A model for learning based on the joint estimation of stochasticity and volatility

Payam Piray* and Nathaniel D. Daw
Princeton Neuroscience Institute and Department of Psychology, Princeton University
*Corresponding author: ppiray@princeton.edu

Abstract

Previous research has stressed the importance of uncertainty for controlling the speed of learning, and of how such control depends on the learner inferring the noise properties of the environment, especially volatility: the speed of change. However, learning rates are jointly determined by the comparison between volatility and a second factor, moment-to-moment *stochasticity*. Yet much previous research has focused on simplified cases corresponding to estimation of either factor alone. Here, we introduce a new learning model, in which both factors are learned simultaneously from experience. By considering the full problem of joint estimation, we highlight a set of previously unappreciated issues, arising from the mutual interdependence of inference about volatility and stochasticity. This interdependence both complicates and enriches the interpretation of previous results, such as those thought to implicate volatility in psychiatric conditions. This provides a novel mechanism for understanding pathological learning in amygdala damage and anxiety disorders.

Introduction

Among the successes of computational neuroscience is a level-spanning account of learning and conditioning, which has grounded biological plasticity mechanisms (specifically, error-driven updating) in terms of a normative analysis of the problem faced by the organism[1–5]. These models recast learning as statistical inference: using experience to estimate the amount of some outcome (e.g., food) expected on average following some cue or action. This is an important subproblem of reinforcement learning, which uses such value estimates to guide choice. The statistical framing has motivated an influential program of investigating the brain's mechanisms for tracking uncertainty about its beliefs, and how these impact learning[6–10].

Uncertainty, in turn, depends upon the noise properties of the values being learned, including both the degree of stochasticity in their measurement (observation noise, the variance of which we call *stochasticity*) and how quickly or how often they change (process noise, the variance of which is known as *volatility*). In general, then, a statistically efficient learner must estimate not just the primary quantity of interest (e.g., the action value) but also parameters describing its noise properties. This perspective has inspired a series of hierarchical Bayesian inference models, which extend inference to either volatility or stochasticity, though typically while treating the other noise parameter as fixed and known to the model.

For volatility, a particularly influential series of theories extends a baseline model known as the Kalman filter to incorporate volatility estimation (but conditional on known stochasticity)[6,11,12]. All else equal, when volatility is higher, the organism is more uncertain about the cue's value (because the true value will on average have fluctuated more following each observation), and so the learning rate (the reliance on each new outcome) should be higher. A series of experiments have reported behavioral and neural signatures of these effects of volatility enhancing learning rate, and also their disruption in relation to psychiatric symptoms[6,8,10,13–23]. Conversely, stochasticity also affects the learning rate, but in the opposite direction: all else equal, when individual outcomes are more stochastic (larger stochasticity), they are *less* informative about the cue's true value and the learning rate, in turn, should be *smaller*. Here again, experiments confirm that people adjust their learning rates in the predicted direction[24,25], and this behavior has been captured by a model that estimates stochasticity (but treating the process noise parameter, in this case a hazard rate, as known).

Altogether, this work has led to a strong argument that the brain's mechanisms for tracking uncertainty, and the inference of the noise parameters that govern it, are crucial to healthy and disordered learning[26]. Although these components seem individually well understood, in this article we argue that important insights are revealed by considering in greater detail the full problem facing the learner: simultaneously estimating both volatility and stochasticity during learning. By introducing a model that performs such joint estimation, and studying its behavior in reinforcement learning tasks, we show that because of the interrelationship between these variables, a full account of any of them interacts with the others in consequential ways.

The key issue is that although the learner's estimates of them play opposite roles on learning rates, from the learner's perspective, they both similarly manifest in noisier, less reliably predictable outcomes. The observation that experienced noise can, to a first approximation, be explained by either volatility or stochasticity – and that these effects might be confused, either by experimenters or by learners – has novel implications. First, previous work apparently showing variation in volatility processing in different groups, such as patients (using a model and tasks that do not vary stochasticity), might instead reflect misidentified abnormalities in processing stochasticity. We suggest that future research should test both dimensions of learning explicitly. Furthermore, from the perspective of a learner inferring volatility and stochasticity, these factors should compete or trade off against one another to best explain observed noise. This means that any dysfunction or damage that impairs detection of stochasticity, should lead to a compensatory increase in inferred volatility, and vice versa: a classic pattern known in Bayesian inference as a failure of "explaining away."

We argue that such compensatory tradeoffs may be apparent both in anxiety disorders and following damage to amygdala. Intolerance of uncertainty is thought to be a critical component of anxiety and a crucial risk factor for developing anxiety disorders[27,28]. Although there has been recent interest in operationalizing this idea by connecting it to statistical learning models and tasks[26,29–36], we and others have focused on apparent abnormalities in processing volatility[13,20,26]. The current model suggests a different interpretation, in which anxiety primarily disrupts inference about stochasticity, but with the additional result that the learner misinterprets noise due to stochasticity as a signal of change, i.e. volatility. We argue that the complementary pattern of explaining away, in which a failure to detect volatility leads to change misattributed to stochasticity, can be appreciated in studies of the amygdala's role in modulating learning rates. In particular, our model suggests that a specific involvement of amygdala in volatility (and the explaining away pattern) explains effects of amygdala damage better than an involvement in learning rates more generally. These sorts of reciprocal interactions also give rise to a richer and subtler set of possible patterns of dysfunction that may help to understand a wide range of other neurological and psychiatric disorders, such as schizophrenia, in which there has been a tendency to study altered processing of uncertainty narrowly in the context of volatility.

The model also sheds new light on experimental phenomena of learning rates in conditioning, and on two classic descriptive theories of conditioning in psychology that have been interpreted as predecessors of the statistical accounts[37,38]. We suggest that seemingly contradictory effects of noise in different experiments, associated with these two theories, can be identified with effects on learning rates of inferred stochasticity vs. volatility, respectively. On this view, different effects will dominate in different experiments depending on which parameter the pattern of the noise suggests.

In the reminder of this article, we present i) a probabilistic model for the joint estimation of volatility and stochasticity from experience; and ii) volatility- and stochasticity- lesioned models in which the corresponding module is damaged. These models highlight the mutual interdependence of inference about volatility and stochasticity and show how the interdependence leads the model to predict paradoxical compensatory behaviors if inference about either factor is damaged. This leads to a rich set

of models for explaining psychiatric and neurological disorders. We use these lesioned models in a series of simulation experiment to explain aspects of pathological behavior observed in anxiety disorders and amygdala damage that are, respectively, explained by the stochasticity and volatility lesion models.

Results

Model

We begin with the Kalman filter, which describes statistically optimal learning from data produced according to a specific form of noisy generation process. The model assumes that the agent must draw inferences (e.g., about true reward rates) from observations (individual reward amounts) that are corrupted by two distinct sources of noise: process noise or *volatility* and outcome noise or *stochasticity* (Figure 1ab). Volatility captures the speed by which the true value being estimated changes from trial to trial (modeled as Gaussian diffusion); stochasticity describes additional measurement noise in the observation of each outcome around its true value (modeled as Gaussian noise on each trial).

For this data generating process, if the true values (i.e., variances) of volatility and stochasticity, v_t and s_t are known, then optimal inference about the underlying reward rate is tractable using a specific application of Bayes rule, here called the Kalman filter[39]. The Kalman filter represents its beliefs about the reward rate at each step as a Gaussian distribution with a mean, m_t, and variance (i.e., uncertainty about the true value), w_t. The update, on every trial, is driven by a prediction error signal, δ_t, and learning rate, α_t. This leads to simple update rules following observation of outcome o_t:

$$\delta_t = o_t - m_t \tag{1}$$

$$\alpha_t = \frac{w_t + v_t}{w_t + v_t + s_t} \tag{2}$$

$$m_{t+1} = m_t + \alpha_t \delta_t \tag{3}$$

$$w_{t+1} = (1 - \alpha_t)(w_t + v_t) \tag{4}$$

This derivation thus provides a rationale for the error-driven update (Eq. 3) prominent in neuroscience and psychology[40], and adds to these a principled account of the learning rate, α_t, which on this view should depend (Eq. 2) on the agent's uncertainty and the noise characteristics of the environment. In particular, Eq. 2 shows that the learning rate is increasing and decreasing, respectively, with volatility and stochasticity. This is because higher volatility increases the chance that the true value will have changed since last observed (increasing the need to rely on the new observation), but higher stochasticity decreases the informativeness of the new observation relative to previous beliefs.

This observation launched a line of research focused on elucidating and testing the prediction that organisms adopt higher learning rates when volatility is higher[6]. But the very premise of these experiments violates the simplifying assumption of the Kalman filter – that volatility is fixed and known to the agent. To handle this situation, new models were developed[11,12] that generalize the Kalman filter to

incorporate learning the volatility v_t as well, arising from Bayesian inference in a hierarchical generative model in which the true v_t is also changing. In this case, exact inference is no longer tractable, but approximate inference is possible and typically incorporates Eqs. 1-4 as a subprocess.

This line of work on volatility estimation inherited from the Kalman filter the view of stochasticity as fixed and known. Of course, in general, all the same considerations apply to stochasticity as well: it must be learned from experience, may be changing, and its value impacts learning rate. Indeed, estimating both parameters is critical for efficient learning, because they have opposite effects on learning rate: whereas volatility increases the learning rate, stochasticity reduces it (Figure 1c). Although some algorithms have been explored for learning when both noise parameters are unknown[41–43], the main other application of this type of model in neuroscience has relied on a different simplification[25], which estimates the stochasticity while treating the hazard rate (the analogue of volatility for changepoint problems) as fixed and known to the model.

Learning these two parameters simultaneously is more difficult because, from the perspective of the agent, larger values of either volatility or stochasticity result in more surprising observations: i.e., larger outcome variance (Figure 1e). However, there is a subtle and critical difference between the effects of these parameters on generated outcomes: whereas larger volatility increases the autocorrelation between outcomes (i.e. covariation between outcomes on consecutive trials), stochasticity reduces the autocorrelation (Figure 1f). This is the key point that makes it possible to dissociate and infer these two terms while only observing outcomes.

We developed a probabilistic model for learning under these circumstances. The data generation process arises from a further hierarchical generalization of these models (specifically the generative model used in our recent work[12]), in which the true value of stochasticity is unknown and changing, as are the true reward rate and volatility (Figure 1d). The goal of the learner is to estimate the true reward rate from observations, which necessitates inferring volatility and stochasticity as well.

As with models of volatility, exact inference for this generative process is intractable. Furthermore, in our experience this problem is also relatively challenging to handle with variational inference, the family of approximate inference techniques used previously (see Discussion). Thus, we have instead used a different standard approximation approach that has also been popular in psychology and neuroscience, Monte Carlo sampling[3,44–46]. In particular, we use particle filtering to track v_t and s_t based on data[47,48]. Our method exploits the fact that given a sample of volatility and stochasticity, inference for the reward rate is tractable and is given by Eqs 1-4, in which s_t and v_t are replaced by their corresponding samples (see Methods; this combination of sequential sampling with exact inference for a subproblem is known as Rao-Blackwellized particle filtering).

Learning under volatility and stochasticity

We now consider the implications of this model for learning under volatility and stochasticity.

A series of studies has used two-level manipulations (high vs low volatility blockwise) to investigate the prediction that learning rates should increase under high volatility[6,13,36,49]. Here volatility has been operationalized by frequent or infrequent reversals (Figure 2a), rather than the smoother Gaussian diffusion that the volatility-augmented Kalman filter models formally assume. Nevertheless, applied to this type of task, these models detect higher volatility in the frequent-reversal blocks, and increase their learning rates accordingly[6,11,12]. The current model (which effectively incorporates the others as a special case) achieves the same blockwise result when stochasticity is held fixed across both blocks (Supplementary Figure 1).

In the preceding line of studies, stochasticity was not manipulated. (Indeed, it was not even independently manipulable because rewards were binary, and the variance of binomial outcomes is determined only by the mean.) However, analogous effects of stochasticity have been seen in another line of studies[7,9,25,50]. In these studies, Nassar and colleagues studied learning rates in a task in which subjects had to predict a value, from observations in which the true value was corrupted, blockwise, by different levels of additive Gaussian noise (i.e., stochasticity) and occasionally "jumping" with a constant hazard rate, analogous to volatility. The main feature of these results relevant to the current model is that these studies have shown that participants' learning rate decreases with increases in the noise level (see also[24]). This effect cannot be explained by models that only consider volatility, and in fact, those models make opposite predictions because they take increased noise as evidence of volatility increase. The current model, however, produces the same blockwise effect as humans: because it correctly infers the change in stochasticity, its learning rate is lower, on average, for higher levels of noise (Supplementary Figure 2). Although we do not intend the current model as a detailed account of how people solve this class of tasks (which is based on a somewhat different generative dynamics), the model can also reproduce other more fine-grained aspects of human behavior in this task, particularly increases in learning rate following switches and scaling of learning rate with the magnitude of error (Supplementary Figure 2).

