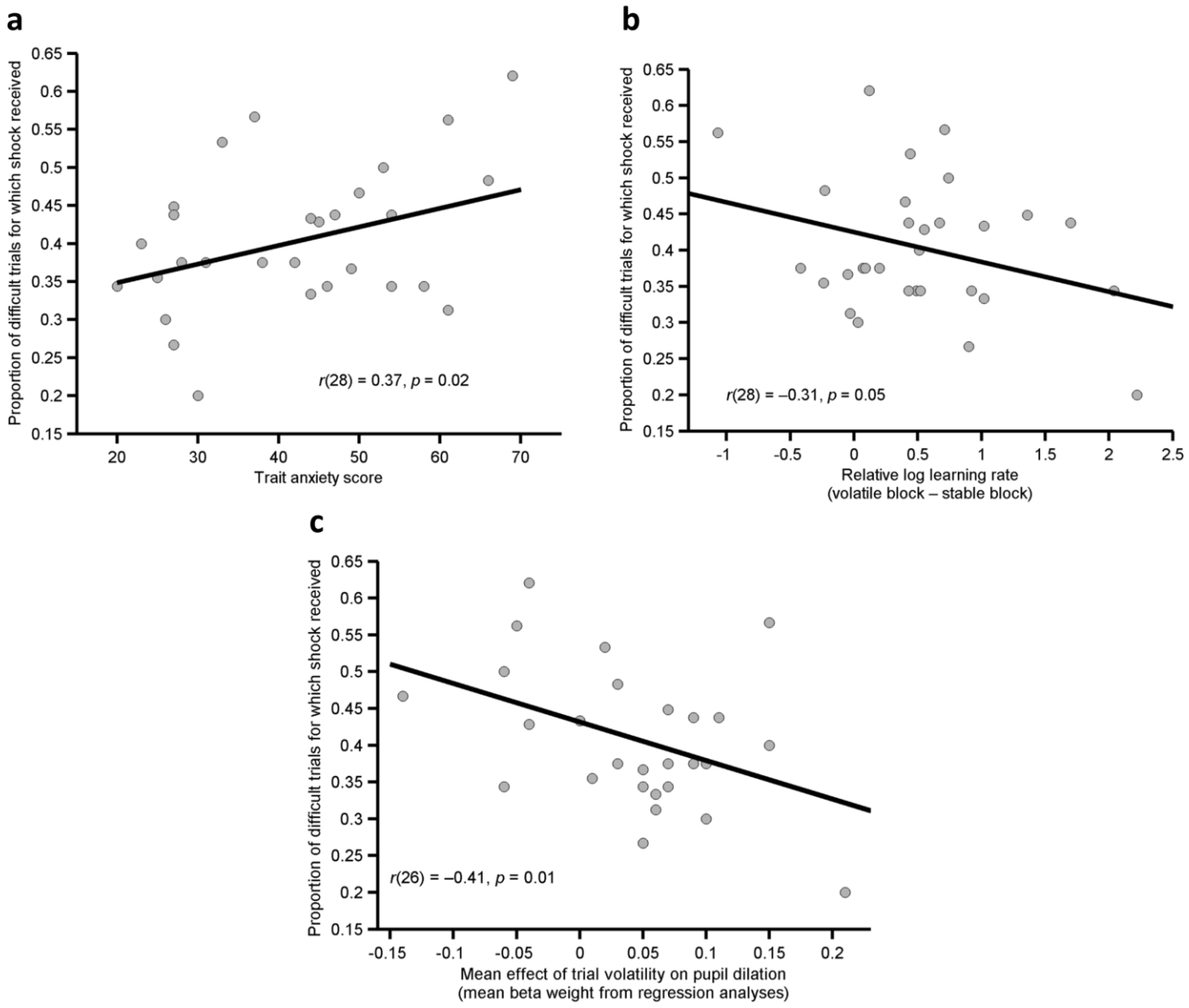
Regression analyses were conducted on a participant by participant basis to examine the extent to which current shock magnitude linked to each shape and the outcomes of the five previous trials predicted choice of shape A versus shape B (e.g. circle versus square) on a given trial. The dependent variable was participant choice (**a,c,d**) or model choice (**b,e,f)** on trial 'n' (coded as 1 if the subject/model chose shape A (e.g. circle) and 0 if shape B (e.g. square)). For the model choice regressions, the coupled Rescorla Wagner (RW) predictor and sofmax action selector learning model was used to simulate choice behavior on a trial to trial basis, using the blockwise estimates of the free parameters (for learning rate, decision temperature and inverse temperature) previously calculated for each participant (see online methods). In each regression analysis, predictor variables comprised shock magnitude for shape A, shock magnitude for shape B, and outcome for trial n-1 to n-5 (coded as 1 if shape A was associated with the shock). As the distribution of parameter estimates across participants was non-normal, median values (horizontal black lines) and interquartile ranges (boxes) are presented. (**a)** Influence of the predictor variables on subject choice, estimated separately for trials in the stable (blue boxes) and volatile (red boxes) blocks. The impact of previous outcomes decreases across time and - consistent with the higher learning rate used by participants in volatile blocks - the impact of prior outcomes on shape choice is greater for more recent trial outcomes (n-1, n-2) in the volatile versus the stable block. (**b)** A parallel analysis was run on the model-derived choices. The similar pattern in (**b**) as for (**a**) is as expected given the good fit of the RW & softmax action selector model to the behavior data (see Supplementary Modeling Note). Plots **c** and **d** illustrate how the influence of prior outcomes on behavioral choice as a function of block volatility differs between low and high trait anxious individuals (as defined using a median split, for ease of graphical illustration). As expected given the inverse relationship between trait anxiety and change in learning rate between stable and volatile blocks (**Fig 2b**), a greater difference between blocks in the effect of recent prior outcomes on choice behavior is observed for low (**c**) than for high (**d**) anxious participants. Plots **e** and **f** illustrate that a parallel pattern is observed when using simulated data from the RW and softmax action selector model. In summary, panels **a**, **c** and **d** illustrate, in a model-free manner, the key behavior of participants assessed by the learning rate analyses reported within the main manuscript. Additionally, the plots shown in panels **b**, **e** and **f** indicate that the coupled RW predictor and softmax action selector is able to reliably capture these behaviors.



Supplementary Figure 4

Relationship between trait anxiety and Bayesian volatility when multiple parameters, including a decay function, are allowed to compete for influence over a dynamic learning rate.

