Special Issue: Probabilistic models of cognition

Bayesian theories of conditioning in a changing world

Aaron C. Courville[1], Nathaniel D. Daw[2] and David S. Touretzky[3]

[1]Département d'Informatique et de Recherche Opérationnelle, Université de Montréal, Montréal QC H3C 3J7, Canada
[2]Gatsby Computational Neuroscience Unit, UCL, Alexandra House, 17 Queen Square, London, WC1N 3AR, UK
[3]Computer Science Department, Center for the Neural Basis of Cognition, Carnegie Mellon University, Pittsburgh, PA 15213, USA

The recent flowering of Bayesian approaches invites the re-examination of classic issues in behavior, even in areas as venerable as Pavlovian conditioning. A statistical account can offer a new, principled interpretation of behavior, and previous experiments and theories can inform many unexplored aspects of the Bayesian enterprise. Here we consider one such issue: the finding that surprising events provoke animals to learn faster. We suggest that, in a statistical account of conditioning, surprise signals change and therefore uncertainty and the need for new learning. We discuss inference in a world that changes and show how experimental results involving surprise can be interpreted from this perspective, and also how, thus understood, these phenomena help constrain statistical theories of animal and human learning.

Introduction

Classical conditioning experiments probe how subjects (typically animals) learn to predict biologically significant events such as food delivery. As with human learning experiments, which are often formally similar, there is a long history of quantitative models of conditioning [1–4], and a more recent interest in reframing them in explicitly statistical terms [5–11]. A statistical viewpoint can offer a principled explanation for why subjects display seemingly peculiar behavior; it can constrain or justify seemingly arbitrary aspects of the previous generation of models; and it can identify theoretically important features underlying a pattern of experimental data. This article focuses on one such feature in animal conditioning experiments: evidence that *surprise* can enhance the speed of learning [3].

We revisit this corner of the literature in this article because, we suggest, this phenomenon bears on a nexus of issues central to modern Bayesian accounts of learning (see also Conceptual Foundations Editorial by Chater, Tenenbaum and Yuille, and the Technical Introduction to this Special Issue; see Supplementary material online). Drawing on work by Dayan and collaborators [5,6,9], we suggest that surprise causes faster learning because it signals *change*, leading to increased *uncertainty* in one's beliefs –

Corresponding author: Courville Aaron C. (courvila@iro.umontreal.ca).

say, about the probability that a bell's ringing will be followed by food delivery. In Bayesian inference, the speed of learning (that is, the credence granted to new evidence over prior beliefs) turns on the reliability (or conversely, uncertainty) accorded to each. Change increases uncertainty, and speeds subsequent learning, by making old evidence less relevant to the present circumstances.

We begin by setting out Pearce's [3] theory of surprise in conditioning, and then consider its parallels with Bayesian inference in a changing world. With this framework in place, we review and reinterpret a number of experimental phenomena. Importantly, this exercise is informative not just about the experiments, but about the theories, because change is a relatively unexplored aspect of the Bayesian model space. For formal simplicity, most Bayesian theories of human [12] and animal [6,9,13] learning (including our own [10,11]) treat the world as unchanging – or at best, changing in a steady, rudimentary fashion. That animals react to surprise by modifying their subsequent learning suggests they are are, in effect, able to make more complicated inferences about when change has occurred [14] or how quickly it is occurring. By understanding what sorts of surprise affect learning and how, we can begin better to understand the structure of the statistical problem animals are solving. Such an enterprise opens new issues for future study in both human and animal learning contexts.

Theories of conditioning

In a classical conditioning experiment, affectively neutral stimuli (such as bells or lights, known as 'conditioned stimuli' or CSs) are repeatedly paired with biologically significant reinforcers (such as food or shock; 'unconditioned stimuli' or USs). Experiments study how animals' predictions of reinforcement develop with experience of various patterns of CS/US pairings; prediction is assessed via 'conditioned responses' (CRs, such as salivation or freezing), which are thought directly to reflect the expectation of a US. In the simplest case, a bell is repeatedly paired with food, and the animal gradually comes to salivate in response to the bell alone.

We focus on a series of classic findings suggesting that surprising events (such as a sudden change in the extent to which some stimulus predicts reinforcement) can speed subsequent predictive learning of the sort described

above. In fact, the experimental protocols in question are convoluted (they involve, at the least, multiple phases to first establish expectations, violate them to produce surprise, and then test subsequent learning in experimental animals compared with non-surprised controls). We therefore defer the details and first discuss two different sorts of abstract theories of the phenomena, associative and Bayesian. These will help to interpret the data we subsequently present, and at the same time raise questions that the data will help to address.