Note that while considered together, these two lines of studies separately demonstrate the two types of effects on learning rates we stress, neither of these lines of work has manipulated stochasticity alongside volatility (though see also[24]). Furthermore, learning of the noise hyperparameters in these studies has largely been explicitly modeled only for either parameter conditional on the other being known. We next consider a variant of this type of task, elaborated to include a 2x2 factorial manipulation of both stochasticity alongside volatility (Figure 2; we also substitute smooth diffusion for reversals). Here, both parameters are constant within the task, but they are unknown to the model. A series of outcomes was generated based on a Markov random walk in which the hidden reward rate is changing according to a random walk and the learner observes outcomes that are noisily generated according to the reward rate.

Figure 2 shows the model's learning rates and how these follow from its inferences of volatility and stochasticity. As above, the model increases its learning rate in the higher volatility conditions; but as expected it also decreases it in the higher stochasticity conditions (Figure 2a). These effects on learning rate arise, in turn (via Eq. 2) because the model is able to correctly estimate the various combinations of volatility and stochasticity from the data (Figure 2bc).

Our model thus suggests a general program of augmenting the standard 2-level volatility manipulation by crossing it with a second manipulation, of stochasticity, and predicts that higher stochasticity should decrease learning rate, separate from volatility effects.

Interactions between volatility and stochasticity

The previous results highlight an important implication of the current model: that inferences about volatility and stochasticity are mutually interdependent. These interrelationships immediately imply a general interpretational issue for experiments that manipulate only one of these noise parameters, and analyze data using a model that attributes all dynamic learning rate effects to one of them. But the details of the interdependence are themselves informative. From the learner's perspective, a challenging problem (simplified away in many of the previous models) is to distinguish volatility from stochasticity when both are unknown, because both of them increase the noisiness of observations. Disentangling their respective contributions requires trading off two opposing explanations for the pattern of observations, a process known in Bayesian probability theory as *explaining away*. Thus, models that neglect stochasticity tend to misidentify stochasticity as volatility and inappropriately modulate learning.

Intriguingly, this situation might in principle arise in neurological damage and psychiatric disorders, if they selectively impact inference about volatility or stochasticity. In that case, the model predicts a characteristic pattern of compensation, whereby learning rate modulation is not merely impaired but *reversed*, reflecting the substitution of volatility for stochasticity or vice versa: a failure of explaining away. Figure 3 shows this phenomenon in the 2x2 design of Figure 2, with two characteristic lesion models. The key point here is that a lesioned model that does not consider one factor (e.g. stochasticity), inevitably makes systematically incorrect inferences about the other factor too. Importantly, previous models that only consider volatility are analogous to the stochasticity lesion model (Figure 3b) and, therefore, make systematically erroneous inference about volatility (Figure 3g) and misadjust learning rate if stochasticity is changing (Figure 3e). This set of lesioned models provide a rich framework for understanding pathological learning in psychiatric and neurologic disorders. Later we show that stochasticity lesion and volatility lesion models explain deficits in learning observed in anxiety and amygdala damage, respectively. But first, we apply the healthy model to reinterpret some long-standing issues about learning rates in animal conditioning.

Stochasticity vs. volatility in Pavlovian learning

Learning rates and their dependence upon previous experience have also been extensively studied in Pavlovian conditioning. In this respect, a distinction emerges between two seemingly contradictory lines of theory and experiment, those of Mackintosh[37] vs Pearce and Hall[38]. Both of these theories concern how the history of experiences with some cue drives animals to devote more or less "attention" to it. Attention is envisioned to affect several phenomena including not just rates of learning about the cue, but also other aspects of their processing, such as competition between multiple stimuli presented in compound. Here, to most clearly examine the relationship with the research and models discussed above, we focus specifically on learning rates for a single cue.

The two lines of models start with opposing core intuitions. Mackintosh[37] argues that animals should pay more attention to (e.g., learn faster about) cues that have in the past been more reliable predictors of outcomes. Pearce and Hall[38] argue for the opposite: faster learning about cues that have previously been accompanied by surprising outcomes, i.e. those that have been *less* reliably predictive.

Indeed, different experiments – as discussed below – support either view. For our purposes, we can view these experiments as involving two phases: a pretraining phase that manipulates stochasticity or surprise, followed by a retraining phase to test how this affects the speed of subsequent (re)learning. In terms of our model, we can interpret the pretraining phase as establishing inferences about stochasticity and volatility, which then (depending on their balance) govern learning rate during retraining. On this view, noisier pretraining might, depending on the pattern of noise, lead to either higher volatility and higher learning rates (consistent with Pearce-Hall) or higher stochasticity and lower learning rates (consistent with Mackintosh).

First consider volatility. It has been argued that the Pearce-Hall[38] logic is formalized by volatility-learning models[2,3,12]. In these models, surprising outcomes during pretraining increase inferred volatility and thus speed subsequent relearning. Hall and Pearce[38,51] pretrained rats with a tone stimulus predicting a moderate shock. In the retraining phase, the intensity of the shock was increased. Critically, one group of rats experienced a few surprising "omission" trials at the end of the pretraining phase, in which the tone stimulus was presented with no shock. The speed of learning was substantially increased following the omission trials compared with a control group that experienced no omission in pretraining. Figure 4a-d shows a simulation of this experiment from the current model, showing that the omission trials lead to increased volatility and faster learning. Note that the history-dependence of learning rates in this type of experiment also rejects simpler models like the Kalman filter, in which volatility (and stochasticity) are taken as fixed; for the Kalman filter, learning rate depends only on the number of pretraining trials but not the particular pattern of observed outcomes. The response probability of the model thus shows the same pattern as response rate for rats (Supplementary Figure 4).

Next, consider stochasticity. Perhaps the best example of Mackintosh's[37] principle in terms of learning rates for a single cue is the 'partial reinforcement extinction effect'[52–54]. Here, for pretraining, a cue is reinforced either on every trial or instead ('partial reinforcement') on only a fraction of trials (Figure 4e). The number of times that the learner encounters the stimulus is the same for both conditions, but the outcome is noisier for the partially reinforced stimulus. The retraining phase consists of extinction (i.e. fully unreinforced presentations of the cue), which occurs faster for fully reinforced cues even though they had been paired with more reinforcers initially. Our model explains this finding (Figure 4f), because it infers larger stochasticity in the partially reinforced condition, leading to slower learning (Figure 4gh). Notably, this type of finding cannot be explained by models which learn only about volatility[6,11,12]. In general, this class of models mistake partial reinforcement for increased volatility (rather than increased stochasticity), and incorrectly predict faster learning.

Note the subtle difference between the two experiments of Figure 4. The surprising omission experiment involves stable pretraining prior to omission, then an abrupt shift, whereas pretraining in the partial reinforcement experiment is stochastic, but uniformly so. Accordingly, though both pretraining phases involve increased noise (relative to their controls) the model interprets the pattern of this noise as more likely reflecting either volatility or stochasticity, respectively, with opposite effects on learning rate. Overall, then, these experiments support the current model's suggestion that organisms learn about stochasticity in addition to volatility. Conversely, the models help to clarify and reconcile the seemingly opposing theory and experiments of Mackintosh and Pearce-Hall, at least with respect to learning rates for individual cues. Indeed, although previous work has noted the relationship between Pearce-Hall surprise, uncertainty, and learning rates[2,3,6,12,20,55], the current modeling significantly clarifies this mapping by identifying it more specifically with volatility, as contrasted against simultaneous inference about stochasticity. Meanwhile, while our basic statistical interpretation of the partial reversal extinction effect has been noted before (e.g. by Gallistel and Gibbon[56]), to our knowledge these previous explanations have not reconciled it with the volatility / Pearce-Hall phenomena. Instead, previous work attempting to map Mackintosh's[37] principle onto statistical models (and distinguish it from Pearce-Hall-like effects) has focused on attention and uncertainty affecting cue combination rather than learning rates[2], which is a complementary but separate idea.

Anxiety and inference about stochasticity vs. volatility

Lesioned models like the ones in Figure 3 are potentially useful for understanding learning deficits in psychiatric disorders, for example anxiety disorders, which have recently been studied in the context of volatility and its effects on learning rate[13,20]. These studies have shown that anxious people are less sensitive to volatility manipulations in probabilistic learning tasks similar to Figure 2. Besides learning rates, an analogous insensitivity to volatility has been observed in pupil dilation[13] and in neural activity in the dorsal anterior cingulate cortex, a region that covaried with learning rate in controls[20].

These results have been interpreted in relation to the more general idea that intolerance of uncertainty is a key foundation of anxiety; accordingly, fully understanding them any require taking account of multiple sources of uncertainty[26], including both volatility and stochasticity. Nevertheless, the primary interpretation of these types of results has been that observed abnormalities are rooted in volatility estimation per se[13,20,26]. Our current model suggests an alternative explanation: that the core underlying deficit is actually with stochasticity, and apparent disturbances in volatility processing are secondary to this, due to their interrelationship.

In particular, these effects and a number of others are well explained by the stochasticity lesion model of Figure 3b, i.e., by assuming that anxious people have a core deficit in estimating stochasticity, and instead treat it as small and constant. As shown in Figure 5b, this model shows insensitivity to volatility manipulation, but in the model that is actually because volatility is misestimated nearer ceiling due to underestimation of stochasticity. This, in turn, substantially dampens further adaptation of learning rate in blocks when volatility actually increases. The elevated learning rate across all blocks leads to hypersensitivity to noise, which prevents anxious individuals from benefitting from stability, as has been

observed empirically[20]. In particular, Piray et al. have studied learning in individuals with low- or high- in trait social anxiety using a switching probabilistic task (Supplementary Figure 6) in which each trial started with a social threatening cue (angry face image). It was found that anxious individuals perform particularly worse than controls in stable trials, whereas their performance is generally matched with controls in volatile trials[20] (Figure 5e). The model shows similar behavior (Figure 5f).

One key prediction of our model, which differs from a volatility-specific account, is that the learning rate is generally higher in anxious people regardless of volatility manipulation or even in tasks that do not manipulate volatility. In fact, Browning et al.[13] do not find evidence to support this prediction: they do not find a significant overall effect of anxiety on learning rate. Of course, it is important not to interpret null results as evidence in favor of the null hypothesis, since a failure to reject the null hypothesis may reflect insufficient power to detect a true effect. Indeed, in Browning's[13] data, while the effect of anxiety on learning rate was not significant overall or in either volatility condition, the point estimate was largest ($r(28)=0.26$, $p=0.16$) in the stable condition, which is also the block in which the model predicts the effect should be statistically strongest (because baseline learning rates, absent any effect of anxiety, are lower).

Importantly, other, larger studies provide positive statistical support for the prediction of elevated learning rate with anxiety[29,31,32]. Note that in delta-rule models, behavior under higher learning rates is closer to win-stay/lose-shift (since higher learning rates weight the most recent outcome more heavily, with full WSLS – dependence only on the most recent outcome – equivalent to a learning rate of 1). Such a strategy has itself been linked to anxiety[31,32]. A notable observation was made in a large (n=122) study by Huang et al.[32], who found anxious people show higher win-stay/lose-shift and this effect is driven by higher lose-shift. Figure 5gh shows results of simulating the proposed model in a task similar to Huang et al.[32] (Supplementary Figure 6). The model shows the same pattern of behavior, with the additional modulation by win vs loss captured because any loss is seen as an evidence for volatility and that results in higher learning rate and a contingency switch. The effect is much less salient for win trials because prediction errors are relatively small in those trials, which substantially dampen any effect of learning rate. Across all trials, the anxious model shows higher learning rate, similar to what Huang et al.[32] found by fitting reinforcement learning models to choice data.

Finally, the lesion model is an extreme case in which a hypothetical stochasticity module is completely eliminated. But this general approach can be extended to less extreme cases in which one module of the model (e.g. stochasticity) has some relative disadvantage in explaining noise. In terms of our model, this can be achieved by having higher update rate parameters for volatility relative to that of stochasticity. These are two main parameters of the model that one can use to explain individual differences across people. For example, the ratio of volatility to stochasticity update rate can be used to capture continuous individual variation in trait anxiety. In this case, the stochasticity lesion model of Figure 3b is an extreme case of this approach in which the stochasticity update rate is zero (thus the ratio of volatility to stochasticity is infinitely large). We have exploited this approach to simulate a result from Browning et al.[13] concerning graded individual differences in anxiety's effect on learning rate adjustment. In particular,

they report (and the model captures; Figure 6) negative correlation between relative learning rate (volatile minus stable) and trait anxiety in the probabilistic switching task of Figure 5a.