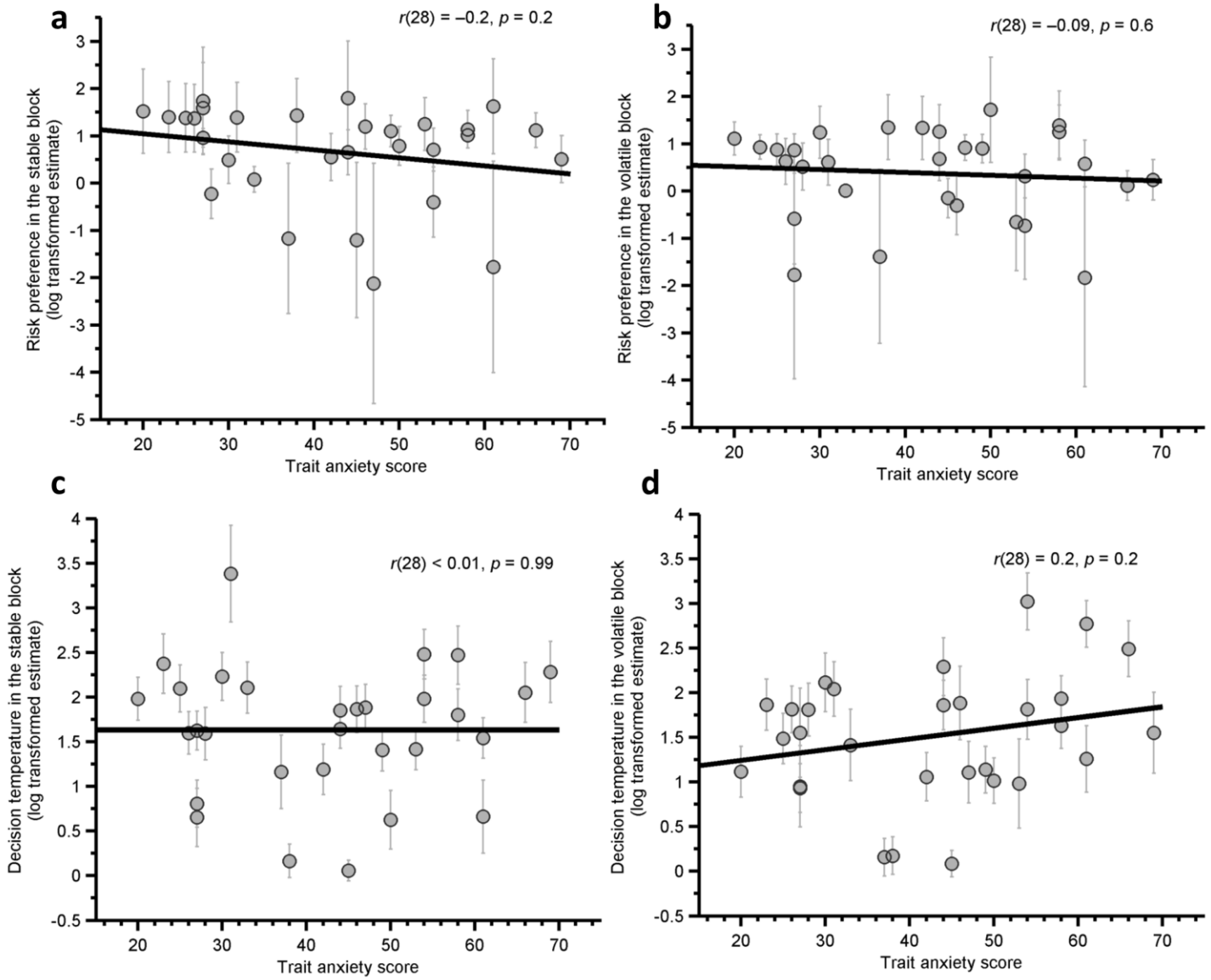
As described in the Supplementary Modeling Note, a more complex model, designed to capture a variety of potential influences on learning rate, was additionally fitted to choice data at the level of the individual participant. On each trial, learning rate was estimated as a weighted sum of three terms: the mean learning rate across trials (α_{mean}), the trial-wise demeaned Bayesian volatility (subject specific weight =ω_{Bayes}) and the trial-wise demeaned exponential decay function (subject specific weight =ω_{Exp}). These parameters along with the slope of the exponential function(λ) and the decision temperature and risk preference terms from the selector model were fitted to participant behavior across the entire task. The resulting parameter estimates were then correlated, across participants, against trait anxiety. Of the six parameters fitted to participants' behavior, only the weight of the volatility term (ω_{Bayes}) was significantly correlated with trait anxiety, illustrated above $r(28) = -0.37$, $p = 0.04$, with none of the other parameters showing a significant relationship with trait anxiety, ps>0.2. Consistent with the findings of the main paper, this result indicates that, even when a variety of factors were allowed to influence a dynamic learning rate, trait anxiety was uniquely associated with reduced influence of environmental volatility upon learning rate. Error bars represent the standard deviation of the parameter estimates for each subject.



Supplementary Figure 5

Shock receipt on difficult trials (where the two options were close in expected value) was positively correlated with trait anxiety (**a**), and negatively correlated with both change in learning rate between stable and volatile blocks (**b**) and modulation of post-outcome pupil dilation by trial volatility (**c**).