Associative learning approaches

Associative learning models [1–4] specify rules for the development of predictions given the stimuli presented. Typically, each stimulus is assumed to have a weight, or 'associative strength', characterizing how strongly it predicts reinforcement. When stimuli are presented in a trial, the amount of reinforcement expected (and, therefore, the degree of behavioral response such as salivation produced) depends on their associative strengths. The weights for the presented stimuli are updated based on the subsequent delivery or nondelivery of reinforcement.

As we have noted, animals respond to surprising events with faster learning. To capture this behavior, some associative learning models [2–4] endow each stimulus with a further attribute known as 'associability', which governs how quickly its associative strength is updated. In the Pearce–Hall [3] version of this idea (Box 1), the associability depends on how accurately the stimulus' associative strength has previously predicted reinforcement – with more accurate stimuli less susceptible to further learning. This feature allows 'surprising' reinforcement or non-reinforcement (such as that caused when a stimulus' predictive relationship with reinforcement abruptly changes) to produce faster learning. (Note that other attentional factors also impact associability [2,4], and these might also have a Bayesian interpretation [6]. Here we focus on surprise.)

This mechanism explains *how* some surprising events give rise to faster learning. To better understand the conceptual issues surrounding *why* they should do so, we will now consider a Bayesian perspective on learning in the presence of changing contingencies.

Bayesian accounts of conditioning and change

In a Bayesian model, Pearce and Hall's [3] concept of surprise can be related to a formal notion of *change*.

Bayesian approaches to conditioning [6,9–11] interpret animals' responding as arising from statistical reasoning about the likelihood of reinforcement, given their experience. Specifically, they take conditioned responding to reflect subjects' estimates of $P(US(t)|CS(t), data(1\ldots t–1))$, the probability of reinforcement given the currently observed stimuli and the training history, $data(1\ldots t–1)$, of conditioned stimuli and reinforcers seen on previous trials.

A standard Bayesian approach to reinforcer prediction would assume the animal begins with a 'world model' or skeletal description of the task contingencies (for instance, how observed stimuli and reinforcers may relate, how they may change over time). The model is specified only up to some set of unknown parameters, $\mathbf{w}$, (controlling, for instance, how probable it is that a bell will be followed by food); the job of the animal is to use experience with bells and food to estimate $\mathbf{w}$, and thus the probability of reinforcement in a particular situation. The learning rule for $\mathbf{w}$ is simply Bayes' rule; but different Bayesian accounts can differ in what sort of world model they assume, and therefore what sorts of regularities in their experience they can learn to capture. Because they focus on, or ignore, different aspects of an animal's experience, different world models can make very different experimental predictions. In the following, we consider different world models and how experimental data might help us to distinguish them.

World models: discriminative and generative

The types of world models used to describe classical conditioning may be grouped into two broad categories, the 'discriminative' and 'generative' approaches. These differ as to whether they attempt to learn about the inputs to a problem as well as the outputs. A discriminative approach to classical conditioning assumes animals model only $P(US(t)|CS(t), data(1\ldots t–1))$, the probability of reinforcement given the current conditioned stimuli. All other stimuli are treated as inputs and not explicitly modeled. The generative approach to classical conditioning assumes that animals learn to predict the full pattern of both stimuli and reinforcement $P(US(t),CS(t)|data(1\ldots t–1))$ rather than just reinforcement. Note that without additional information about stimulus delivery, $P(CS(t)|data(1\ldots t–1))$, it is not possible to reconstruct a generative model from a discriminative one.

We will interpret the conditioning data we present in terms of our latent variable model of conditioning [10,11], which takes a generative approach, assuming that animals learn to predict both conditioned and unconditioned stimuli. Our world model (Box 2) treats correlated patterns of stimuli and reinforcers as arising from unitary events, called 'latent causes', which are unobservable in themselves but whose occurrence may be inferred from the patterns of observable stimuli. A number of other approaches [6,7,9] instead assume a discriminative view.
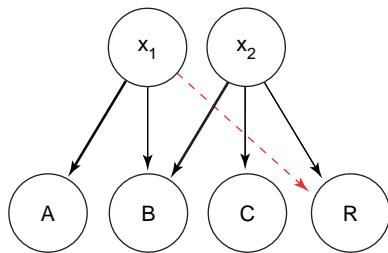
World models: change

Another important aspect of world models in classical conditioning is whether and how they incorporate the

---

Box 1. The Pearce–Hall model

The Pearce–Hall model [3] updates the associative strength V_i for each stimulus i present on trial t according to:

$$\Delta V_i(t) \propto \alpha_i(t)\lambda(t)$$

where $\lambda(t)$ is the magnitude of reinforcement delivered on the trial and $\alpha_i(t)$ is the associability of the stimulus.
  The associability is modulated by surprise, defined as:

$$\alpha_i(t) = |\lambda(t-1) - V_i(t-1)|$$

that is, the absolute value of the difference between the previous trial's actual reinforcement and that which had been predicted on the basis of the stimulus.