Amygdala damage and inference about volatility vs. stochasticity

The opposite pattern of compensatory effects on inference is evidently visible in the effects of amygdala damage on learning. The amygdala plays an important role in associative learning[57,58]. Although some researchers have emphasized a role of the amygdala as a site of association between conditioned- and unconditioned- stimulus in conditioning per se, other authors (drawing on evidence from human neuroimaging work, single-cell recordings, and lesion studies) have proposed that the amygdala is involved in a circuit for controlling or adjusting learning rates[55,58–62]. Most informative from the perspective of our model are lesion studies in rats[59,63,64],which we interpret as supportive an involvement specifically in processing of volatility, rather than learning rates or uncertainty more generally. These experiments examine a surprise-induced upshift in learning rate similar to the Pearce-Hall experiment from Figure 4. Lesions to the central nucleus of the amygdala attenuate this effect, suggesting a role in volatility processing. But an important detail of these results with respect to our model's predictions is that the effect is not merely attenuated but reversed. This reciprocal effect supports perhaps the most central feature and prediction of our model: that volatility trades off against a (presumably anatomically separate) system for stochasticity estimation.

Figure 7 shows their serial prediction task and results in more detail. Rats performed a prediction task in two phases. A group of rats in the 'consistent' condition performed the same prediction task in both phases. The 'shift' group, in contrast, experienced a sudden change in the contingency in the second phase. Whereas the control rats showed elevation of learning rate in the shift condition manifested by elevation of food seeking behavior in the very first trial of the test, the amygdala lesioned rats showed the opposite pattern. Lesioned rats showed significantly smaller learning rate in the shift condition compared with the consistent one: a reversal of the surprise-induced upshift.

We simulated the model in this experiment. To model a hypothetical effect of amygdala lesion on volatility inference, we assumed that lesioned rats treat volatility as small and constant. As shown in Figure 7, the model shows an elevated learning rate in the shift condition for the control rats, which is again due to increases in inferred volatility after the contingency shift. For the lesioned model, however, surprise is misattributed to the stochasticity term as an increase in inferred volatility cannot explain away surprising observations (because it was held fixed). Therefore, the contingency shift inevitably increases stochasticity and thereby decreases the learning rate. Notably, the compensatory reversal in this experiment cannot be explained using models that do not consider both volatility and stochasticity terms.

A similar pattern of effects of amygdala lesions, consistent with our theory, is seen in an experiment on non-human primates. In a recent report by Costa et al.[61], it has been found that amygdala lesions in monkeys disrupt reversal learning with deterministic contingencies, moreso than a reversal task with stochastic contingencies. This is striking since deterministic reversal learning tasks are much easier. Similar to the previous experiment, our model explains this finding because large surprises caused by the

contingency reversal are misattributed to the stochasticity in lesioned animals (because volatility was held fixed), while control animals correctly attribute them to the volatility term (Figure 8; see Supplementary Figures 7 for performance of the model and Supplementary Figure 8 for simulation of the model in all probabilistic schedules tested by Costa and colleagues[61]). This effect is particularly large in the deterministic case because the environment is very predictable before the reversal and therefore the reversal causes larger surprises than those for the stochastic one. Similar findings have been found in a study in human subjects with focal bilateral amygdala lesions[65], in which patients tend to show more deficits in deterministic reversal learning than stochastic one. Again, these experimental findings are not explained by a Kalman filter or models that only consider the volatility term.

Overall, then, these experiments support the current model's picture of dueling influences of stochasticity and volatility. Furthermore, the current model helps to clarify the precise role of amygdala in this type of learning, relating it specifically to volatility-mediated adjustments.

Discussion

A central question in decision neuroscience is how the brain learns from the consequences of choices given that these can be highly noisy. To do so effectively requires simultaneously learning about the characteristics of the noise, as has been emphasized most strongly in a prominent line of work on how the brain tracks the volatility of the environment. Here we revisit this problem for the more realistic case when both volatility and a second noise parameter, stochasticity, must be simultaneously estimated.

While various experiments have, mostly separately, shown that humans can adjust learning rates in response to manipulations of either type of noise, models of how they do so have focused primarily on estimating either parameter while taking the other as known. This skirts the more difficult problem of distinguishing types of noise. To solve this problem and investigating its consequences for learning, we built a probabilistic model for learning in uncertain environments that tracks volatility and stochasticity simultaneously. Using this model to simulate a number of experiments across conditioning, psychiatry and lesion studies, we show a consistent theme whereby the interdependence of inference about these two noise parameters gives rise to patterns of effects that could not be appreciated in previous models that considered estimating either type of noise separately.

The importance of dissociating these forms of noise, and some aspects of their interaction, have been noted previously. For instance, Pulcu and Browning[26] emphasize the inadequacy of existing experiments for dissociating volatility vs. stochasticity learning, and raise the possibility that in principle, people might confuse them. In Nassar et al.'s study[25], the volatility-like hazard rate parameter (though viewed from the model's perspective as fixed and known) is fit as a per-subject free parameter construed as an individual difference. The empirical and model-fitting results showcase a dependence of the (inferred) stochasticity parameter upon the (fit/known) hazard rate, consistent with the bidirectional pattern of interdependence we posit. Building on all these ideas, we build and simulate a model to showcase the potential interdependence between these two types of inference across a range of situations. One important caveat, given the range of applications we consider, is that we abstract away details of the many individual

studies to emphasize their parallelism with respect to our main point of interest. Thus, for instance, we neglect valence-dependent modulation of learning which is likely an additional dimension important both in anxiety[20,26,29,36] and in studies of amygdala[61]. Relatedly, as our goal is to showcase the range of situations in which parallel issues may arise, we acknowledge that different explanations may exist for many individual results.

Our work builds most directly on a rich line of theoretical and experimental work on the relationship between volatility and learning rates[6,8,13,15,66–68]. There have been numerous reports of volatility effects on healthy and disordered behavioral and neural responses, often using a two-level manipulation of volatility like that from Figure 2[6,8,10,13–23,36]. Our modeling suggests that it will be informative to drill deeper into these effects by augmenting this task to cross this manipulation with stochasticity so as more clearly to differentiate these two potential contributors[24]. For example, in tasks that manipulate (and models that consider) only volatility, it can be seen from Equations (1-4) that the timeseries of several quantities all covary together, including the estimated volatility v_t, the posterior uncertainty w_t, and the learning rate α_t. It can therefore be difficult in general to distinguish which of these variables is really driving tantalizing neural correlates related to these processes, for instance in amygdala and dorsal anterior cingulate cortex[6,55]. The inclusion of stochasticity (which increases uncertainty but decreases learning rate) would help to drive these apart.

Indeed, another related set of learning tasks considered prediction of continuous outcomes corrupted by stochasticity, i.e. additive Gaussian noise[7,9,25,43,50] which could provide another foundation for factorial manipulations of the sort we propose. Indeed, a number of these studies (complementary to the volatility studies) included multiple levels of stochasticity and showed learning rate effects[7,9,25,50,69]. The models used in these studies have largely used a complementary simplification to the volatility one: they estimate stochasticity, but conditional on a known value for the hazard rate (equivalent to volatility). Interestingly, rather than overall adjustment to noise statistics, these studies more explicitly emphasized the detection of *discrete* changes in the environment and the resulting local adjustments of the learning rate. From a modeling perspective, inference under change at discrete changepoints (occurring at some hazard rate) raises issues quite analogous to change due to more gradual diffusion (with some volatility). Thus, in practice it has been common and effective to apply models for one sort of change to tasks actually involving the other [6,11,12], a substitution also in part licensed by approximate models of changepoint detection that (as with the volatility models and the Kalman filter for continuous change) also reduce learning to error driven updates with a time-varying learning rate [25,70]. Thus, although we build the current work on a generative model with continuous rather than abrupt changepoints, we do not mean this as a substantive claim, as we expect our main substantive points (concerning the inference about noise vs. change hyperparameters) would play out analogously in other variants. In any case, research into the neural substrates of changepoint detection is highly relevant to the change problem conceived in terms of volatility as well (see [10] for a recent review).

The current framework's tendency to elide the distinction between discrete and continuous change (but distinguish both from stochasticity) is also the basis of an important, but subtle, distinction from another

prominent dichotomy previously proposed: that between "expected" and "unexpected" types of uncertainty [71,72]. While it might appear that these categories correspond, respectively, to stochasticity and volatility as we define them, that is not actually the case. Formally, this is because the Dayan and Yu model (in its most detailed form, Dayan and Yu 2003) arises from a Kalman filter augmented with additional discrete changepoints: i.e. both diffusion and jumps. The focus of that work was distinguishing the special effects of surprising jumps ("unexpected uncertainty"), which were hypothesized to recruit a specialized neural interrupt system. Meanwhile, all other uncertainty arising in the baseline Kalman filter (i.e., that from both stochasticity and volatility; the posterior variance w_t in Eq. 4) is lumped together under "expected uncertainty." That said (although we see this as a misreading of the earlier work) our impression is that later authors' use of these terms actually tends to comport more with our distinction than the original definition [26], i.e. to take unexpected and expected uncertainty as synonymous with volatility and stochasticity as we define them. In any case, Yu and Dayan did not consider the problem considered here, of estimating the noise hyperparameters for learning under uncertainty.

The most important feature of our model is the competition it induces between volatility and stochasticity to "explain away" surprising observations. This leads to a predicted signature of the model in cases of lesion or damage affecting inference about either type of noise: disruption causing neglect of one of the terms leads to overestimation of the other term. For example, if a module responsible for volatility learning were disrupted, the model would overestimate stochasticity, because surprising observations that are due to volatility would be misattributed to the stochasticity. This allowed us to revisit the role of amygdala in associative learning and explain some puzzling findings about its contributions to reversal learning.

Similar explanations grounded in the current model may also be relevant to a number of psychiatric disorders. Abnormalities in uncertainty and inference, broadly, have been hypothesized to play a role in numerous disorders, including especially anxiety and schizophrenia. More specifically, abnormalities in volatility-related learning adjustments have been reported in patients or people reporting symptoms of several mental illnesses[13,14,16–23,36]. The current model provides a more detailed framework for better dissecting these effects, though this will ideally require a new generation of experiments manipulating both factors.

In the present work, we have developed these ideas mostly in terms of pathological decision making in anxiety, which is one of the areas where earlier work on volatility estimation has been strongest and where further refinement using our theory seems most promising[30,33,35,73]. We considered an account by which anxious individuals systematically misidentify outcomes occurring due to chance (stochasticity) as instead a signal of change (volatility)[32]. This account offers a contrary interpretation of a pattern of effects that had been taken to indicate that volatility sensitivity is instead deficient in anxiety[26]. Although some null effects in the study of Browning et al.[13] do not support this account, we view it as an overall better account of the pattern of data across several studies[20,29,31,32]. Our account is also broadly consistent with studies suggesting that anxious individuals might feel overwhelmed when faced with uncertainty[34] and fail to make use of long-term statistical regularities[32]. These are also hypothesized to be related to

symptoms of excessive worry[27,28]. Misestimating stochasticity – moreso than volatility – also seems consonant with the idea that anxious individuals tend to fail to discount negative outcomes occurring by chance (i.e., stochasticity) and instead favor alternative explanations like self-blame[74]. This hypothesis is also consistent with the observation that acquisition of fear conditioning tends to be enhanced in anxious individuals[75,76]. Finally, although a simple increase in learning rate seems harder to reconcile with generally slower extinction of Pavlovian fear learning in anxiety[75], this probably reflects the well-known fact that extinction is not simply unlearning of the original associations, but instead is dominated by additional processes[77,78]. This includes in particular statistical inference about latent contexts[5], which is likely to be affected by both stochasticity and volatility in ways that should be explored in future work.

More generally, this modeling approach, which quantifies misattribution of stochasticity to volatility and vice versa, might be useful for understanding various other brain disorders that are thought to influence processing of uncertainty and have largely been studied in the context of volatility in the past decade[14,16,17,19,21,22,66,79]. As another example, positive symptoms in schizophrenia have been argued to result from some alterations in prior vs likelihood processing, perhaps driven by abnormal attribution of uncertainty (or precision) to top-down expectations[80]. But different such symptoms (e.g. hallucinations vs. delusions) manifest in different patients. One reason may be that these relate to disruption at different levels of a perceptual-inferential hierarchy, i.e. with hallucination vs delusion reflecting involvement of more or less abstract inferential levels, respectively[81–83]. In this respect, the current model may provide a simple and direct comparative test, since stochasticity enters at the perceptual, or outcome, level (potentially associated with hallucination) but volatility acts at the more abstract level of the latent reward (and may be associated with delusion; see Figure 1).

Our work also touches upon a historical debate in the associative learning literature about the role of outcome stochasticity (i.e., in our terms, noise) in learning. One class of theories, most prominently represented by Mackintosh[37], proposes that attention is preferentially allocated to cues that are most reliably predictive of outcomes; whereas Pearce and Hall[60] suggest the opposite: that attention is attracted to surprising misprediction. We address only a subset of the experimental phenomena involved in this debate (those involving learning rates for cues presented alone), but for this subset we offer a very clear resolution of the apparent conflict. Our approach and goals also differ from classic work in this area. A number of important models of attention in psychology also attempt to reconcile these theories by providing more phenomenological models that hybridize the two theories to account for various and often paradoxical experimental work[84–87]. Our goal is different and is more descended from a tradition of normative theories that provide a computational understanding of psychological phenomena from first principles by first addressing what is the computational problem that the corresponding neural system is evolved to solve[2,88].