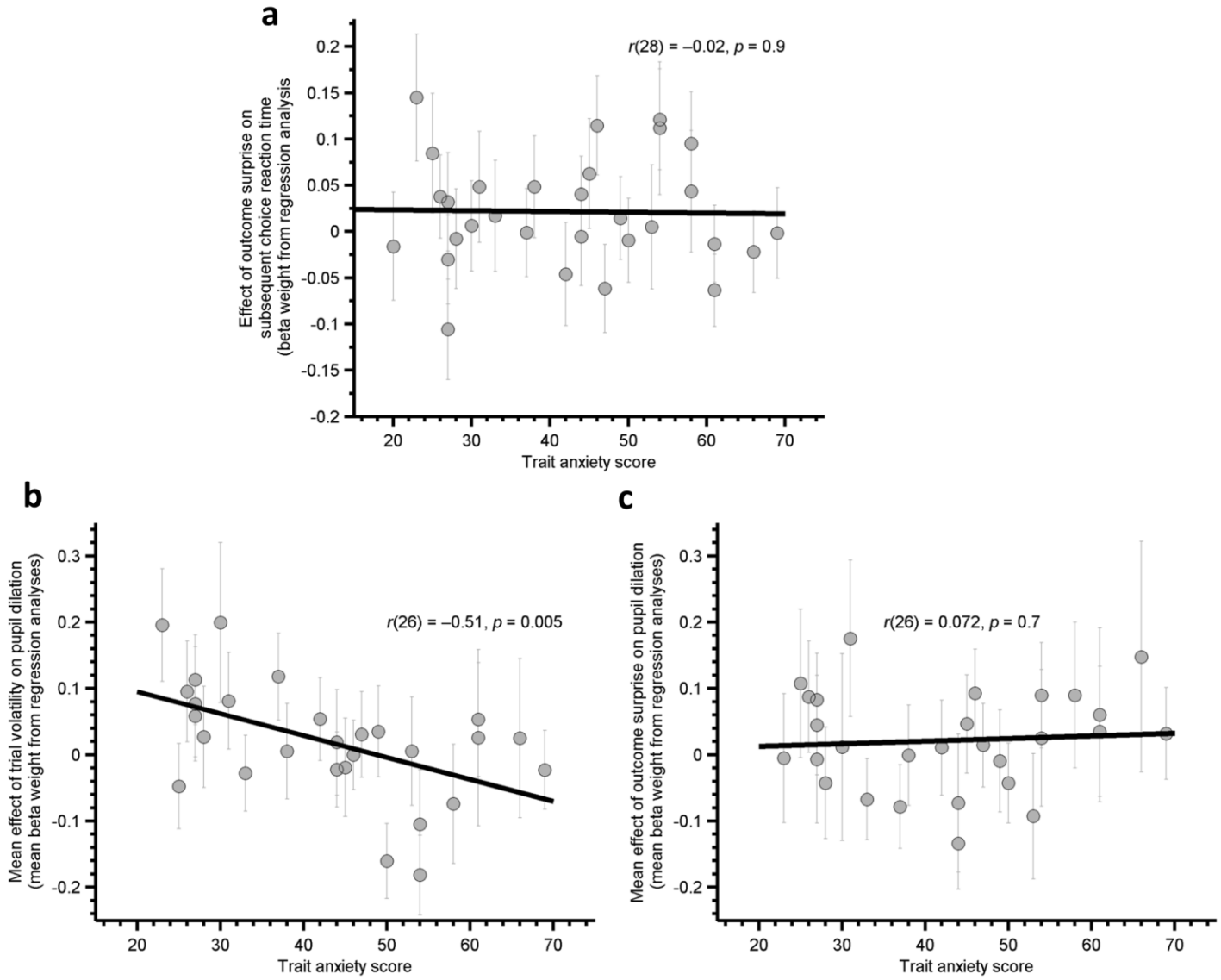
The advantage of flexible adaptation of learning rate to changes in environmental volatility is that it allows learning to occur at an optimal rate for a given environment. As subjects with higher levels of trait anxiety were less able to adjust their learning between stable and volatile blocks, they would be predicted to more often incorrectly judge the action most likely to result in shock. This in turn would be most likely to have an observable effect on trials where the two options were close in expected value. Following Behrens et al[8], we defined 'difficult trials' as those in which the Pascalian value of the two shapes differed by less than 5. (**a**) Trait anxiety was positively correlated with the proportion of 'difficult' trials on which shock was received. (**b**) We further confirmed that this 'proportion of difficult trials where shock was received' index of performance was negatively related to adjustment of learning rate between stable and volatile task blocks, $r(28) = –0.31$, $p = 0.05$, 1-tailed (**b**), as well as to the pupil response to trial volatility, $r(26) = –0.41$, $p = 0.01$, 1-tailed (**c**).



Supplementary Figure 6

No significant association was observed between Trait Anxiety and Risk Preference or Inverse Decision Temperature.

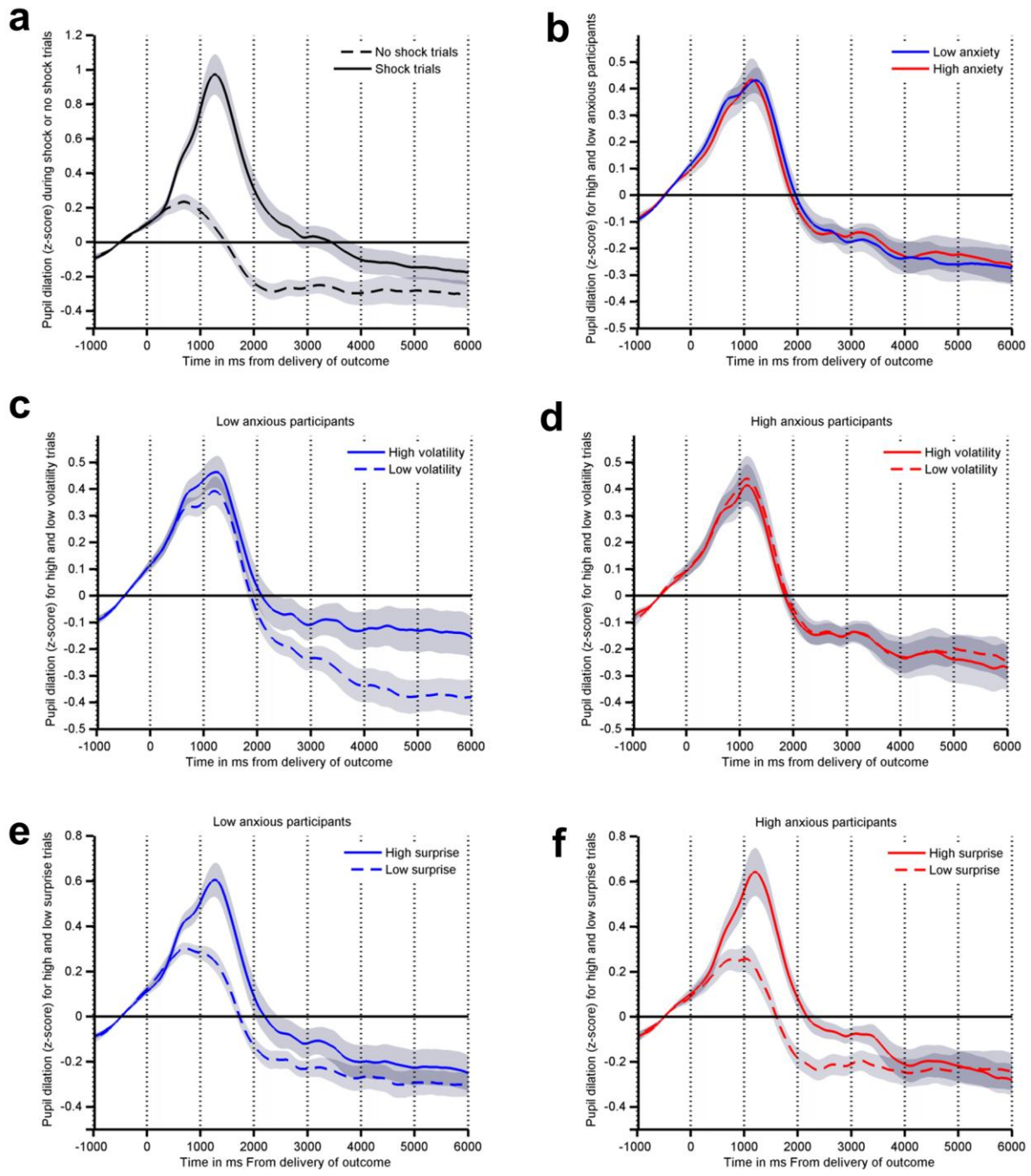
The softmax action selector contained two free parameters: risk preference (the tendency to minimize shock probability versus shock magnitude) and inverse decision temperature (the degree to which a participant used expected values, i.e. estimated shock probability x shock magnitude, for each shape to guide choices). These parameters were estimated from participants' choices in the stable and volatile blocks of the aversive learning task (see Online Methods). Panels **a**-**d** show the relationship between the log transformed parameter estimates and trait anxiety. There was no significant relationship between trait anxiety and risk preference as calculated for either the stable (**a**; $r(28) = –0.23$, $p = 0.2$) or volatile (**b**; $r(28) = –0.09$, $p = 0.6$) task blocks. Trait anxiety was also not associated with the mean value of this parameter across blocks or with the difference score between blocks (ps > .1). There was also no relationship between trait anxiety and inverse decision temperature as estimated for the stable (**c**; $r(28) < 0.01$, $p = 0.99$) or volatile (**d**; $r(28) = 0.24$, $p = 0.2$) blocks. Trait anxiety was also not associated with mean decision temperature across blocks, or with the difference in decision temperature between blocks (*p*s > .1). Error bars represent the standard deviation of the parameter estimates for each subject.