<table>
<tr><td>

Box 2. A latent variable model for classical conditioning

According to the latent variable theory [10,11], animals assume that patterns of observed events within their environment are attributable to a set of unobservable 'latent' causes. Learning is interpreted as an attempt to recover the parameters of the generative model that gave rise to the observed events.

In the theory, animals model the occurrence of binary valued causes $x(t)$ and observables $y(t)$ (both stimuli and reinforcers) on each trial. Causes are activated randomly on each trial, and each stimulus arrives (or doesn't) according to an independent flip of a biased coin. The bias is determined by the sum of weights, w_t, from all active causes to the stimulus, so that the activation of a cause promotes a particular pattern of observable stimuli (those to which it has positive weights).

Thus, in Figure I below, the activation of the hidden cause $x_1(t)$ promotes the appearance of stimuli A and B (thicker lines denote larger weights) and opposes (red dashed line for negative weight) reinforcement (R); the activation of the hidden cause $x_2(t)$ promotes B, C, and reinforcement. The weights therefore describe likely patterns of co-occurring stimuli and reinforcers, and so, given learned weights, the likelihood of reinforcement given some observed stimuli can easily be estimated. Additionally, we assume that the parameters w_t may change from trial to trial. The rate of change for the weight parameters is controlled by an additional parameter θ_t. Estimating the values of w_t therefore requires estimating the speed at which they are currently changing, which can also be done using Bayes' rule.

Figure I. Latent variable model (see text for details)

</td></tr>
</table>

possibility of change in the task contingencies (for instance, in the predictive relationship between a bell and food). An animal's beliefs about changing contingencies can be incorporated into the Bayesian perspective by specifying how the parameters **w** evolve over time. In particular, one can specify how the parameters in effect at time t, $\mathbf{w}_t$, relate to $\mathbf{w}_{t-1}$, the parameters at time $t-1$ (e.g. by a random walk) [9,13]. Of course, one might also specify that the parameters do not change (that $\mathbf{w}_t = \mathbf{w}_{t-1}$).

Animals working from different basic assumptions about how the world changes will behave differently as a result of drawing different statistical conclusions from their experience. For instance, if the model assumes that change occurs sporadically [14] or at different rates, then the standard Bayesian process of learning and prediction will implicitly involve drawing inferences about whether a change has occurred, or how quickly the world is currently changing. As detailed below, these inferences will affect the speed of learning, and it is this process we suggest corresponds to Pearce and Hall's notion of surprise [3]. On the other hand, if the model assumes no change [10,11] or

change at a constant rate (as in the Kalman filter model of Kakade and Dayan [9,13]), then evidence of change, in itself, will have no such effect.

For the following, we assume that the parameters of our generative model (Box 2), $\mathbf{w}_t$, change via a random diffusion process, and that the amount of change on each trial is controlled by parameters θ_t which are themselves unknown and changing. Inferring the parameters $\mathbf{w}_t$ involves inferring θ_t. This formulation therefore posits that animals learn, effectively, about the speed, θ_t, at which parameters change.

Change and uncertainty

A key concept in Bayesian inference is uncertainty. A Bayesian agent estimates not just parameters $\mathbf{w}_t$ from experience, but also its uncertainty (or lack of confidence) in these estimates. As originally noted by Dayan and Long [5], uncertainty about model parameters plays the same role as 'associability' in associative learning theories such as the Pearce–Hall model. Greater uncertainty about $\mathbf{w}_t$ grants increased influence to newly observed data when the two are combined, through Bayes' rule, to produce a new estimate. Analogously in associative learning models, increasing the associability weights current data more heavily.

The possibility that parameters have changed gives rise to uncertainty about them, because if they have changed, previous evidence is less informative about their present value. When rapid change occurs, it will produce surprising events (that is, events not well captured by the previously inferred values of the model parameters). For this reason, in a model that estimates the speed of change (Box 2), suprising events will be evidence of rapid change. This will lead to increased uncertainty and faster subsequent learning – analogous to the relationship between surprise and learning rate in the Pearce–Hall model. Thus key to this account is how events in the world affect the model's beliefs about the speed of change, which we examine and illustrate below.

Theories of conditioning: summary

We have reviewed associative and Bayesian viewpoints on classical conditioning, surprise, and change. In the following, we will use these theories (Box 2) to examine a number of empirical phenomena involving surprise and learning. This will serve to illustrate the ideas about change we have put forward; to suggest an interpretation of the data in terms of these statistical concepts; and, also, to investigate and constrain an appropriate Bayesian model of conditioning phenomena. On the last point, recall that Bayesian theories can differ in assuming different structure in their world model; for instance, generative versus discriminative models or different accounts of change. One way to get an empirical indication of what sort of information is represented in animals' world models is to understand what sorts of information affects their behavior. Therefore, in the following, we ask whether different sorts of surprise affect animals' behavior. If, for instance, surprising CSs (neutral stimuli) affect behavior, then that would support a generative model, which

incorporates such stimuli, over a discriminative model, which does not attempt to predict them.