Any probabilistic model relies on a set of explicit assumptions about how observations have been generated, i.e. a generative model, and also an inference procedure to estimate the hidden parameters that are not directly observable. Such inference algorithms typically reflect some approximation strategy because exact inference is not possible for most important problems, including our generative model

(Figure 1). In previous work in this area, we and others have relied on variational approaches to approximate inference, which factors difficult inference problems into smaller tractable ones, and approximates the answer as though they were independent[11,12]. Interestingly, although one of the most promising successes of this approach in neuroscience has been in hierarchical Kalman filters with volatility inference, we found it difficult to develop an effective variational filter for the current problem, when stochasticity is unknown. The core problem, in our hands, was that effective explaining away between the two noise types was difficult to achieve using simplified variational posteriors that omitted aspects of their mutual dependency.

Interestingly, there are other algorithms that, in principle, address similar learning problems. These include using an explicitly variational approach extending the HGF (code is publicly available as hgf_jget in the TAPAS toolbox[89], but has not been documented or tested in published articles), augmenting the variational HGF with mixture models[41], an analogous simplified learning rule based more on neural considerations[42], and an exact model for tracking hazard rates under a particular case of changepoint detection[43]. While these have not yet been applied to the full range of problems we investigate here, we suspect that future work investigating the approximate approaches will find challenges in explaining away. In any case, in the current modeling, we have adopted a different estimation method based on Monte Carlo sampling, in particular a variant of particle filtering that preserves many of the advantages of variational methods by incorporating exact conditional inference for a subset of variables[47]. The inference model employed here combines Kalman filtering for estimation of reward rate[39] conditional on the volatility and stochasticity, with particle filtering for inference about these[48]. One drawback of the particle filter, however, is that it requires tracking a number of samples on every trial. In practice, we found that a handful (e.g. 100) of particles results in a sufficiently good approximation.

Finally, in this study, we only modeled the effects of volatility and stochasticity on learning rate. However, uncertainty affects many different problems beyond learning rate, and a full account of how subjects infer volatility and stochasticity (and how these, in turn, affect uncertainty) may have ramifications for many other behaviors. Thus, there have been important statistical accounts of a number of such problems, but most of them have neglected either stochasticity or volatility, and none of them have explicitly considered the effects of *learning* the levels of these types of noise. These problems include cue- or feature-selective attention[2]; the explore-exploit dilemma[90,91]; and the partition of experience into latent states, causes or contexts[5,78,92]. The current model, or variants of it, is more or less directly applicable to all these problems and should imply new predictions about the effects of manipulating either type of noise across many different behaviors.

Methods

Description of the model

Recall that outcome on trial t, o_t, in our model depends on three latent variables, the reward rate, stochasticity and volatility. The reward rate on trial t, x_t, has Markov-structure dynamics:

$$x_t = x_{t-1} + e_t, \tag{5}$$

where e_t is a (zero-mean) Gaussian noise with variance given by volatility. Therefore, we have:

$$p(x_t|x_{t-1}, v_t) = N(x_t|x_{t-1}, v_t), \tag{6}$$

where v_t is volatility. We define the inverse volatility, $z_t = v_t^{-1}$, which is the preferred formulation here as it has been used in previous studies for its analytical plausibility[12]. Outcomes were generated based on the reward rate and stochasticity according to a Gaussian distribution:

$$p(o_t|x_t, u_t) = N(o_t|x_t, s_t), \tag{7}$$

where s_t is the stochasticity with $y_t = s_t^{-1}$.

For volatility and stochasticity, we assumed a multiplicative noise on their inverse, which is an approach that has been shown to give rise to analytical inference when considered in isolation (but not here) [93,94]. Specifically, the dynamics over these variables are given by $z_t = {\eta_v}^{-1} z_{t-1}\epsilon_t$, where $0 < \eta_v < 1$ is a constant and ϵ_t is a random variable in the unit range and a Beta-distribution $p(\epsilon_t) = \mathrm{B}(\epsilon_t|0.5\eta_v(1-\eta_v)^{-1}, 0.5)$. Note that the conditional expectation of z_t is given by z_{t-1}, because $E[\epsilon_t] = \eta_v$. We assume a similar and independent dynamics for y_t parametrized by the constant η_s: $y_t = {\eta_s}^{-1} y_{t-1}\varepsilon_t$.

In our implementation, we parametrized the model with $\lambda_v = 1 - \eta_v$ and $\lambda_s = 1 - \eta_s$, respectively. This is because these parameters can be interpreted as the update rate for volatility and stochasticity, respectively. In other words, larger values of λ_v and λ_s result in faster update of volatility and stochasticity, respectively. Intuitively, this is because a smaller λ_v increases the mean of ϵ_t and results in a larger update of z_t. Since volatility is the inverse of z_t, therefore, smaller λ_v results in slower update of volatility. This has been formally shown in our recent work[12]. In addition to these two parameters, this generative process depends on initial value of volatility and stochasticity, v_0 and s_0.

For inference, we employed a Rao-Blackwellised Particle Filtering approach[47], in which the inference about z_t and s_t were made by a particle filter[48] and, conditional on these, the inference over x_t was given by the Kalman filter (i.e. Eqs (1-4)). The particle filter is a Monte Carlo sequential importance sampling method, which keeps track of a set of particles (i.e. samples). The algorithm performs three steps on each trial. First, in a prediction step, each particle is transitioned to the next step based on the generative process. Second, weights of each particle are updated based on the probability of observed outcome:

$$b_t^l \propto N(o_t | m_{t-1}^l, w_{t-1}^l + v_t^l + s_t^l), \quad (8)$$

where b_t^l is the weight of particle l on trial t, m_{t-1}^l and w_{t-1}^l are estimated mean and variance by the Kalman filter on the previous trial (Eqs. 1-4), and v_t^l and s_t^l are volatility and stochasticity samples (i.e. the inverse of z_t^l and y_t^l). In this step, particles were also resampled using the systematic resampling procedure if the ratio of effective to total particles falls below 0.5. In the third step, the Kalman filter was used to update the mean and variance. In particular, for every particle, Eqs 1-4 were used to define α_t^l and update m_t^l and w_t^l for every particle. Learning rate and estimated reward rate on trial t was then defined as the weighted average of all particles, in which the weights were given by b_t^l. We have used particle filter routines implemented in MATLAB.

Finally, we should note that our results are not dependent on the specific generative process that we have assumed here. In particular, it is possible to define a generative process that diffuses according to Gaussian noise. In such a generative model, random variables related to volatility and stochasticity diffuse according to independent Gaussian noise:

$$p(z_t | z_{t-1}) = N(z_t | z_{t-1}, \sigma_v^2), \quad (9)$$

$$p(y_t | y_{t-1}) = N(y_t | y_{t-1}, \sigma_s^2), \quad (10)$$

where volatility and stochasticity are respectively defined as $v_t = \exp(z_t)$ and $s_t = \exp(y_t)$, and σ_v and σ_s are model parameters that play analogous role as λ_v and λ_s above, respectively. The reward rate and outcomes are then generated based on the same process as the previous generative model (Equations 5-7). Inference about reward rate also remains the same (Equations 1-4). Our simulations show that, as long as the particle filter was used for inference about volatility and stochasticity, such a process can successfully recover true unknown volatility and stochasticity (Supplementary Figure 3).

Simulation details

In simulations related to Figures 1-3, timeseries were generated according to the Markov random walk with constant volatility and stochasticity. For these simulations, we assumed $\lambda_v = 0.1$ and $\lambda_s = 0.1$; $v_0 = 1$ (average over small and large true volatility) and $s_0 = 2$ (average over small and large true stochasticity). For lesioned models in Figure 3, the corresponding lesion variable was assumed to be fixed at its initial value throughout the task.

For the conditioned suppression experiment presented in Figure 4, the weak and strong shock was 0.3 and 1, respectively, plus a small noise with variance of 10^{-2}. The noise for the partial reinforcement experiment was assumed to be 10^{-4}. 100 trials were used for training. We assumed 5 omission trials for the omission condition of conditioned suppression experiment. Model parameters in Figure 4 were $\lambda_v = \lambda_s = 0.2$, and $v_0 = s_0 = 0.1$. For corresponding Supplementary Figures 4-5, the response probability of the model was calculated based on a softmax with a decision noise parameter of 5.

Reward rate in Figure 5a was 0.8 in the stable block and switching between 0.25 and 0.75 in the volatile block with the outcome variance of 0.01. For simulations presented in Figure 5a-d, model parameters were similar to those used for simulations in Figure 4 and the stochasticity for the lesioned models was assumed to be 0.001. Volatile condition in Figure 5ef was defined as trials with no contingency switch in their preceding 10 trials, similar to Piray et al.[20]. For the simulation presented in Figure 5h, we followed Huang et al.[32] who fitted a number of reinforcement learning models to choice data, in which they simplified the task to its core features that are directly related to reinforcement learning. Furthermore, we made a further simplification here by considering only two choices. Probability of reward in tasks by Piray et al. and Huang et al. presented in Figure 5 are plotted in Supplementary Figure 6. Outcome variance was assumed to be 0.01. Model parameters were similar to those used for simulations in Figure 4. The stochasticity for the lesioned models in these two simulations were assumed to be 0.05. For simulating choice, we used the softmax with a decision noise of 3.

In Figure 6, the task was a probabilistic switching task with stable and volatile blocks similar to the task of Browning et al.[13]. Outcome variance was assumed to be 0.01. The median correlation presented in the inset of Figure 6b is the Spearman rank correlation across 100 sets of simulations. For each set, 30 artificial subjects were generated, which only differed in their volatility- and stochasticity- update rate parameters. To have a relatively uniform model trait anxiety (i.e. volatility to stochasticity update rate), every set was further divided to 3 subsets (each containing 10 artificial subjects), in which the mean of model trait anxiety was 0.5, 1, 3, respectively. We further ensured that the model trait anxiety is greater than 0.26 and smaller than 4. These values were chosen to relatively reflect the distribution of trait anxiety in Browning et al.'s[13] data (Figure 6a). Furthermore, the volatility update rate was drawn randomly between 0 and 0.2 and the stochasticity update rate was calculated according to the model trait anxiety. A fixed and small initial volatility and stochasticity was used for all artificial subjects ($v_0 = s_0 = 0.001$).

For simulating the experiment in Figure 7, reward timeseries was generated with a very small outcome variance, 10^{-6}. Here, the model was trained to predict both the tone (give the light) and the reward (given the light) on every trial. Model parameters were $\lambda_v = \lambda_s = 0.2$, and $v_0 = s_0 = 0.5$. Volatility was assumed for the lesioned model to be small and fixed (0.25e-6). Figure 7b shows the average learning rate on the last trial of phase 2 (i.e. the first trial of the test) for the first cue across all simulations. Figure 7c-f shows volatility and stochasticity signals for the first cue. For simulation of the reversal task in Figure 8, a small outcome variance similar, 10^{-6}, was used for generating outcomes. Model parameters were the same as those used in Figure 7. Volatility for the lesioned model was assumed to be small and fixed at 0.01. We have used softmax with a decision noise parameter as the choice model. We assumed that the decision noise parameter is 3 and 1 for the control and lesioned animals, respectively. These parameters were used to reproduce the general reduction of performance in lesioned animals, which is independent of the difference between the deterministic vs stochastic task in the two groups explained by our learning model.

For Supplementary Figure 2, we followed the design of the prediction task by Nassar et al.[25]. The timeseries were generated according to a hidden reward rate plus a noise term, in which the variance of

noise was either 1 (small) or 9 (large). The reward rate was subject to a random jump in the range of 0 to 10. Change points occurred after at least five trials plus a random draw according to an exponential distribution with a rate of 0.05 (i.e. mean 20). The initial volatility and stochasticity were both assumed to be 5 (average over small and large true stochasticity). We further assumed that $\lambda_v = 0.4$ and $\lambda_s = 0.2$, which reflects the instructions given to subjects about possible jumps in the underlying reward rate. Simulation and model parameters in Supplementary Figure 1 were the same as those in Figure 5a. For all simulations, we have assumed initial reward rate prior to be a Gaussian with mean 0 and variance 100 (Figures 2-3 and Supplementary Figure 3) or 1 (all other simulations). Simulations were repeated sufficiently to have negligible sampling error (i.e. invisible standard error of the mean). Thus, simulations presented in Figures 1, 2, 3 and Supplementary Figure 3 were repeated 10000 times; conditioning simulations presented in Figures 4 and 7 were repeated 40000 times; and all other simulations were repeated 1000 times. All simulations were conducted with 100 particles.

