Published in final edited form as:

Anim Cogn. 2017 September ; 20(5): 867–880. doi:10.1007/s10071-017-1107-5.

Sure enough: efficient Bayesian learning and choice

Brad R Foley and

The Department of Molecular and Computational Biology, The University of Southern California.
1050 Childs Way, Los Angeles, California. 90089

Paul Marjoram

Preventative Medicine, Keck School of Medicine, The University of Southern California. 2001 N.
Soto Street, Los Angeles, California. 90032

Abstract

Probabilistic decision-making is a general phenomenon in animal behaviour, and has often been interpreted to reflect the relative certainty of animals' beliefs. Extensive neurological and behavioral results increasingly suggest that animal beliefs may be represented as probability distributions, with explicit accounting of uncertainty. Accordingly, we develop a model that describes decision-making in a manner consistent with this understanding of neuronal function in learning and conditioning. This first-order Markov, recursive Bayesian algorithm is as parsimonious as its minimalist point-estimate, Rescorla-Wagner analogue. We show that the Bayesian algorithm can reproduce naturalistic patterns of probabilistic foraging, in simulations of an experiment in bumblebees. We go on to show that the Bayesian algorithm can efficiently describe the behavior of several heuristic models of decision-making, and is consistent with the ubiquitous variation in choice that we observe within and between individuals in implementing heuristic decision-making. By describing learning and decision-making in a single Bayesian framework, we believe we can realistically unify descriptions of behaviour across contexts and organisms. A unified cognitive model of this kind may facilitate descriptions of behavioural evolution.

Keywords

Bayesian; learning; decision-making; uncertainty; foraging

1 Introduction

It is a truth so obvious that it is effectively invisible, that animals search, and choose among available options probabilistically, rather than deterministically. This includes animals as diverse as fruit flies, nematodes, guppies and bees (Foley et al. 2015; Saltz 2011; Kauffman et al. 2011; Abrahams 1989; Real 1981). Even in the face of strong flocking instincts,

Tel.: +1-213-740-2410, Fax: +1-213-740-1109, bradfole@usc.edu.

6 Compliance with ethical standards

Conflict of interest: Author Foley declares that he has no conflict of interest. Author Marjoram declares that he has no conflict of interest.

Ethical Approval: This article does not contain any studies with human participants or animals performed by any of the authors

animals distribute themselves among available options (Michelena et al. 2009; Jovani and Mavor. 2011) instead of converging on a single “best” option. Probabilistic behaviour might be a simple result of environmental or behavioural stochasticity, but it may also reflect features of the animals’ cognition. Given that variability is inevitable in a noisy world, we might expect well-adapted organisms to have mechanisms to account for probabilistic outcomes (Chater et al. 2010). Experiments have shown that animals learn, and make choices, based on the strength of their beliefs, the reliability of cues, and the likelihood of outcomes (Yu and Dayan 2005; Rushworth and Behrens 2008; Körding 2007; Gershman 2015; Knill and Pouget 2004; Real 1981). Animals, then, incorporate uncertainty into many important decisions. Uncertainty may be incorporated into the learning process at the neurological level (Roesch et al. 2012; Pouget et al. 2013), and the decision-making level. Models that explicitly account for uncertainty in decision-making and learning, will potentially be important tools.

While the importance of reasoning with uncertainty is generally appreciated, it is most common to model belief using point estimates (Roche et al. 1996)—for instance, by estimating an “expected mean reward”—and the most commonly utilised learning models, e.g. the Rescorla Wagner model, treat belief as a point estimate (e.g. Roesch et al. 2012). Binary choice can be modelled with point estimate models, by comparing belief-values and adding an extrinsic noise component (Stephens, 1985) to produce probabilistic behaviour (Stephens 1985; Biernaskie et al. 2009; Real 1981). But we can, instead, model uncertainty explicitly, by treating beliefs as *probability distributions* with some measure of centrality (for instance, an expected mean reward), and a variance around that value, to explicitly model the uncertainty associated with the belief (reviewed in Ma and Jazayeri 2014). We will follow Ma and Jazayeri (2014) and define an animal’s “belief” here, and throughout, as some internal representation of the state of the external world. We define “uncertainty” as the representation of a probability distribution around that *belief*, as opposed to sensory noise, external reward variability, or behavioural stochasticity. We do not mean to imply conscious awareness by these terms. We also note that, by these general definitions, belief and uncertainty may be fully encoded in very simple systems (Pouget et al. 2013; Roesch et al. 2012; Valone 2006; Griffiths et al. 2012).

One of the most common frameworks for explicitly modeling belief-updating while incorporating uncertainty is Bayesian reasoning. In the Bayesian approach, belief is represented as a *probability distribution*, rather than a point estimate (Körding 2007; Chater et al. 2010). Beliefs are then updated according to the support provided by new information. As such, uncertainty is central to modeling learning in a Bayesian framework—the more uncertain an animal is about its current beliefs, and the more it trusts incoming information, the faster the animal should learn. In behaviour, Bayesian cognitive models have primarily been applied in the context of two classes of optimal foraging problems—patch-switching with resource depletion (Iwasa et al. 1981; Olsson and Holmgren 1998; Olsson and Brown 2006; Green 2006), and information gathering before making a best-choice (Luttbeg 1996; Luttbeg and Warner 1999). In particular, the Comparative Bayes model of Luttbeg (1996) posits a recursive Bayesian updating mechanism for learning, that relies on repeated exploration among uncertain options, followed by a final choice, once uncertainty has been sufficiently reduced. This model has been shown to plausibly explain behaviours as complex

as evaluating prospective home patches (Selonen and Hanski 2010), food items (Luttbeg and Langen 2004), and mates (Luttbeg 1996); and solves the problem of assuming that animals *a priori* know the value and distribution of potential rewards in the environment (Real 1990).