Learning from surprising reinforcement

The core idea of the Pearce–Hall model (Box 1) is that surprising reinforcement or non-reinforcement gives rise to increased associabilities and faster learning. As we have discussed, a Bayesian perspective on this idea is that surprising events signal change, thereby increasing uncertainty and speeding learning. We now describe some experiments that demonstrate these effects.

The experimental designs are, necessarily, fairly complex, as the experiments involve several stages. The usual structure is to start with some protocol that would cause slow learning, and then show that this effect can be reversed by subsequently introducing some surprising change in contingencies, speeding learning.

Surprising reinforcers accelerate learning

To unpack a typical experiment, we start with the 'latent inhibition' (LI) effect [15] (see Table 1, Experiment 1): if a stimulus is repeatedly presented non-reinforced, prior to being paired with reinforcement, then acquisition of the stimulus-reinforcement relationship is slowed. Actually, latent inhibition can occur even if the stimulus preexposures are *reinforced* (WS: Weak Shock; see Table 1 acronyms) – in one experiment [16], a stimulus was repeatedly pre-exposed together with a small shock, which retarded acquisition when the stimulus was subsequently paired with a larger shock. But serially combining both sorts of preexposure (WS+LI) – either of which alone would slow subsequent acquisition – cancels their effect [3]. The interpretation is that the surprise caused by the change from weakly reinforced to non-reinforced preexposure speeds learning, reversing the slowing.

Surprise may also account for the reversal of another well-known effect that retards learning: blocking (Experiment 2 of Table 1). In blocking (BL), the acquisition of an association between a stimulus, A, and reinforcement is weakened ('blocked') if, on the training trials, presentations of A and reinforcement are accompanied by a second stimulus, B, that had already been associated with reinforcement. This effect can be reversed by a number of 'unblocking' treatments that introduce various surprising events, facilitating improved learning. In one version [19], the reinforcer is switched (between two equally aversive choices, footshock and ice-water dunking) during the blocking phase, and this surprising change reverses the blocking effect (UB).

Figures 1a and 1b illustrate the interpretation of this result in terms of Bayesian change detection. Here, the reinforcer switch causes faster change to be inferred, speeding learning and reversing the effect of blocking.

In short, the experiments discussed here suggest that animals can indeed detect surprising changes in reinforcer delivery and adjust their learning rate accordingly, as predicted by the Pearce–Hall model and also by Bayesian theories in which the occurrence or speed of change can be tracked. Other Bayesian theories, those that assume change is nonexistent or constant, would not produce such surprise-related effects. We now consider the behavioral effects of another sort of change, which suggest further constraints on the space of consistent Bayesian models.

Surprising neutral stimuli also accelerate learning

A further question is whether surprising neutral stimuli (CSs like lights or tones) can produce similar effects on the speed of learning as do surprising reinforcers. Such

Table 1. A summary of experiments that pertain to issues of surprise and change

Group	Phase 1	Interm	Phase 2	Rate of acquisition
Experiment 1: Latent inhibition and reinforced preexposure				
Control	–	–	$A+$	Fast
LI	$A-$	–	$A+$	Slow
WS	$A+$ (weak)	–	$A+$	Slow
WS+LI	$A+$ (weak)	$A-$	$A+$	Fast
Experiment 2: Unblocking by reinforcer switch				
Control	–	–	$A+$	Fast
BL	$B+_1$	–	$AB+_1$	Slow
UB	$B+_1$	–	$AB+_2$	Fast
Experiment 3: Latent inhibition and overshadowing				
Control	–	–	$A+$	Fast
LI	$A-$	–	$A+$	Slow
OV	–	–	$AB+$	Slow
LI+OV	$A-$	–	$AB+$	Fast
Experiment 4: Partial reinforcement extinction effect				**Rate of extinction**
EXT	$A+$	–	$A-$	Fast
PREE	$A+/A-$	–	$A-$	Slow