Data availability

Simulation data are available at https://github.com/payampiray/stochasticity_volatility_learning. Data by Piray et al. [20] presented in Figure 5e are publicly available at https://github.com/payampiray/piray_etal_2019_JNeurosci. Data by Browning et al.[13] plotted in Figure 6a are publicly available as a Source Data file to the corresponding paper. Source data are provided with this paper.

Code availability

All simulations were conducted using custom code written in MATLAB (2018a). Codes are available at https://github.com/payampiray/stochasticity_volatility_learning.

References

1. Dayan, P. & Long, T. Statistical Models of Conditioning. in 117–123 (1998).
2. Dayan, P., Kakade, S. & Montague, P. R. Learning and selective attention. *Nat. Neurosci.* **3 Suppl**, 1218–1223 (2000).
3. Courville, A. C., Daw, N. D. & Touretzky, D. S. Bayesian theories of conditioning in a changing world. *Trends Cogn. Sci. (Regul. Ed.)* **10**, 294–300 (2006).
4. Daunizeau, J. *et al.* Observing the observer (I): meta-bayesian models of learning and decision-making. *PLoS ONE* **5**, e15554 (2010).
5. Gershman, S. J., Blei, D. M. & Niv, Y. Context, learning, and extinction. *Psychol Rev* **117**, 197–209 (2010).
6. Behrens, T. E. J., Woolrich, M. W., Walton, M. E. & Rushworth, M. F. S. Learning the value of information in an uncertain world. *Nat. Neurosci.* **10**, 1214–1221 (2007).
7. Nassar, M. R. *et al.* Rational regulation of learning dynamics by pupil-linked arousal systems. *Nat. Neurosci.* **15**, 1040–1046 (2012).
8. Iglesias, S. *et al.* Hierarchical prediction errors in midbrain and basal forebrain during sensory learning. *Neuron* **80**, 519–530 (2013).
9. McGuire, J. T., Nassar, M. R., Gold, J. I. & Kable, J. W. Functionally dissociable influences on learning rate in a dynamic environment. *Neuron* **84**, 870–881 (2014).
10. Soltani, A. & Izquierdo, A. Adaptive learning under expected and unexpected uncertainty. *Nat. Rev. Neurosci.* **20**, 635–644 (2019).

11. Mathys, C., Daunizeau, J., Friston, K. J. & Stephan, K. E. A bayesian foundation for individual learning under uncertainty. *Front Hum Neurosci* **5**, 39 (2011).
12. Piray, P. & Daw, N. D. A simple model for learning in volatile environments. *PLoS Comput. Biol.* **16**, e1007963 (2020).
13. Browning, M., Behrens, T. E., Jocham, G., O'Reilly, J. X. & Bishop, S. J. Anxious individuals have difficulty learning the causal statistics of aversive environments. *Nat. Neurosci.* **18**, 590–596 (2015).
14. Brazil, I. A., Mathys, C. D., Popma, A., Hoppenbrouwers, S. S. & Cohn, M. D. Representational Uncertainty in the Brain During Threat Conditioning and the Link With Psychopathic Traits. *Biol Psychiatry Cogn Neurosci Neuroimaging* **2**, 689–695 (2017).
15. Farashahi, S. *et al.* Metaplasticity as a Neural Substrate for Adaptive Learning and Choice under Uncertainty. *Neuron* **94**, 401-414.e6 (2017).
16. Lawson, R. P., Mathys, C. & Rees, G. Adults with autism overestimate the volatility of the sensory environment. *Nat. Neurosci.* **20**, 1293–1299 (2017).
17. Powers, A. R., Mathys, C. & Corlett, P. R. Pavlovian Conditioning-Induced Hallucinations Result from Overweighting of Perceptual Priors. *Science* **357**, 596–600 (2017).
18. Katthagen, T. *et al.* Modeling subjective relevance in schizophrenia and its relation to aberrant salience. *PLoS Comput. Biol.* **14**, e1006319 (2018).
19. Paliwal, S. *et al.* Subjective estimates of uncertainty during gambling and impulsivity after subthalamic deep brain stimulation for Parkinson's disease. *Scientific Reports* **9**, 14795 (2019).
20. Piray, P., Ly, V., Roelofs, K., Cools, R. & Toni, I. Emotionally Aversive Cues Suppress Neural Systems Underlying Optimal Learning in Socially Anxious Individuals. *J. Neurosci.* **39**, 1445–1456 (2019).
21. Cole, D. M. *et al.* Atypical processing of uncertainty in individuals at risk for psychosis. *NeuroImage: Clinical* **26**, 102239 (2020).
22. Deserno, L. *et al.* Volatility Estimates Increase Choice Switching and Relate to Prefrontal Activity in Schizophrenia. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging* **5**, 173–183 (2020).
23. Diaconescu, A. O., Wellstein, K. V., Kasper, L., Mathys, C. & Stephan, K. E. Hierarchical Bayesian models of social inference for probing persecutory delusional ideation. *Journal of Abnormal Psychology* **129**, 556–569 (2020).
24. Lee, S., Gold, J. I. & Kable, J. W. The human as delta-rule learner. - PsycNET. *Decision* **7**, 55–66 (2020).
25. 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).
26. Pulcu, E. & Browning, M. The Misestimation of Uncertainty in Affective Disorders. *Trends in Cognitive Sciences* **23**, 865–875 (2019).
27. Dugas, M. J., Gagnon, F., Ladouceur, R. & Freeston, M. H. Generalized anxiety disorder: a preliminary test of a conceptual model. *Behav Res Ther* **36**, 215–226 (1998).
28. Ladouceur, R., Gosselin, P. & Dugas, M. J. Experimental manipulation of intolerance of uncertainty: a study of a theoretical model of worry. *Behav Res Ther* **38**, 933–941 (2000).
29. Aylward, J. *et al.* Altered learning under uncertainty in unmedicated mood and anxiety disorders. *Nature Human Behaviour* **3**, 1116–1123 (2019).
30. Gagne, C., Dayan, P. & Bishop, S. J. When planning to survive goes wrong: predicting the future and replaying the past in anxiety and PTSD. *Current Opinion in Behavioral Sciences* **24**, 89–95 (2018).
31. Harlé, K. M., Guo, D., Zhang, S., Paulus, M. P. & Yu, A. J. Anhedonia and anxiety underlying depressive symptomatology have distinct effects on reward-based decision-making. *PLOS ONE* **12**, e0186473 (2017).
32. Huang, H., Thompson, W. & Paulus, M. P. Computational Dysfunctions in Anxiety: Failure to Differentiate Signal From Noise. *Biological Psychiatry* **82**, 440–446 (2017).
33. Huys, Q. J. M., Daw, N. D. & Dayan, P. Depression: a decision-theoretic analysis. *Annu. Rev. Neurosci.* **38**, 1–23 (2015).
34. Luhmann, C. C., Ishida, K. & Hajcak, G. Intolerance of uncertainty and decisions about delayed, probabilistic rewards. *Behav Ther* **42**, 378–386 (2011).
35. Paulus, M. P. & Yu, A. J. Emotion and decision-making: affect-driven belief systems in anxiety and depression. *Trends Cogn. Sci.* **16**, 476–483 (2012).
36. Pulcu, E. & Browning, M. Affective bias as a rational response to the statistics of rewards and punishments. *eLife* **6**, e27879 (2017).

37. Mackintosh, N. J. A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review* **82**, 276–298 (1975).
38. Pearce, J. M. & Hall, G. A model for Pavlovian learning: variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychol Rev* **87**, 532–552 (1980).
39. Kalman, R. E. A New Approach to Linear Filtering and Prediction Problems. *Transactions of the ASME--Journal of Basic Engineering* **82**, 35–45 (1960).
40. Kakade, S. & Dayan, P. Acquisition and extinction in autoshaping. *Psychol Rev* **109**, 533–544 (2002).
41. Moens, V. & Zénon, A. Learning and forgetting using reinforced Bayesian change detection. *PLoS Comput Biol* **15**, (2019).
42. Silvetti, M., Vassena, E., Abrahamse, E. & Verguts, T. Dorsal anterior cingulate-brainstem ensemble as a reinforcement meta-learner. *PLOS Computational Biology* **14**, e1006370 (2018).
43. Wilson, R. C., Nassar, M. R. & Gold, J. I. Bayesian On-line Learning of the Hazard Rate in Change-Point Problems. *Neural Comput* **22**, 2452–2476 (2010).
44. Griffiths, T. L., Navarro, D. J. & Sanborn, A. N. A More Rational Model of Categorization. *Proceedings of the Annual Meeting of the Cognitive Science Society* **28**, (2006).
45. Daw, N. D. & Courville, A. C. The rat as particle filter. in *Advances in Neural Information Processing Systems 20* (eds. Platt, J. C., Koller, D., Singer, Y. & Roweis, S. T.) 369–376 (Curran Associates, Inc., 2008).
46. Brown, S. D. & Steyvers, M. Detecting and predicting changes. *Cogn Psychol* **58**, 49–67 (2009).
47. Doucet, A., Freitas, N. de, Murphy, K. P. & Russell, S. J. Rao-Blackwellised Particle Filtering for Dynamic Bayesian Networks. in *Proceedings of the 16th Conference on Uncertainty in Artificial Intelligence* 176–183 (Morgan Kaufmann Publishers Inc., 2000).
48. Doucet, A. & Johansen, A. M. *A tutorial on particle filtering and smoothing: fifteen years later*. (2011).
49. Behrens, T. E. J., Hunt, L. T., Woolrich, M. W. & Rushworth, M. F. S. Associative learning of social value. *Nature* **456**, 245–249 (2008).
50. Nassar, M. R. *et al.* Age differences in learning emerge from an insufficient representation of uncertainty in older adults. *Nat Commun* **7**, (2016).
51. Hall, G. & Pearce, J. M. Restoring the Associability of a Pre-Exposed CS by a Surprising Event. *The Quarterly Journal of Experimental Psychology Section B* **34**, 127–140 (1982).
52. Gibbon, J., Farrell, L., Locurto, C. M., Duncan, H. J. & Terrace, H. S. Partial reinforcement in autoshaping with pigeons. *Animal Learning & Behavior* **8**, 45–59 (1980).
53. Rescorla, R. A. Within-subject partial reinforcement extinction effect in autoshaping. *The Quarterly Journal of Experimental Psychology B: Comparative and Physiological Psychology* **52B**, 75–87 (1999).
54. Haselgrove, M., Aydin, A. & Pearce, J. M. A partial reinforcement extinction effect despite equal rates of reinforcement during Pavlovian conditioning. *J Exp Psychol Anim Behav Process* **30**, 240–250 (2004).
55. Li, J., Schiller, D., Schoenbaum, G., Phelps, E. A. & Daw, N. D. Differential roles of human striatum and amygdala in associative learning. *Nat. Neurosci.* **14**, 1250–1252 (2011).
56. Gallistel, C. R. & Gibbon, J. Time, rate, and conditioning. *Psychol Rev* **107**, 289–344 (2000).
57. Phelps, E. A., Lempert, K. M. & Sokol-Hessner, P. Emotion and decision making: multiple modulatory neural circuits. *Annu. Rev. Neurosci.* **37**, 263–287 (2014).
58. Averbeck, B. B. & Costa, V. D. Motivational neural circuits underlying reinforcement learning. *Nat. Neurosci.* **20**, 505–512 (2017).
59. Holland, P. C. & Gallagher, M. Amygdala circuitry in attentional and representational processes. *Trends Cogn. Sci. (Regul. Ed.)* **3**, 65–73 (1999).
60. Roesch, M. R., Esber, G. R., Li, J., Daw, N. D. & Schoenbaum, G. Surprise! Neural Correlates of Pearce-Hall and Rescorla-Wagner Coexist within the Brain. *Eur J Neurosci* **35**, 1190–1200 (2012).
61. Costa, V. D., Dal Monte, O., Lucas, D. R., Murray, E. A. & Averbeck, B. B. Amygdala and Ventral Striatum Make Distinct Contributions to Reinforcement Learning. *Neuron* **92**, 505–517 (2016).
62. Homan, P. *et al.* Neural computations of threat in the aftermath of combat trauma. *Nat. Neurosci.* **22**, 470–476 (2019).
63. Holland, P. C. & Gallagher, M. Amygdala central nucleus lesions disrupt increments, but not decrements, in conditioned stimulus processing. *Behav. Neurosci.* **107**, 246–253 (1993).
64. Holland, P. C. & Schiffino, F. L. Mini-review: Prediction errors, attention and associative learning. *Neurobiol Learn Mem* **131**, 207–215 (2016).