Supplementary Figure 7

Analyses of reaction time and pupil dilation data using non-Bayesian estimates of trial volatility and outcome surprise

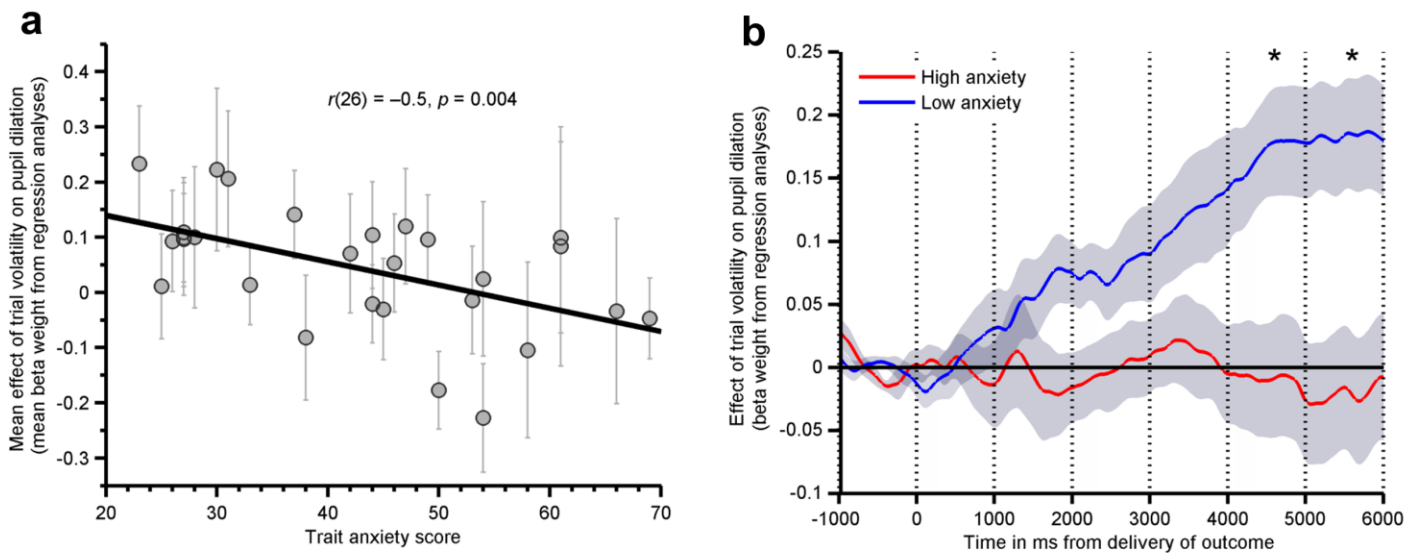
Analyses of the pupil dilation and reaction time data reported in this paper utilized estimates of trial volatility and outcome surprise derived from a Bayesian learner (see Supplementary Modeling Note). Here, we replicated these analyses using non-Bayesian estimates of volatility and surprise. Volatility was simply coded as '1' for all trials in the volatile block and '0' for all trials in the stable block. In addition, a non-Bayesian surprise regressor was created by coding all trials in which the less predictive stimulus was associated with the shock as being high surprise ('1') and the other trials as low surprise ('0'). These measures of volatility and surprise were entered into the regression models predicting reaction time slowing (**a**) and pupil response (**b,c**) together with the other, control, predictor variables previously used. As in the main analyses, the beta estimates for these regressors were correlated against trait anxiety. The results of these additional analyses replicated those reported in the paper: trait anxiety did not influence the degree to which participants slowed their response following a surprising outcome, $r(28) = –0.02$, $p = 0.9$ (**a**). Further trait anxiety was associated with a significantly reduced pupil response during high volatility trials, $r(26) = –0.51$, $p = 0.005$ (**b**), but showed no significant relationship with the pupil response to surprising outcomes, $r(26) = 0.07$, $p = 0.7$ (c). In summary, the relationship reported between trait anxiety and effects of volatility and outcome surprise held regardless of the manner in which these indices were calculated. Error bars represent the standard deviations of the parameter estimates from the behavioral model (**a**) and the regression coefficients (beta weights) from the pupil analysis (**b, c**) for each subject.