Experiment 1: Pre-exposure with unreinforced stimuli (A–) retards acquisition, a phenomenon known as latent inhibition (LI) [17]; pre-exposure of stimulus pairings with a weak shock (A+) also retards acquisition (WS) [16]. Surprisingly, the two pre-exposure effects cancel each other when combined (WS+LI) [18].
Experiment 2: Blocking (BL) is the retardation of learning that arises as a result of being reinforced in the presence of a pretrained CS (B). The blocking effect is significantly reduced (unblocking or UB) if the reinforcer is changed between pretraining (Phase 1) and the compound training (Phase 2). In one version of the experiment the reinforcer is changed from a footshock ($+_1$) to an ice-water dunking ($+_2$) [19].
Experiment 3: Overshadowing (OV) occurs when acquisition to one stimulus (A) is slowed by the presence of another simulus (B) during training. Whereas latent inhibition and overshadowing both cause slow learning on their own, once again the effects seem to counteract each other when combined (OV+LI) [20].
Experiment 4: Extinction (EXT) is the loss of a conditioned response to a stimulus (A) that occurs when a trained conditioned stimulus (A) is repeatedly presented without reinforcement (Phase 2: A–). If the reinforced presentations (Phase 1: A+) are interspersed with non-reinforced presentations (Phase 1: A–) (on 50% of trials), then subsequent extinction occurs more slowly, a phenomenon known as the partial reinforcement extinction effect (PREE).
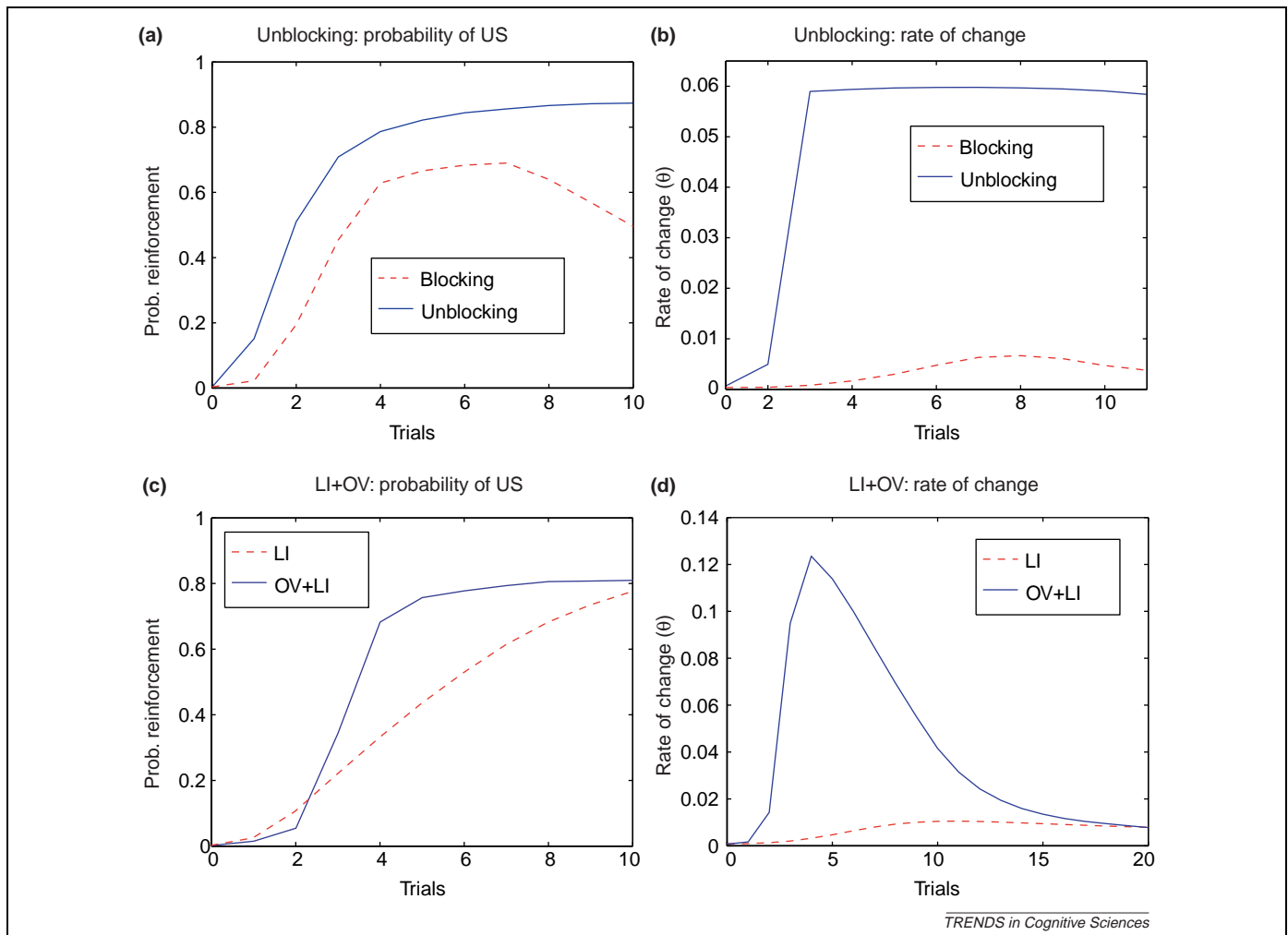
Figure 1. The probability of reinforcement (i.e. the modeled conditioned responding) and the rate of change estimated by our Bayesian model for Experiments 2 and 3 from Table 1. The schematic plots are derived from inference using the world model presented in Box 2 and observations drawn from the experimental designs. The modeled responses (a,c) qualitatively illustrate behavior patterns observed experimentally; inferences about the rate of change (b,d) help to explain why the learning curves differ between conditions. Experiment 2: unblocking with a qualitative change in reinforcement. **(a)** The estimated probability of reinforcement and **(b)** the estimated rate of change as a function of the number of Phase 2 'blocking' trials, for both the blocking and unblocking training schedules. Experiment 3: overshadowing counteracts latent inhibition. **(c)** The estimated probability of reinforcement and **(d)** the estimated rate of change as a function of the number of Phase 2 trials, either A+ trials in the case of latent inhibition (LI) or AB+ in the case of latent inhibition plus overshadowing (LI+OV).