65. Hampton, A. N., Adolphs, R., Tyszka, M. J. & O'Doherty, J. P. Contributions of the amygdala to reward expectancy and choice signals in human prefrontal cortex. *Neuron* **55**, 545–555 (2007).
66. Diaconescu, A. O. *et al.* Inferring on the Intentions of Others by Hierarchical Bayesian Learning. *PLoS Comput Biol* **10**, (2014).
67. de Berker, A. O. *et al.* Computations of uncertainty mediate acute stress responses in humans. *Nat Commun* **7**, 10996 (2016).
68. Khorsand, P. & Soltani, A. Optimal structure of metaplasticity for adaptive learning. *PLoS Comput. Biol.* **13**, e1005630 (2017).
69. Diederen, K. M. J. & Schultz, W. Scaling prediction errors to reward variability benefits error-driven learning in humans. *Journal of Neurophysiology* **114**, 1628–1640 (2015).
70. Wilson, R. C., Nassar, M. R. & Gold, J. I. A Mixture of Delta-Rules Approximation to Bayesian Inference in Change-Point Problems. *PLOS Computational Biology* **9**, e1003150 (2013).
71. Dayan, P. & Yu, A. Uncertainty and Learning. *IETE Journal of Research* **49**, 171–181 (2003).
72. Yu, A. J. & Dayan, P. Uncertainty, neuromodulation, and attention. *Neuron* **46**, 681–692 (2005).
73. Hartley, C. A. & Phelps, E. A. Anxiety and decision-making. *Biol. Psychiatry* **72**, 113–118 (2012).
74. Beck, A. T. *Depression: Causes and Treatment*. (University of Pennsylvania Press, 1970).
75. Duits, P. *et al.* Updated meta-analysis of classical fear conditioning in the anxiety disorders. *Depress Anxiety* **32**, 239–253 (2015).
76. Lissek, S. *et al.* Classical fear conditioning in the anxiety disorders: a meta-analysis. *Behav Res Ther* **43**, 1391–1424 (2005).
77. Bouton, M. E. Context and behavioral processes in extinction. *Learn Mem* **11**, 485–494 (2004).
78. Redish, A. D., Jensen, S., Johnson, A. & Kurth-Nelson, Z. Reconciling reinforcement learning models with behavioral extinction and renewal: implications for addiction, relapse, and problem gambling. *Psychol Rev* **114**, 784–805 (2007).
79. Reed, E. J. *et al.* Paranoia as a deficit in non-social belief updating. *eLife* **9**, (2020).
80. Stephan, K. E., Baldeweg, T. & Friston, K. J. Synaptic plasticity and dysconnection in schizophrenia. *Biol. Psychiatry* **59**, 929–939 (2006).
81. Baker, S. C., Konova, A. B., Daw, N. D. & Horga, G. A distinct inferential mechanism for delusions in schizophrenia. *Brain* **142**, 1797–1812 (2019).
82. Horga, G. & Abi-Dargham, A. An integrative framework for perceptual disturbances in psychosis. *Nature Reviews Neuroscience* **20**, 763–778 (2019).
83. Wengler, K., Goldberg, A., Chahine, G. & Horga, G. Hallucinations and Delusions Relate to Distinct Hierarchical Alterations in Intrinsic Neural Timescales. *Biological Psychiatry* **87**, S179–S180 (2020).
84. Le Pelley, M. E. The role of associative history in models of associative learning: a selective review and a hybrid model. *Q J Exp Psychol B* **57**, 193–243 (2004).
85. Haselgrove, M., Esber, G. R., Pearce, J. M. & Jones, P. M. Two kinds of attention in Pavlovian conditioning: evidence for a hybrid model of learning. *J Exp Psychol Anim Behav Process* **36**, 456–470 (2010).
86. Pearce, J. M. & Mackintosh, N. J. Two theories of attention: a review and a possible integration. in *Attention and Associative Learning: From Brain to Behaviour* (eds. Mitchell, C. & Le Pelley, M. E.) 11–39 (Oxford University Press, 2010).
87. Le Pelley, M. E., Mitchell, C. J., Beesley, T., George, D. N. & Wills, A. J. Attention and associative learning in humans: An integrative review. *Psychol Bull* **142**, 1111–1140 (2016).
88. Marr, D. *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. (W. H. Freeman and Company, 1982).
89. Aponte, E. *et al. TAPAS - Translational Algorithms for Psychiatry-Advancing Science*. (2020).
90. Gittins, J. C. Bandit Processes and Dynamic Allocation Indices. *Journal of the Royal Statistical Society. Series B (Methodological)* **41**, 148–177 (1979).
91. Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B. & Dolan, R. J. Cortical substrates for exploratory decisions in humans. *Nature* **441**, 876–879 (2006).
92. Wilson, R. C., Takahashi, Y. K., Schoenbaum, G. & Niv, Y. Orbitofrontal cortex as a cognitive map of task space. *Neuron* **81**, 267–279 (2014).
93. West, M. On Scale Mixtures of Normal Distributions. *Biometrika* **74**, 646–648 (1987).

94. Gamerman, D., dos Santos, T. R. & Franco, G. C. A Non-Gaussian Family of State-Space Models with Exact Marginal Likelihood. *J. Time Ser. Anal.* **34**, 625–645 (2013).

Acknowledgments

We thank Sam Zorowitz, Peter Dayan, Yoel Sanchez Araujo, and Guillermo Horga for helpful discussions. This work was supported by grants IIS-1822571 from the National Science Foundation, part of the CRNCS program, and 61454 from the John Templeton Foundation.

Author Contributions

P.P. and N.D.D. designed the study and wrote the manuscript. P.P. performed analyses.

Competing Interests

The authors declare no competing interests.

Figure legends

Figure 1. Statistical difference between volatility and stochasticity. a-b) Examples of generated time-series based on a small and large constant volatility parameter given a small (a) or a large (b) constant stochasticity parameter are plotted. c) Given a surprising observation (e.g. a negative outcome), one should compute how likely the outcome is due to the stochasticity (left balloon) or due to the volatility (right balloon). Dissociating these two terms is important for learning, because they have opposite influences on learning rate. d-e) It is possible to infer both volatility and stochasticity based on observed outcomes, because these parameters have dissociable statistical signatures. In particular, although both of them increase variance (d), but they have opposite effects on autocorrelation (e). In particular, whereas volatility increases autocorrelation, stochasticity tends to reduce it. Here, 1-step autocorrelation (i.e. correlation between trial t and $t-1$) was computed for 100 time-series generated with parameters defined in b and c. Small and large parameters for volatility were 0.5 and 1.5 and for stochasticity were 1 and 3, respectively. f) Structure of the (generative) model: outcomes were stochastically generated based on a probabilistic model depending on reward rate, stochasticity and volatility. Only outcomes were observable (the gray circle), and value of all other parameters should be inferred based on outcomes. The observed outcome is given by the true reward rate, x_t, plus some noise whose variance is given by the stochasticity, s_t. The reward rate itself depends on its value on the previous trial plus some noise whose variance is given by the volatility, v_t. Both volatility and stochasticity are dynamic and probabilistic Markovian variables generated noisily based on their value on the previous trial. Thus, the model has two modules, volatility and stochasticity, which compete to explain experienced noise in outcomes. See Methods for formal treatment of the model. Errorbars in (d-e) are standard error of the mean calculated across 10000 simulations and are too small to be visible. Source data are provided as Source Data file.

Figure 2. Performance of the model in task with constant but unknown volatility and stochasticity parameters. Outcomes were generated according to the same procedure and parameters as those used in Figure 1 (see Figure 1ab for example outcome time-series seen by the model). a) Learning rate in the model varies by changes in both the true volatility and stochasticity. Furthermore, these parameters have

opposite effects on learning rate. In contrast to volatility, higher stochasticity reduces the learning rate. b) Estimated volatility captures variations in true volatility. c) Estimated stochasticity captures variations in the true stochasticity. In (a-c), average learning rate, estimated volatility and stochasticity in the last 20 trials were plotted over all simulations (100 simulations). d-f) Learning rate, volatility and stochasticity estimates by the model for small true volatility. g-i) The three signals are plotted for the larger true volatility. Estimated volatility and stochasticity by the model capture their corresponding true values. Errorbars are standard error of the mean computed over 10000 simulations and are too small to be visible. See also Supplementary Figure 3 for further simulation analysis. Source data are provided as Source Data file.

Figure 3. Behavior of the lesioned model. a) Stochasticity and volatility module inside the model compete to explain experienced noise. b-c) Two characteristic lesioned models produce seemingly contradictory behaviors, because if the stochasticity module is lesioned, noise due to stochasticity is misattributed to volatility (b), and vice versa (c). d-f) Mean learning rate is plotted for the 2x2 design of Figure 2 for the healthy and lesioned models. For both lesion models, lesioning does not merely abolish the corresponding effects on learning rate, but reverses them. Thus, the stochasticity lesion model shows elevated learning rate with increases in stochasticity (e), and the volatility lesion model shows reduced learning rate with increases in volatility (f). This is due to misattribution of the noise due to the lesioned factor to the existing module. g) The stochasticity lesion model makes erroneous inference about volatility and increases its volatility estimate in higher unpredictable environments. h) The volatility lesion model makes erroneous inference about stochasticity and increases its stochasticity estimate for higher volatile environments. In fact, both lesion models are not able to distinguish between volatility and stochasticity and therefore show similar pattern for the remaining module. For the healthy model, volatility and stochasticity estimates are the same as Figure 2b and 2c, respectively. Simulation and model parameters were the same as those used in Figure 2. Errorbars reflect standard error of the mean computed over 10000 simulations and are too small to be visible. Source data are provided as Source Data file.

Figure 4. The model explains puzzling issues in Pavlovian learning. a-d) Pearce and Hall's conditioned suppression experiment. The design of experiment[51], in which they found that the omission group show higher speed of learning than the control group (a). b) Median learning rate over the first trial of the retraining. The learning rate is larger for the omission group due to increases of volatility (c), while stochasticity is similar for both groups (d). The model explains partial reinforcement extinction effects (e-h). e) The partial reinforcement experiment consists of a partial condition in which a light cue if followed by reward on 50% of trials and a full condition in which the cue is always followed by the reward. f) Learning rate over the first trial of retraining has been plotted. Similar to empirical data, the model predicts that the learning rate is larger in the full condition, because partial reinforcements have relatively small effects on volatility (g), but it considerably increases stochasticity (h). Errorbars reflect standard error of the mean over 40000 simulations and are too small to be visible. See Supplementary Figures 4 and 5 for empirical data and corresponding response probability by the model.

Figure 5. The stochasticity lesion model shows a pattern of learning deficits associated with anxiety. Behavior of the lesioned model as the model of anxiety, in which stochasticity is assumed to be small and constant, is shows along the control model. a-d) Behavior of the models in the switching task of Figure 2 is shown. An example of estimated reward by the models shows that the anxious model is more sensitive to noisy outcomes (a), which dramatically reduces sensitivity of the learning rate to volatility manipulation in this task (b). This, however, is primarily related to inability to make inference about stochasticity, which leads to misestimation of volatility (c-d). e-f) The model explains the data reported by Piray et al.[20], in which the high (social) anxiety group did not benefit from stability as much as the low anxiety group (e). The model shows the same behavior (f). g-h) The model explains the data by Huang et al.[32], in which the anxious group showed higher lose-shift behavior compared to the control group (g). The model shows the same behavior (g), which is due to higher learning rate in the anxious group (inset). Errorbars in (b), (f) and (h) reflect standard error of the mean over 1000 simulations and are too small to be visible. Data in (e) and (g) are, respectively, adapted from Piray et al. [20] and Huang et al.[32], in which mean and standard error of the mean are plotted (obtained over n=44 and n=122 independent samples, respectively.) Source data are provided as Source Data file.

Figure 6. The model explains effects of trait anxiety as a continuous index on learning. a) Data by Browning et al.[13] show a significant negative correlation between relative log learning rate and trait anxiety in a probabilistic switching task with stable and volatile blocks. b) The model shows a similar pattern. The inset shows the median rank correlation between trait anxiety and the relative learning rate. Model trait anxiety is defined as the ratio of volatility to stochasticity update rates (thus higher if the stochasticity update rate is small). The lesion model of anxiety (Figure 5) is a special case in which the stochasticity update rate is zero. Errorbars reflect standard error of the median over 1000 simulations and are too small to be visible. Source data are provided as Source Data file.

Figure 7. The model displays the behavior of amygdala lesioned rats in associative learning. a) The task used for studying the role of amygdala in learning by Holland and Gallagher[59,63,64]. Rats in the 'consistent' condition received extensive exposure to a consistent light–tone in a partial reinforcement schedule (i.e. only half of trials led to reward). In the 'shift' condition, however, rats were trained on the same light-tone partial reinforcement schedule in the first phase, but the schedule shifted to a different one in the shorter second phase, in which rats received light-tone-reward on half of trials and light-nothing on the other half. b) Empirical data showed that while the contingency shift facilitates learning in the control rats, it disrupts performance in lesioned rats. c) learning rate in the last trial of second phase shows the same pattern. This is because the shift increases volatility for the control rats (d) but not for the lesioned rate (e). In contrast, the contingency shift increases the stochasticity for the lesioned rats substantially more than that for the control rats, which results in reduced learning rate for the lesioned animals (f-g). The gray line shows the starting trial of the second phase. Data in (b) was originally reported in [63] and reproduced here from [64]. Errorbars in other (c-g) reflect standard error of the mean over 40000 simulations and are too small to be visible. See also supplementary Table 1. Source data are provided as Source Data file.