Supplementary Figure 8

Baseline corrected, z-transformed eyetracker traces from the aversive learning task**.**

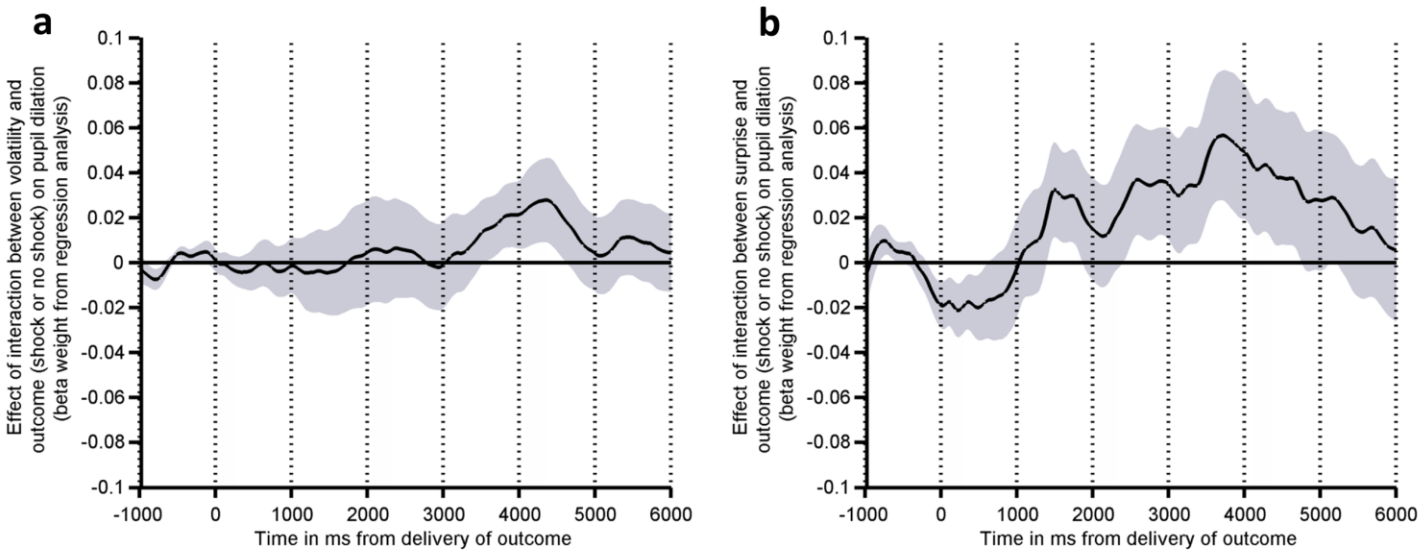
(a). Mean pupil dilation across participants separated into trials in which shock was administered or no shock was administered. **(b).** Mean pupil dilation across trials shown separately for high and low trait anxious participants. Effect of volatility on pupil dilation shown separately for low (**c**) and high (**d**) trait anxious participants. Effect of surprise on pupil dilation shown separately for low (**e**) and high (**f**) trait anxious participants. NB for illustration, participants have been separated into high and low anxious groups based on a median split on trait anxiety scores. Similarly trials were classified as high/low volatility or surprise based on a median split on these variables. Lines show mean value (across participants) of z-transformed, baseline corrected pupil diameter. Shaded regions represent SEM.



Supplementary Figure 9

The relationship between trait anxiety and post outcome pupil dilation as a function of trial-wise estimates of volatility, controlling for trial number within each block.

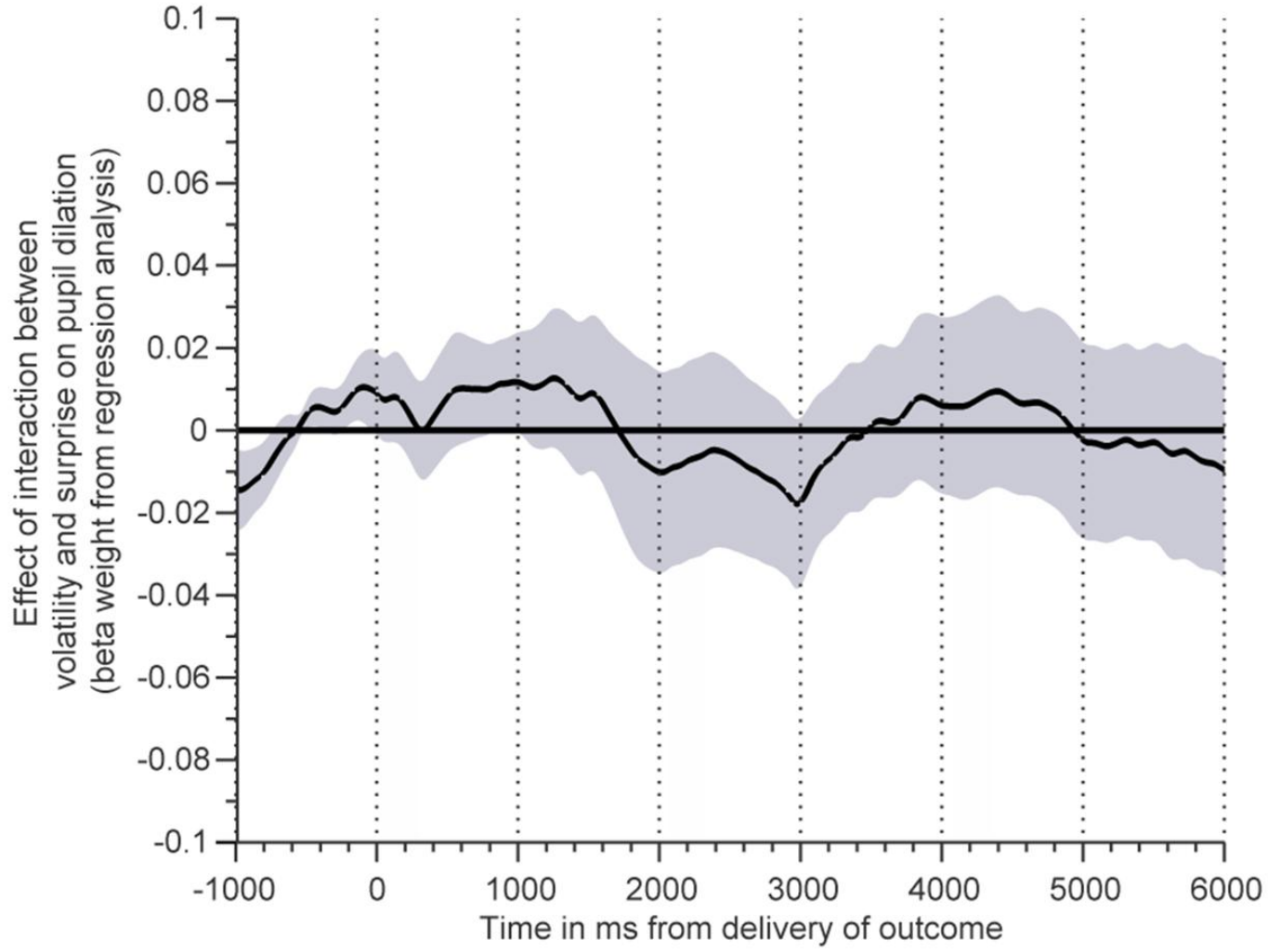
This figure illustrates the same analyses as those described in **Figure 4** of the main manuscript with the exception that regressors coding for the trial number within each task block have been included in the participant-wise regression analyses of the effect of volatility on post outcome pupil dilation. The results illustrated in **Figure 4** remain robust with the inclusion of these extra control parameters. Specifically, a strong negative correlation between trait anxiety and the effect of volatility upon pupil dilation post outcome is still observed (**a**), $r(26) = –0.5$, $p = 0.004$. As before, this is driven by low but not high trait anxious participants showing a modulatory effect of volatility on the post outcome pupil response (**b**), a median split on trait anxiety levels is used here as in **Figure 4**. Asterisks indicate 1s time bins in which Bonferonni corrected t-tests differed between the groups at $p < .05$ corrected, 2-tailed. Error bars in panel **a** represent the standard deviations of the regression coefficients (beta weights) from the pupil analysis for each subject. Shaded regions in panel **b** represent the standard error of the mean.