an effect would not be captured by the Pearce–Hall model [3] (which computes surprise only in terms of expected reinforcement), nor for similar reasons by Bayesian theories that take a discriminative approach (that is, those that model only reinforcement but not neutral stimuli, and cannot therefore detect change in them).

Contrary to these accounts, such an effect may be demonstrated by another experiment in which two treatments that would slow conditioning counteract one another when combined (Experiment 3 of Table 1) [20]. Recall that in latent inhibition (LI), a stimulus, A, is pre-exposed (here, unreinforced), retarding acquisition when it is subseqently paired with reinforcement. Blaisdell and colleagues [20] demonstrate that this effect is reversed if a novel stimulus, B, is introduced during the reinforcement phase. Without the stimulus pre-exposure (which should retard learning), the inclusion of B (known as 'over-shadowing'; OV) would also retard learning [21], but, here again, the combination of treatments (LI+OV) reduces

the deficit. On the Bayesian analysis we have described, this is because of the surprising inclusion of neutral stimulus B, signaling an increased rate of change (Figures 1c and 1d), facilitating acquisition and counteracting latent inhibition.

There are a number of other demonstrations of surprising non-reinforcing stimuli affecting learning rates. In a serial conditioning task introduced by Wilson *et al.* [22] and subsequently used extensively by Holland and collaborators [23], a surprising change in how well a light predicts a neutral tone stimulus speeds subsequent conditioning.

That surprising stimuli should affect learning suggests, on a Bayesian account, that they must be part of the world model being learned. That is, such experiments seem to support generative models over discriminative models of conditioning. We finally consider another sort of surprise that is relevant to the structure of the Bayesian world model: surprise due to stochastic reinforcer delivery.
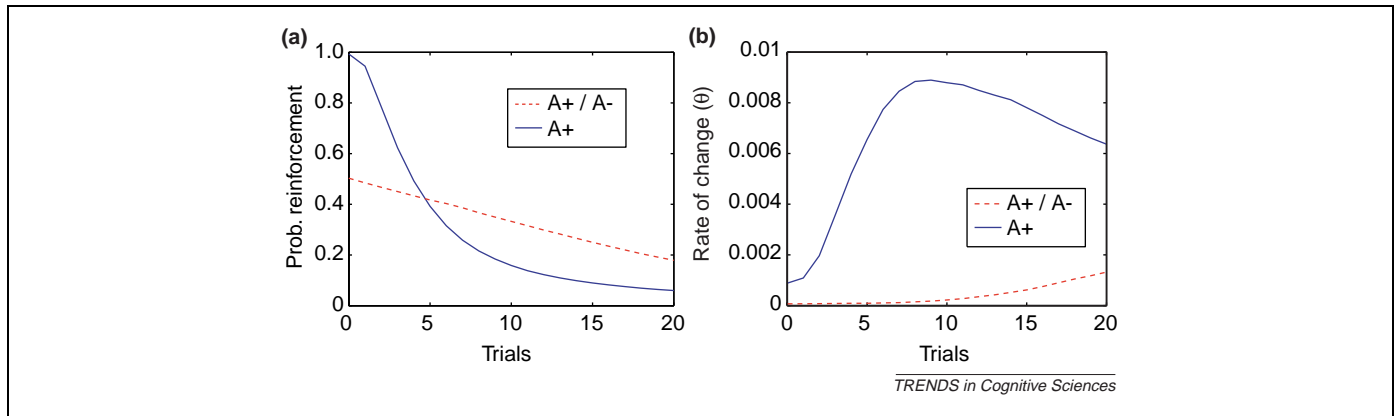
Figure 2. The partial reinforcement extinction effect: **(a)** the estimated probability of reinforcement and **(b)** the rate of change estimated during extinction trials ($A-$) for both a fully reinforced and partially reinforced stimulus A. The schematic plots are derived via inference over the world model presented in Box 2 using observations drawn from Experiment 4 of Table 1.