Figure 8. The model displays the behavior of amygdala lesioned monkeys in probabilistic reversal learning. a) The probabilistic reversal learning task by Costa et al.[61]. The task consists of 80 trials, in which animals chose one of the two stimuli by making a saccade to it and fixating on the chosen cue. A probabilistic reward was given following a correct choice. The stimulus-reward contingency was reversed in the middle of the task (on a random trial between trials 30-50). The task consists of different schedules, but we focus here on 60%/40% (stochastic) and 100%/0% (deterministic), which show the clearest difference in empirical data. b) Performance of animals in this task. In addition to the general reduced performance by the lesioned animals, their performance was substantially more disrupted in the deterministic- than stochastic- reversal. c) Performance of the model in this task shows the same pattern. d-i) Learning rate, volatility and stochasticity signals for the deterministic (d-f) and stochastic task (g-i). Solid and dashed line are related to acquisition and reversal phase, respectively. Deterministic reversal increases the learning rate in the control animals due to increases in volatility, but not in the lesioned monkeys, in which it reduces the learning rate due to the increase of the stochasticity. The reversal in the stochastic task has very small effects on these signals, because stochasticity is relatively large during both acquisition and reversal. Data in (b) are adapted from Costa et al.[61], in which mean and standard error of the mean are plotted. Errorbars in other panels reflect standard error of the mean over 1000 simulations and are too small to be visible. See also Supplementary Figures 7 for choice time-series and Supplementary Figure 8 for simulation of the model in all four probabilistic schedules tested by Costa et al.[61] and corresponding empirical data. Source data are provided as Source Data file.

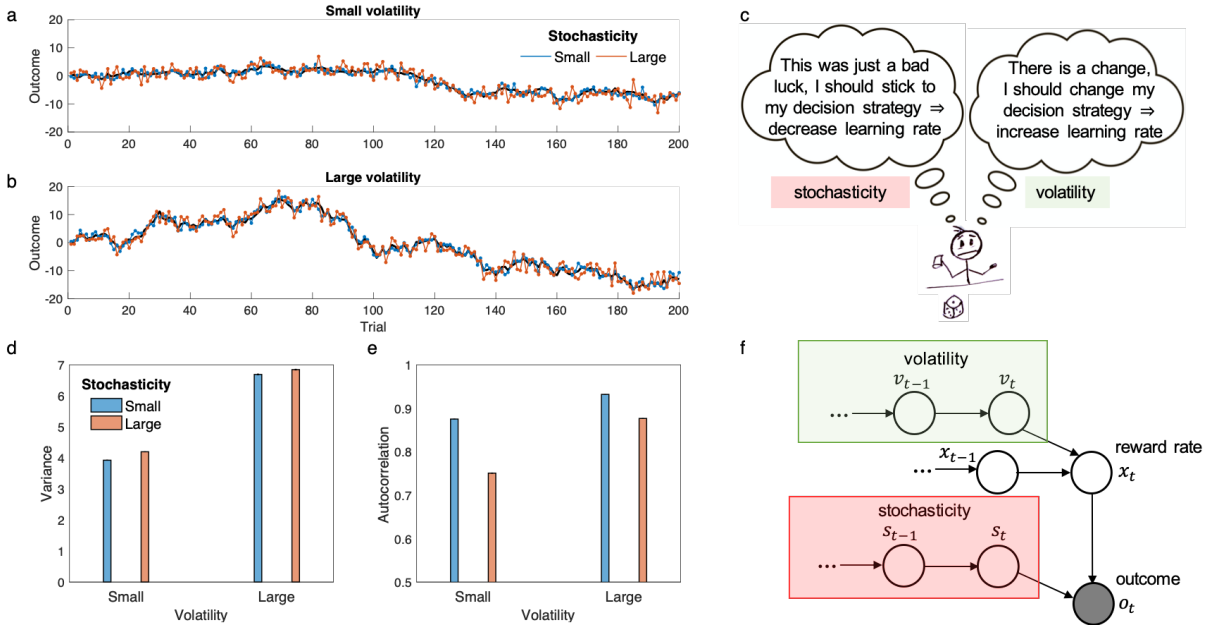
a
Small volatility
Stochasticity
Small
Large
Outcome
b
Large volatility
Outcome
Trial
c
This was just a bad luck, I should stick to my decision strategy ⇒ decrease learning rate
There is a change, I should change my decision strategy ⇒ increase learning rate
stochasticity
volatility
d
Stochasticity
Small
Large
Variance
Small
Large
Volatility
e
Autocorrelation
Small
Large
Volatility
f
volatility
v_{t-1}
v_t
reward rate
x_{t-1}
x_t
stochasticity
s_{t-1}
s_t
outcome
o_t

Fig 1

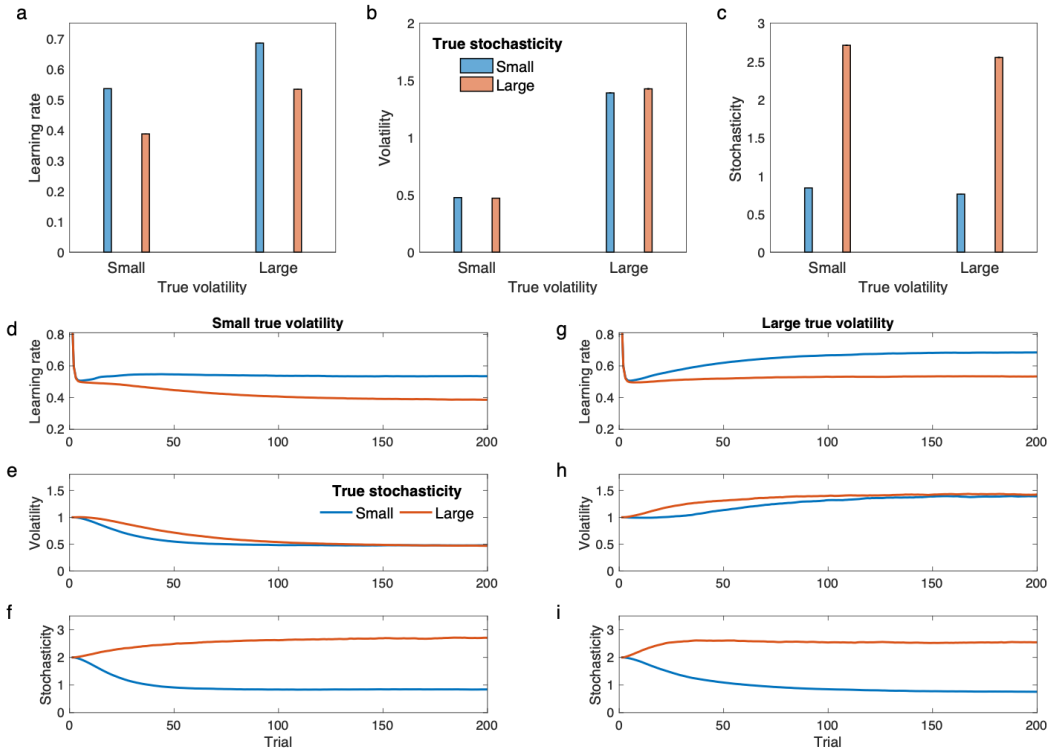
a
Learning rate
Small
Large
True volatility
b
True stochasticity
Small
Large
Volatility
Small
Large
True volatility
c
Stochasticity
Small
Large
True volatility
d
Small true volatility
Learning rate
e
True stochasticity
Small
Large
Volatility
f
Stochasticity
Trial
g
Large true volatility
Learning rate
h
Volatility
i
Stochasticity
Trial

Fig 2

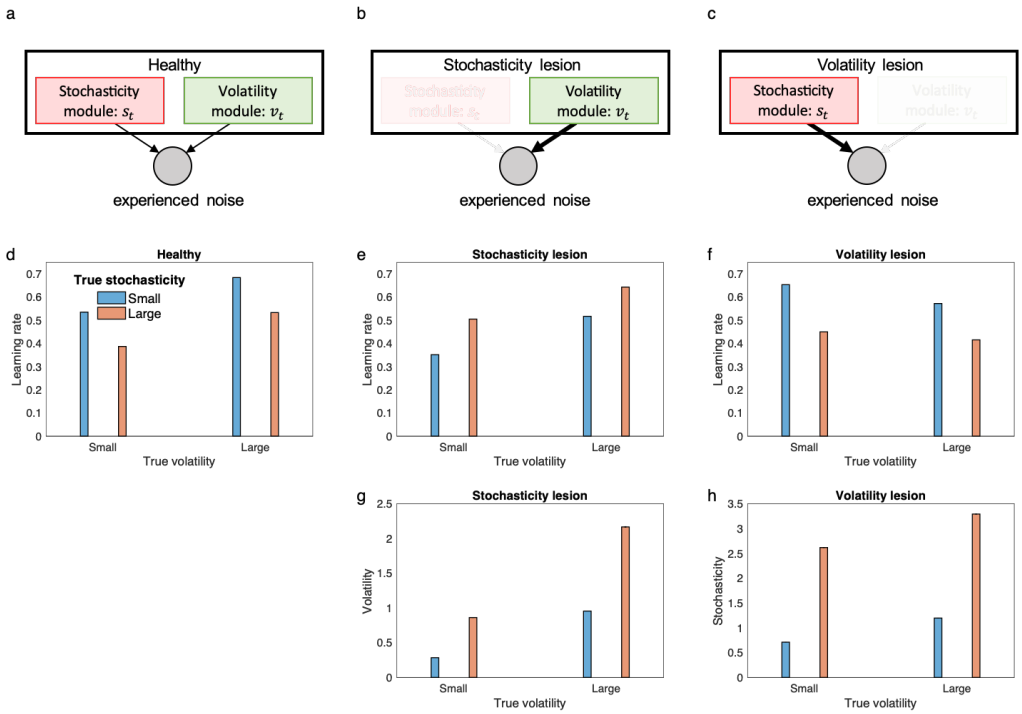
a
Healthy
Stochasticity module: s_t
Volatility module: v_t
experienced noise
b
Stochasticity lesion
Volatility module: v_t
experienced noise
c
Volatility lesion
Stochasticity module: s_t
experienced noise
d
Healthy
True stochasticity
Small
Large
Learning rate
True volatility
e
Stochasticity lesion
Learning rate
True volatility
f
Volatility lesion
Learning rate
True volatility
g
Stochasticity lesion
Volatility
True volatility
h
Volatility lesion
Stochasticity
True volatility

Fig 3

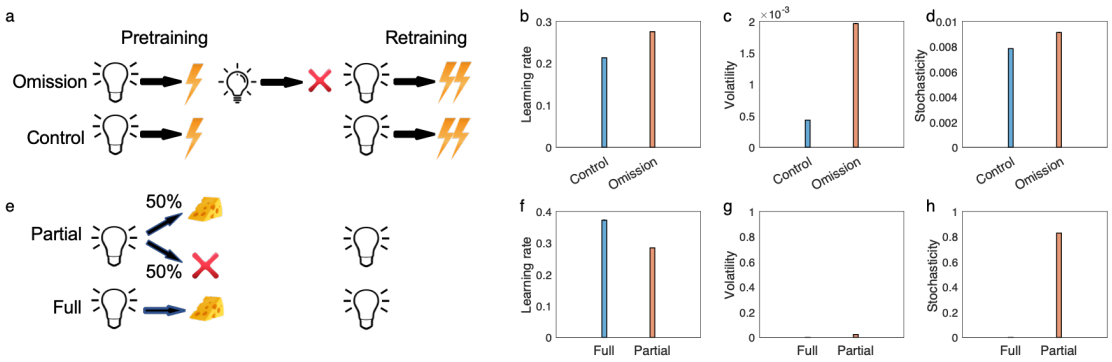
a
Pretraining
Retraining
Omission
Control
e
50%
Partial
50%
Full
b
Learning rate
Control
Omission
c
Volatility
Control
Omission
d
Stochasticity
Control
Omission
f
Learning rate
Full
Partial
g
Volatility
Full
Partial
h
Stochasticity
Full
Partial