Supplementary Figure 10

Results from additional pupil analyses investigating the interaction between outcome (shock received versus no shock received), and volatility (**a**) and surprise (**b**).

The delivery of a shock causes a large dilation of participants' pupils (**Figure S8a**). In order to determine whether effects of volatility and surprise were modulated by outcome (shock received versus no shock received), we reran the pupil regression analyses descripted in the main manuscript with additional regressors encoding the interaction between volatility and outcome and the interaction between surprise and outcome. Neither interaction term had a significant influence on pupil dilation (t-tests of mean regressor values against 0, $p > 0.1$ for all time bins). Consistent with this, adding these terms to the model reduced the goodness of fit of the model by an adjusted R^2 of 0.0002. Shaded areas represent the standard error of the mean.



Supplementary Figure 11

Interaction of estimated trialwise volatility and surprise on pupil dilation post outcome.

The pupil analysis was rerun with the inclusion of a surprise x volatility interaction term. There was no significant interaction of surprise x volatility upon pupil dilation in any of the 1s post-outcome time bins (t-test of mean regressor value against 0, $p > 0.3$ for all time bins). While the post-outcome main effect of surprise on pupil dilation precedes that of volatility (**Figure 3**), trialwise volatility changes very slowly (see Supplementary Modeling Note) – i.e. the volatility of a given trial n and that of the following trial n+1 is highly correlated ($r > .9$). Hence this regression analysis should capture any effect of volatility on trial n on the response to surprise in trial n+1. Shaded areas represent standard error of the mean.

Supplementary Modeling note

Decision choice data: estimating change in learning rate.

The primary measure of interest in this paper was change in learning rate between stable and volatile task blocks. This was calculated by fitting a simple learning model[8] to participants' choice data which consisted of a Rescorla-Wagner predictor[10] coupled to a softmax based action selector. The predictor updated its estimate of outcome probabilities using the equation:

$$r_{(i+1)} = r_{(i)} + \alpha\varepsilon_{(i)}$$

Here $r_{(i+1)}$ is the estimated outcome probability for the i +1st trial, $r_{(i)}$ is the estimated outcome probability for the ith trial, α is the learning rate and $\varepsilon_{(i)}$ is the prediction error on the ith trial.

The selector transforms these predictions into action probabilities as follows. First, it estimates the negative value or 'aversiveness' of the two options. This is mathematically equivalent to estimated value for reward versions of the bandit task but reflects the fact that the outcome concerned (shock administration) is negatively valenced. Hence a high probability large magnitude shock is of high 'negative' value. The equation below assumes, for ease of reference, that the circle and square stimuli were used in the task:

$$g_{circle\ (i+1)} = F\left(r_{(i+1)}, \gamma\right) f_{circle(i+1)}$$

$$g_{square\ (i+1)} = F\left(1 - r_{(i+1)}, \gamma\right) f_{square(i+1)}$$

Here $g_{\text{circle}(i+1)}$ and $g_{\text{square}(i+1)}$ are the estimated negative values of the stimuli on the i +1st trial, $f_{circle(i+1)}$ and $f_{square(i+1)}$ are the known shock magnitudes for the two stimuli, and $F(\mathrm{r}, \gamma)$ is a linear transform within the bounds of 0 and 1:

$$F(r, \gamma) = \max\left[\min[(\gamma(r - 0.5) + 0.5), 1], 0\right]$$

The risk preference parameter γ allows the model to place greater weight on outcome magnitude ($\gamma < 1$) or outcome probability ($\gamma > 1$) when calculating the expected value. Effectively, this allows

Nature Neuroscience: doi:10.1038/nn.3961

the model to flexibly capture the extent to which a given participant prefers to minimize the probability or the magnitude of potential shocks.

Given, as outlined above, that the estimated values are negative (the probability x magnitude of an aversive outcome), the final action probabilities were generated using the following probability distribution:

$$P_{(choice=circle)} = \frac{1}{1 + \exp^{(-\beta(g_{(square)} - g_{(circle)})}}$$

Here β is the inverse decision temperature and controls the degree to which the expected values are used in determining the shape chosen.

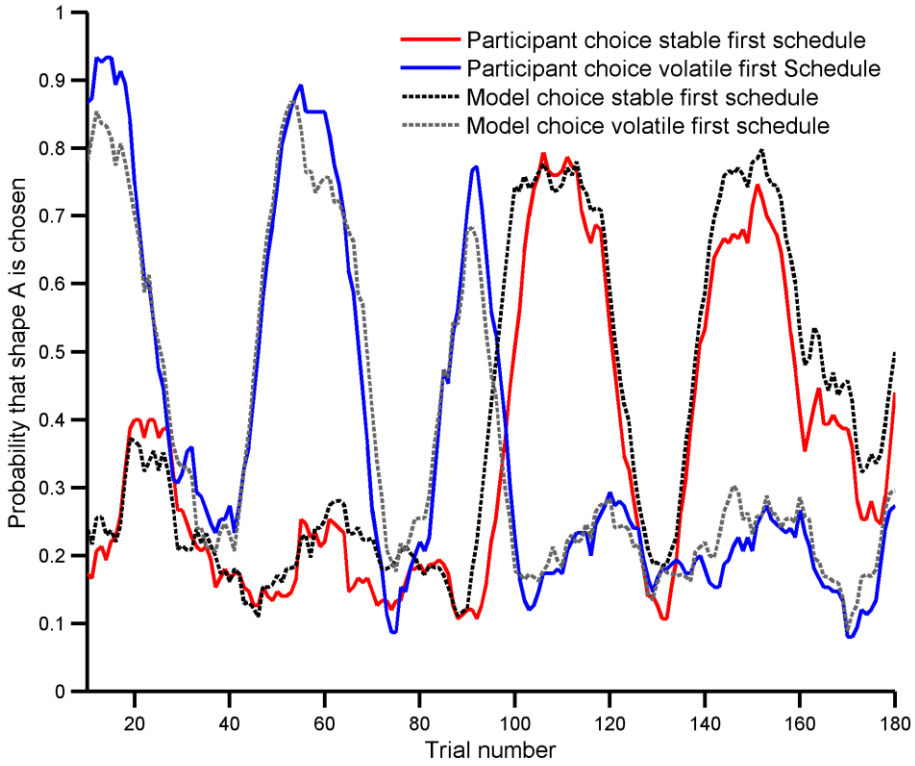
In total, the combined predictor and selector model has 3 free parameters (α, β, γ). These parameters were estimated separately for the stable and volatile blocks, for each participant, as in prior modeling of the structurally equivalent reward task[8]. Parameter estimates were obtained by calculating a joint posterior probability density function over all three parameter values, given the choice behavior of the participant. The values of the parameters were discretized in the joint function using the following limits: learning rate was represented using 30 points spaced equally from log (0.01) to log(1), risk preference using 15 points ranging from log(0.1) to log(10) and inverse temperature using 20 points ranging from log(1) to log(100). Individual parameter estimates (and the standard deviation (SD) of those estimates) were then defined as the expected value (and SD) of the marginal probability density function over the given parameter, obtained by direct integration[8]. We note that the SD of individual level parameter estimates is represented by error bars on the scatter plots in **Figures 2-4** and **Supplementary Figures S2, S6 and S9**. For completeness, we examined whether learning rates for the volatile and stable blocks differed in their estimation noise (i.e. the SD of the parameter estimates), and whether this varied as a function of trait anxiety. Of note, unlike the estimated parameters themselves, the SDs of the estimates are highly positively skewed requiring non-parametric statistical tests to be adopted in these analyses. Using a related