Fixed-rate stochastic reinforcement might not accelerate learning

Because the Pearce–Hall [3] model computes surprise as the mismatch between observed and predicted reinforcement (Box 1), it makes a characteristic prediction concerning experiments (known as 'partial reinforcement') in which a stimulus is reinforced randomly some fraction of the time. (For instance, reinforcement might be delivered, or not, according to a coin flip.) In this case, as reinforcement can never be exactly predicted, the model predicts constant surprise and persistent fast learning. By contrast, our Bayesian analysis predicts no such effect, because of the structure of the generative model we have assumed. That generative model assumes that reinforcer delivery will be stochastic, and attempts only to predict the *probability* of reinforcement. Surprise occurs when there is evidence of change in the probabilities. Random variation under constant probabilities (as with a sequence of coin flips) will not be surprising.

One test of these accounts is how quickly responding to partially reinforced stimuli extinguishes when reinforcement ceases. The Pearce–Hall model predicts that elevated associability should result in faster extinction for partially reinforced stimuli relative to fully reinforced stimuli. The opposite is true: responding to partially reinforced stimuli extinguishes more slowly [24–26]. This is known as the partial reinforcement extinction effect (PREE; Table 1, Experiment 4).

As first noted by Gallistel and Gibbon [7], the PREE is actually to be expected on a statistical account, because a downward change in the probability of reinforcement is harder to spot against a partially reinforced background. Figure 2a illustrates the idea in our Bayesian analysis. Here, when extinction begins, the absence of reinforcement is less surprising if reinforcer omissions were anyway expected owing to prior partial reinforcement. As a result, slower change is inferred, leading to less uncertain model weights, a smaller learning rate, and slower extinction.

In principle, the inferential effects we describe could be concealing additional effects of the learning rate of the sort predicted by Pearce–Hall. However, Gallistel and Gibbon

[7] replot PREE extinction times in terms of the number of expected reinforcements omitted (compensating, on their analysis, for the different amounts of experience needed to spot a change); thus transformed, the data still reveal no residual effect of Pearce–Hall style stochasticity sensitivity.

That said, it is clear that the PREE involves multiple interacting factors [27,28]; and a recent experiment [26] seems to reject the mathematical form, although not the basic insight, of Gallistel and Gibbon's [7] account. We also should note that there is some empirical support for the Pearce–Hall prediction of persistently elevated associability in partial reinforcement [29], but measuring associability using an orienting response (which is thought to reflect it), rather than through its signature effects on learning, on which we have concentrated here. Further experiments are required to unravel fully the effects of stochasticity on uncertainty and associability, and their differential effects on behavior. However, provisionally, the PREE suggests that an appropriate Bayesian world model is one that predicts the probability, rather than the absolute occurrence, of reinforcement.

Conclusions

We have presented a Bayesian perspective on the issues surrounding animal reasoning with respect to change in the environment. The idea – that changing contingencies lead to uncertainty about reinforcement probabilities, which in turn leads to faster learning – provides a tentative but parsimonious explanation for why animals behave as they do in the presence of change (see also Box 3 for future research areas). This viewpoint also puts inference about change at the center of the learning problem. Bayesian learning in a changing world involves rich computational and inferential issues that have not been fully explored either experimentally or theoretically. (see also Editorial 'Where next?' in this issue). Work on change in Bayesian models has been largely confined to the animal learning literature; it will be of particular interest to investigate the same issues in the context of

human learning, which has recently featured very much parallel models and tasks.

Supplementary data

Supplementary data associated with this article can be found at doi:10.1016/j.tics.2006.05.004

References

1 Rescorla, R.A. and Wagner, A.R. (1972) A theory of Pavlovian conditioning: The effectiveness of reinforcement and non-reinforcement. In *Classical Conditioning, 2: Current Research and Theory* (Black, A.H. and Prokasy, W.F., eds), pp. 64–69, Appleton-Century-Crofts

2 Mackintosh, N.J. (1975) A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychol. Rev.* 82, 532–552

3 Pearce, J.M. and Hall, G. (1980) A model for Pavlovian learning: Variation in the effectiveness of conditioned but not unconditioned stimuli. *Psychol. Rev.* 87, 532–552

4 Le Pelley, M.E. (2004) 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

5 Dayan, P. and Long, T. (1998) Statistical models of conditioning. In *Advances in Neural Information Processing Systems Vol. 10* (Kearns M.J. *et al.*, eds), pages 117–123, MIT Press

6 Dayan, P. *et al.* (2000) Learning and selective attention. *Nat. Neurosci.* 3, 1218–1223

7 Gallistel, C.R. and Gibbon, J. (2000) Time, rate and conditioning. *Psychol. Rev.* 107, 289–344

8 Courville, A.C. and Touretzky, D.S. (2001) Modeling temporal structure in classical conditioning. In *Advances in Neural Information Processing Systems Vol. 14* (Dietterich, T.G. *et al.*, eds), pp. 3–10, MIT Press