Fig 4

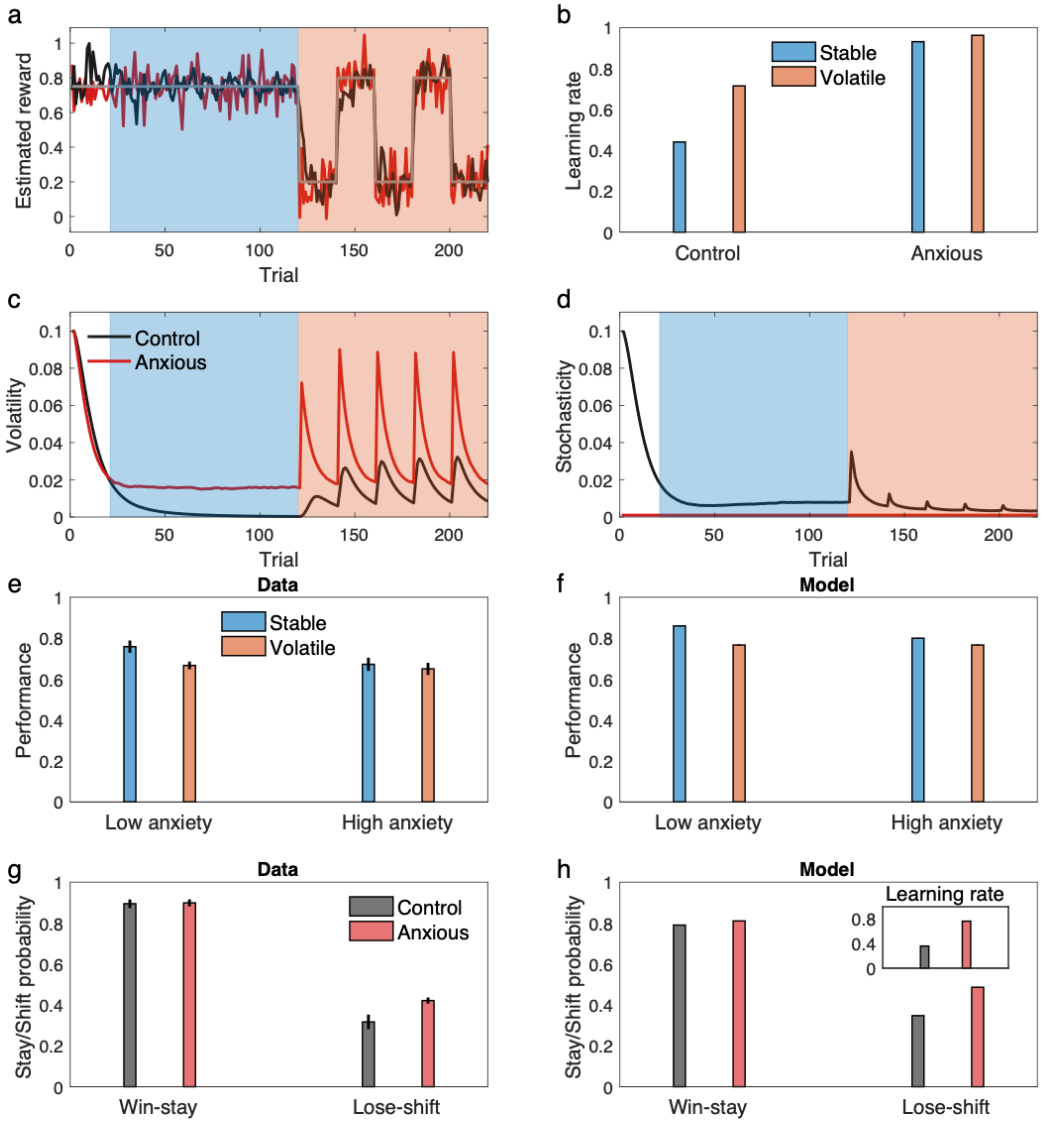
a
Estimated reward
Trial
b
Learning rate
Stable
Volatile
Control
Anxious
c
Volatility
Control
Anxious
Trial
d
Stochasticity
Trial
e
Data
Performance
Stable
Volatile
Low anxiety
High anxiety
f
Model
Performance
Low anxiety
High anxiety
g
Data
Stay/Shift probability
Control
Anxious
Win-stay
Lose-shift
h
Model
Stay/Shift probability
Learning rate
Win-stay
Lose-shift

Fig 5

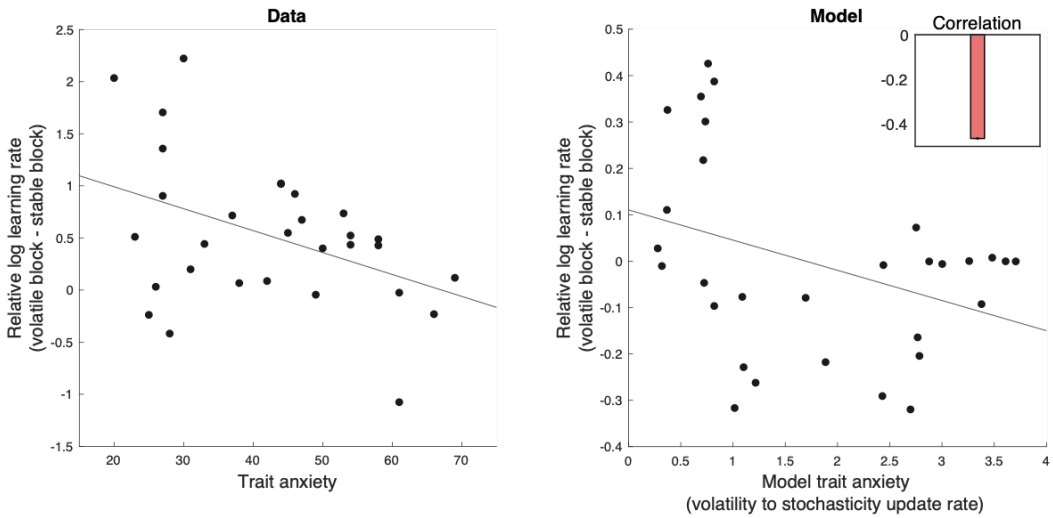
Data
Relative log learning rate
(volatile block - stable block)
Trait anxiety
Model
Correlation
Model trait anxiety
(volatility to stochasticity update rate)

Fig 6

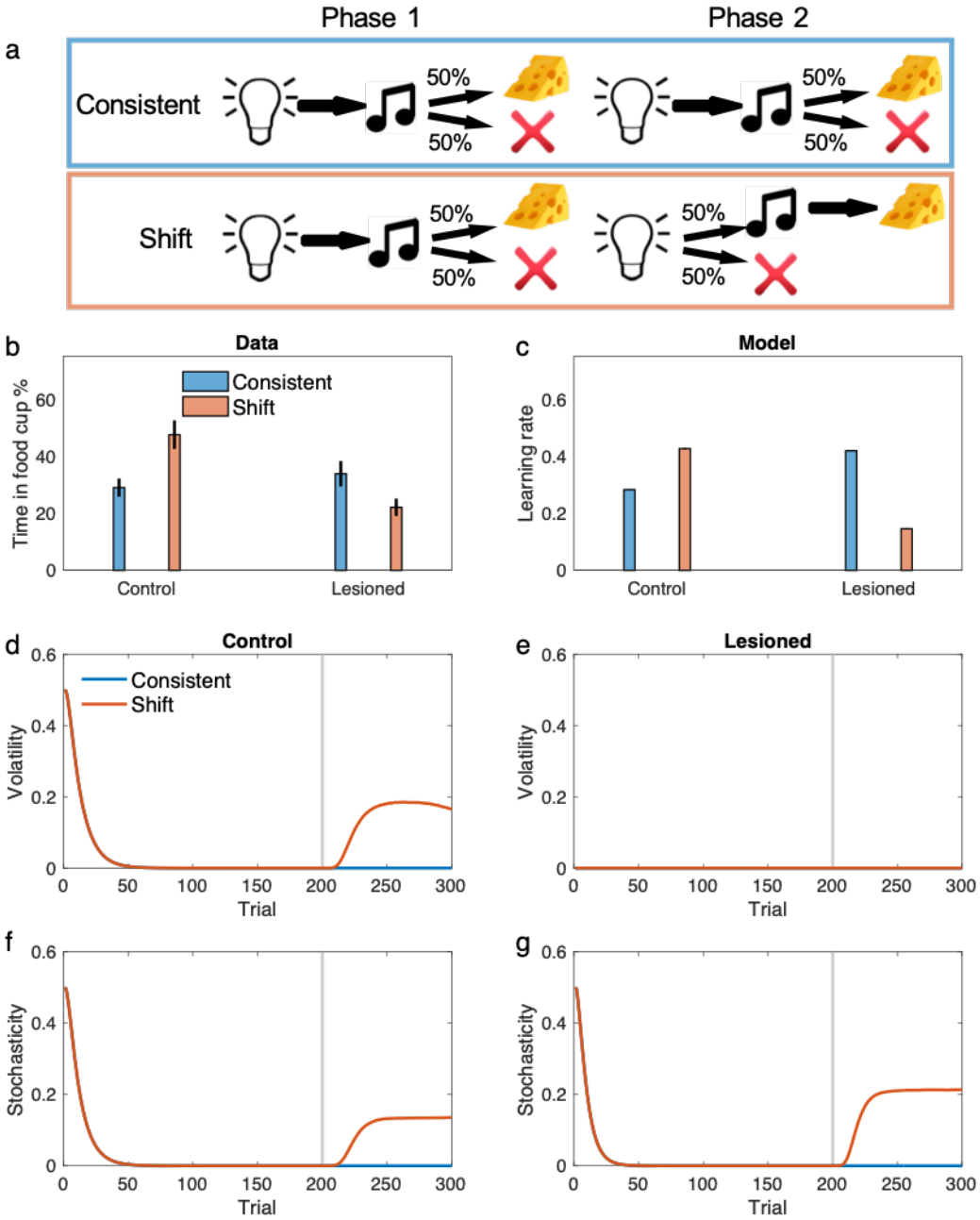
a
Phase 1
Phase 2
Consistent
Shift
50%
50%
b
Data
Consistent
Shift
Time in food cup %
Control
Lesioned
c
Model
Learning rate
Control
Lesioned
d
Control
Consistent
Shift
Volatility
Trial
e
Lesioned
Volatility
Trial
f
Stochasticity
Trial
g
Stochasticity
Trial

Fig 7

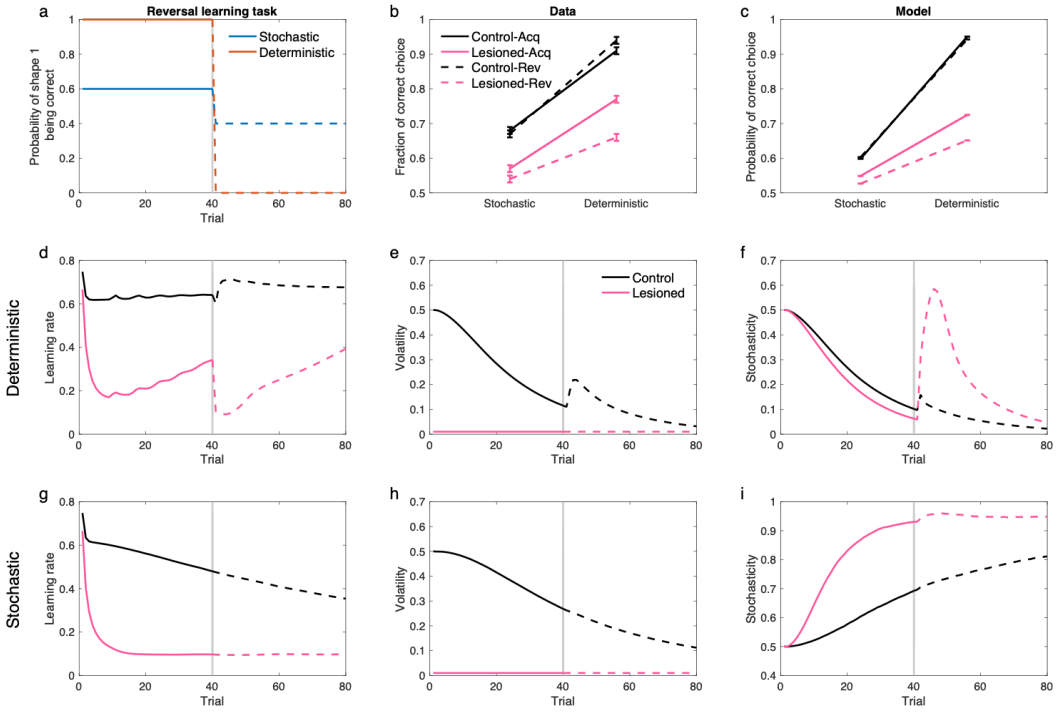
a
Reversal learning task
Stochastic
Deterministic
Probability of shape 1 being correct
Trial
b
Data
Control-Acq
Lesioned-Acq
Control-Rev
Lesioned-Rev
Fraction of correct choice
Stochastic
Deterministic
c
Model
Probability of correct choice
Stochastic
Deterministic
Deterministic
d
Learning rate
Trial
e
Control
Lesioned
Volatility
Trial
f
Stochasticity
Trial
Stochastic
g
Learning rate
Trial
h
Volatility
Trial
i
Stochasticity
Trial

Fig 8