samples sign test, across participants, there was no significant difference in the standard deviation of the learning rate parameter estimates for the stable versus volatile blocks (p=0.1). Further, there was no relationship between trait anxiety and the standard deviation of the learning rate parameter estimate for either the stable (Spearman's rho = 0.01, p = 0.9) or volatile (Spearman's rho = 0.03, p = 0.9) task blocks.

When fitting the model parameters to participant choice, the first 10 trials of each block were omitted to allow comparison with the eyetracking data (here, initial trials were excluded to avoid luminance related changes in pupil size at the beginning of the blocks[14]). Additionally, trials in which no response was made were excluded (mean = 0.5% of trials). As the three free parameters are multiplicative (i.e. they act to multiply value or probability terms in the equations presented above), their logarithms were used when testing their relationship, across participants, with trait anxiety. Similarly change in the parameters between blocks was calculated as the difference of their log values.

An illustration of the ability of the model to recapitulate participant choice is provided in the figure below, which presents learning curves derived from participant choice on top of those derived from the model.

Nature Neuroscience: doi:10.1038/nn.3961



Learning Curves Illustrating Participant and Model Choice During the Learning Task. The plot shown here overlays the choice behavior of participants who completed the "stable first" task schedule (red line) or the "volatile first" task schedule (blue line) with the output (i.e. probability of choosing shape A) of the behavioral (Rescorla Wagner & Softmax Action Selector) model (dotted grey/black lines). In order to more clearly illustrate the effects of learning on participant choice and model output, the effect of shock magnitude (which varies randomly across trials) has been reduced by smoothing both participant choices and model predictions using a running average of 10 trials (i.e. this reduces the trial by trial variations attributable to changes in shock magnitude while preserving the lower frequency effects of learning). The model output was produced by running the behavioral model once for each participant, using the individually fitted parameter values, on the task trial sequences. As can be seen the model closely describes participant behavior during the task for both schedules.

To more formally test the model's predictions of participant behavior, we discretized the model's predictions (i.e., on a given trial, if it gave >50% likelihood that a participant would choose option A this was labeled as correct prediction of choice if the participant choose A). This inevitably leads to a relatively coarse measure of model performance, as trials where the model gives 90% likelihood of option A being chosen and 51% likelihood of option A being chosen are treated the same. Nevertheless, this binarized index still demonstrated that the model correctly predicted 81% of

Nature Neuroscience: doi:10.1038/nn.3961

choices across all trials and 76% of choices on 'difficult trials' (defined as in Behrens et al.[8] as those where the expected value of the two options differed by less than 5). Additional comparison of this model against alternatives (two simplified models using only information about outcome probability not magnitude, and vice versa, and one keeping β and γ fixed across blocks) are given in the table below.

Model	Number of parameters	Negative Log Likelihood (lower is better)	AIC (lower is better)	Pseudo-R^2 (higher is better)	Adjusted Pseudo-R^2 (higher is better)
A					
(i) Full Model as used in main manuscript. (Uses prior outcomes and shock magnitude to guide choice)	6 (α,β,γ per block)	2217.6	4795.3	0.382	0.332
(ii) Prior outcome history model. (Uses only previous outcomes to guide choice)	4 (α, β per block)	2520.3	5280.7	0.300	0.267
(iii) Outcome value model (Uses only shock magnitudes associated with the two shapes to guide choice)	2 (β per block)	3516.8	7153.7	0.016	-0.0007
B					
Full Model with Single Risk Preference (γ) and Inverse Temperature (β) across both blocks	4 (α per block, single β,γ across blocks)	2305.9	4851.8	0.36	0.326

Model fits to participants' choice behavior. (A). The full behavioral model (i), as used for the analyses reported in the main manuscript and in prior published studies of reward-learning[8], assumes that participants combine information about outcome probability and magnitude when making their choices. An alternate possibility is that subjects predominantly used only one of these two sources of information. Hence, we examined the fit to participants' behavior of the full model (i) relative to two simplified version of this model: (ii) a model which ignores shock magnitude information and (iii) a model which uses shock magnitude information but does not adapt its expected outcome probability on the basis of previous outcomes. We report two measures of model fit (negative log likelihood and the Pseudo-R^2 measure) as well as model fit statistics which penalize additional parameters (the Akaike Information Criteria, AIC, and the adjusted Pseudo-R^2 measure). As can be seen, the full model outperforms both the more limited models on all measures. (B) The

same measures of model fit are given for a simplified version of the full model which uses 2 learning rate parameters (one for the stable task block and one for the volatile task block) but only one risk preference and one inverse temperature parameter across the whole task. This simpler model performs marginally less well than the full model reported in the main manuscript (even when penalizing the additional parameters in the full model) and also has the disadvantage that it assumes that all of the effect of block volatility loads onto the learning rate parameter (rather than the risk preference or inverse temperature parameters). The full model allows these parameters to vary between blocks and thus does not rely on this assumption, also allowing for empirical examination of the relationship between anxiety and each of these parameter estimates. Note. The Pseudo-R^2 and adjusted Pseudo-R^2 measures were calculated as recommended by McFadden (McFadden, D., 1974, "Conditional Logit Analysis of Qualitative Choice Behavior" in Frontiers in Econometrics, ed. Zarembka, P., 105–142, Academic Press). α= learning rate; γ = risk preference, β = inverse decision temperature.