9 Kakade, S. and Dayan, P. (2002) Acquisition and extinction in autoshaping. *Psychol. Rev.* 109, 533–544

10 Courville, A.C. *et al.* (2003) Model uncertainty in classical conditioning. In *Advances in Neural Information Processing Systems Vol. 16* (Thrun, S. *et al.*, eds), pp. 977–984, MIT Press

11 Courville, A.C. *et al.* (2004) Similarity and discrimination in classical conditioning: A latent variable account. In *Advances in Neural Information Processing Systems Vol. 17* (Saul, L.K. *et al.*, eds), pp. 313–320, MIT Press

12 Tenenbaum, J.B. and Griffiths, T.L. (2001) Structure learning in human causal induction. In *Advances in Neural Information Processing Systems Vol. 13* (Leen, T.K. *et al.*, eds), pp. 59–65, MIT Press

13 Dayan, P. Kakade, S. *et al.* (2001) Explaining away in weight space. In *Advances in Neural Information Processing Systems Vol. 13* (Leen, T.K. *et al.*, eds), pp. 451–457, MIT Press

14 Dayan, P. and Yu, A. (2001) ACh, uncertainty, and cortical inference. In *Advances in Neural Information Processing Systems Vol. 14* (Dietterich, T.G. *et al.*, eds), pp. 189–196, MIT Press

15 Lubow, R.E. (1973) Latent inhibition. *Psychol. Bull.* 79, 398–407

16 Hall, G. and Pearce, J.M. (1979) Latent inhibition of a CS during CS–US pairings. *J. Exp. Psychol. Anim. Behav. Process.* 5, 31–42

17 Lubow, R.E. and Moore, A.U. (1959) Latent inhibition: the effect of nonreinforced preexposure to the conditional stimulus. *J. Comp. Physiol. Psychol.* 52, 415–419

18 Hall, G. and Pearce, J.M. (1982) Restoring the associability of a preexposed CS by a surprising event. *Q. J. Exp. Psychol. B.* 34, 127–140

19 Blaisdell, A.P. *et al.* (1997) Unblocking with qualitative change of unconditioned stimulus. *Learn. Motiv.* 28, 268–279

20 Blaisdell, A.P. *et al.* (1998) Overshadowing and latent inhibition counteract each other: Support for the comparator hypothesis. *J. Exp. Psychol. Anim. Behav. Process.* 24, 335–351

21 Pavlov, I.P. (1927) *Conditioned Reflexes: An Investigation of the Physiological Activity of the Cerebral Cortex*, Oxford University Press

22 Wilson, P.N. *et al.* (1992) Restoration of the orienting response to a light by a change in its predictive accuracy. *Q. J. Exp. Psychol.* 44B, 17–36

23 Holland, P.C. (1997) Brain mechanisms for changes in processing of conditioned stimuli in Pavlovian conditioning: Implications for behavior theory. *Anim. Learn. Behav.* 25, 373–399

24 Gibbon, J. *et al.* (1980) Partial reinforcement in autoshaping with pigeons. *Anim. Learn. Behav.* 8, 45–59

25 Rescorla, R.A. (1999) Within-subject partial reinforcement extinction effect in autoshaping. *Q. J. Exp. Psychol. B.* 52, 75–87

26 Haselgrove, M. *et al.* (2004) A partial reinforcement extinction effect despite equal rates of reinforcement during Pavlovian conditioning. *J. Exp. Psychol. Anim. Behav. Process.* 30, 240–250

27 Pearce, J.M. *et al.* (1997) Partial reinforcement in appetitive Pavlovian conditioning with rats. *Q. J. Exp. Psychol. B.* 50, 273–294

28 Rescorla, R.A. (1999) Partial reinforcement reduces the associative change produced by nonreinforcement. *J. Exp. Psychol. Anim. Behav. Process.* 25, 403–414

29 Kaye, H. and Pearce, J.M. (1984) The strength of the orienting response during Pavlovian conditioning. *J. Exp. Psychol. Anim. Behav. Process.* 10, 90–109

30 Yu, A.J. and Dayan, P. (2005) Uncertainty, neuromodulation, and attention. *Neuron* 46, 681–692

31 Gallistel, C.R. *et al.* (2001) The rat approximates an ideal detector of changes in rates of reward: Implications for the law of effect. *J. Exp. Psychol. Anim. Behav. Process.* 27, 354–372

32 Kurth-Nelson, Z. and Redish, A.D. (2004) μ-agents: Action-selection in temporally dependent phenomena using temporal difference learning over a collective belief structure. *Soc. Neurosci. Abstr.* 30, 2071

33 Gallistel, C.R. *et al.* (2004) The learning curve: Implications of a quantitative analysis. *Proc. Natl. Acad. Sci. U. S. A.* 101, 13124–13131