The Bayesian learner.

The Bayesian Learner has previously been described in detail by Behrens and colleagues[8]. Briefly, the learner estimates the likelihood of shock delivery following a given choice on trial i+1 given the outcome of the choice made on the current trial (i) and the previously estimated probability of that outcome. It takes advantage of the Markovian assumption that the full history of estimated outcome probabilities does not need to be used, but that this can be approximated by using the estimate of outcome probability from trial i alone. Under this assumption, the probability that a given choice (and by subtraction from 1, the probability for the other choice) will result in shock delivery on the next trial, i.e (r_{i+1}), is represented using a beta distribution with mean r_i and a width parameter *V*. *V* is defined as exp(v), where v is the estimated volatility of the environment.

$$p(r_{i+1}|r_i, v) \backsim \beta(r_i, V)$$

In other words, the estimated volatility v is the log of the variance parameter (V) of this beta distribution. A second parameter used is k, which reflects uncertainty in the current estimate of environmental volatility – this is high in environments that rapidly transit between stable and volatile periods. Formally, *K* = exp(*k*) and is the width parameter of a normal distribution, centered on v_i, which represents the probability distribution of volatility on the following trial $(i+1)$:

$$p(v_{i+1}|v_i, k) \backsim N(v_i, K)$$

Nature Neuroscience: doi:10.1038/nn.3961

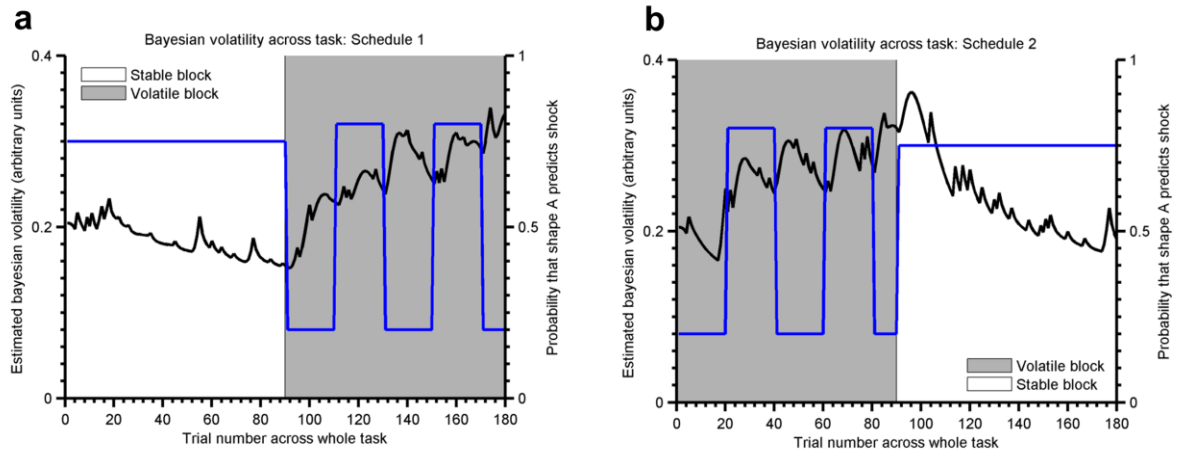
The joint probability of r_{i+1}, v_{i+1} and k is estimated following the outcome (y) of each trial as:

$$p(r_{i+1}, v_{i+1}, k|y_{\leq i+1}) \propto p(y_{i+1}|r_{i+1}) \int \left[\iint p(r_i, v_i, k|y_{\leq i}) p(v_{i+1}|v_i, k) dv_i \right] p(r_{i+1}|r_i, v_{i+1}) dr_i$$

Trialwise estimates of the individual parameters (r_{i+1}, $v_{i+1,}$ k) are obtained by marginalization of the joint probability function. In addition, trial-wise estimates of surprise are calculated by taking the negative logarithm of the conditional probability of the outcome on trial i, given the expectation of the Bayesian learner[22,] i.e:

$$surprise_i = -\log\left(p(y_i|E(r_i))\right)$$

The figure below illustrates the trialwise estimates of volatility obtained from the Bayesian Learner across the aversive learning task.



Environmental volatility v_i as estimated by the Bayesian learner on a trial-wise basis across the stable and volatile task blocks. Panels (a) and (b) depict the estimates that correspond to presentation of the stable block prior to, or following, the volatile block, respectively. Estimated volatility (black lines, left hand axes), can be seen to change gradually across each block of 90 trials as the learner accrues information about the (in)stability of task contingencies. This gradual accrual of information means that the initial trials of each block provide least information about environmental volatility, and a large number of trials are needed to assess the extent to which participants can alter their learning rate in response to volatility. The underlying outcome probabilities of the task are included as blue lines (right hand axes).

Nature Neuroscience: doi:10.1038/nn.3961

Modelling alternative potential influences on learning rate

As described above, participants' learning was estimated within each task block using a Rescorla-Wagner predictor[10] coupled to a softmax based action selector. This enabled us to estimate changes in learning rate between the volatile and stable blocks of the task. There are, however, other potential influences on learning rate that are not captured by this model and which may conceivably be influenced by trait anxiety. In particular, learning rates often start high and decay exponentially during the course of learning (see, for example, Schiller, D. et al., 2008, "From fear to safety and back: reversal of fear in the human brain." *The Journal of neuroscience:* 28, 11517-11525). We therefore tested whether, using a more complex learning model in which a variety of factors compete for influence over learning rate, trait anxiety would be specifically associated with volatility related effects on participants' learning rate.

The learning rate (α) on trial (i) of the alternative model was calculated as:

$$\alpha_{(i)} = \alpha_{mean} + \omega_{Bayes}(v_i - \bar{v}) + \omega_{Exp}(e^{-\lambda i} - \overline{e^{-\lambda}})$$

Here v_i is the Bayesian volatility on trial i, $\bar{v}$ represents the expected (mean) value of volatility across the task, $e^{-\lambda i}$ is a decaying exponential function and $\overline{e^{-\lambda}}$ is the expected value of this function. The learning rate on a given trial is a weighted sum of three terms: the mean learning rate across trials (α_{mean}), the trial-wise demeaned Bayesian volatility (subject specific weight = ω_{Bayes}) and the trial-wise demeaned exponential decay function (subject specific weight = ω_{Exp}). These parameters along with the slope of the exponential function(λ) and the decision temperature and risk preference terms from the selector model were fitted to participant behavior across the entire task (parameter fits were calculated by the same Bayesian procedure as used in the main study analysis). The results of this analysis are presented in **Figure S4**.

Nature Neuroscience: doi:10.1038/nn.3961