Here, we develop a more general Bayesian model of learning than these previous foraging models, based on the Rescorla Wagner (RW) model of conditioned response learning. We propose a single additional step in addition to Bayesian updating to account for residual uncertainty, similar to the way that asymptotic variance in certainty is modelled in other conditioning paradigms (Kakade and Dayan 2002; Yi 2007; Gibbon et al. 1984), consistent with the frequent observation that animals (including primates) typically do not converge on complete certainty in their beliefs even after extended reinforcement, whether examining conditioned responses (Reger et al. 2012), or neural correlates of conditioning and uncertainty (Rushworth and Behrens 2008). This additional step may also model “forgetting”, that is the attenuation of conditioning, or extinction, with time (Kraemer and Golding 1997; White 2001; Rescorla 2004). We will then demonstrate how this very general cognitive model can lead, with minimal additional assumptions, to a diversity of realistic behaviours and behavioural strategies. Belief models that incorporate uncertainty promise to clarify diverse puzzles in decision-making, such as systematic “mistakes” in foraging (Janmaat et al. 2013); and speed-accuracy tradeoffs (Liu and Watanabe 2012). They may potentially unify disparate decision-making heuristics under a single, plausible, cognitive framework (for instance, see the diffusion models of Pelé and Sueur 2013).

This model is not meant to be a normative model of how animals “should” make decisions, from a rationalist perspective; i.e. we are not attempting a “Bayesian Fundamentalist” argument (Jones and Love 2011). Nor is it meant to model optimal decision-making, from the perspective of perfect knowledge of the external world. It is a descriptive model, based on current, empirical observations of learning and conditioning. We model animal decision-making, emerging from the set of beliefs (with uncertainty) of a very simple animal. We then explore the performance of our model in an illustrative case, using the bee foraging data of Cnaani et al. (2006), to establish whether this model can plausibly explain real behaviour.

2 Models

2.1 Case study in bee foraging

Here, we describe the way an animal simultaneously learns about and exploits its environment. This allows us to model how moment-to-moment changes in belief translate into behaviour. The processes of exploitation (choice) vs exploration (learning) are sometimes partitioned in models of behavior (Kaelbling et al. 1996; Oster and Heinrich 1976; Luttbeg 1996), but it is clear that both can co-occur (Krakauer and Rodríguez-Gironés 1995). Learning with decision-making is often studied in the context of foraging (Roche et al. 1996). Accordingly, as our example dataset, we will refer to a study of bee foraging: floral learning and exploitation (Cnaani et al. 2006). Bees are a classic example of foraging (Greggers and Menzel 1993; Real 1981), in changing environments (Oster and Heinrich 1976), in a Bayesian manner (Biernaskie et al. 2009). We emphasize that our example has an entirely different structure from that found in the more commonly studied optimal foraging models with resource depletion. It is also a simpler scenario than optimum models for

learning and exploiting resources in changing environments (Biernaskie et al. 2009; Wilson et al. 2010; Gershman et al. 2010).

In the study of Cnaani et al. (2006), researchers manipulated energetic rewards, with two different signals of nectar quality associated with odor cues. In one manipulation, higher energetic reward was associated with increased nectar *concentration*. In the other, higher energetic reward was associated with increased nectar *volume*. A third manipulation (their “reciprocal” treatment, which we will not consider) kept the energetic reward constant, but varied the concentration and volume associated with flower odors.

The bee behavior illustrates several very common patterns of animal learning. In each case, bees quickly learned to identify the most rewarding cues. Consistent with earlier studies (Real 1981; Real et al. 1990), the pattern of learning showed three distinct features. First, learning appeared to follow an S-shaped distribution in which learning proceeded rapidly when the reward values initially changed, and finally leveled off. Second, bees never entirely stopped foraging from both options. Even when one option was clearly better than the other, they foraged probabilistically among both options (the final asymptotes in Fig. 1 range from 0.023 to 0.225). Third, different cues of floral quality resulted in very different levels of learning (equilibrium foraging preference)—bees learned most efficiently when high nectar concentration, rather than high volume, signaled the best reward. This was true even in reciprocal volume-concentration manipulations, where Net Energy Gain calculations suggested the most valuable reward was the high-volume, low-concentration nectar (Cnaani et al. 2006).

2.2 Learning: Rescorla Wagner

There are a number of point-estimate models of animal associative learning, (c.f. Kaelbling et al. 1996). In particular, delta models are well-supported by behavioral and neurological evidence (e.g. Roesch et al. 2012). In a delta model, learning occurs when there is a difference between the observed and expected association for an unconditioned stimulus and a conditioned cue. This difference is the “delta” of the delta model. The most common delta model is the Rescorla-Wagner (RW). RW models the strengthening of an association, or weight, w_i (see Table 1), between an unconditioned stimulus, or reward, x , and a conditioned stimulus (or cue, c_i) of type i . Here, c_i can be 0 or 1 (present or absent). An unconditioned stimulus might be a reward, such as sugar, and the cue might be a floral character, like color or odor. The change in association $w_{i,t}$ between the cue and the reward at time t , is proportional to the difference between the anticipated result, $a_{i,t}$ and the measured value, $x_{i,t}$:

$$\Delta w_{i,t} = \alpha_i (x_{i,t} - a_{i,t}) \cdot c_{i,t} \quad (1)$$

The term α_i denotes the salience of the cue of type i . In this system we are assuming the cue and the reward have a one to one correspondence. In general, there may be multiple cues associated with one or more outcomes, in which case we would use a more general form for Eq 1. Typically the RW equation contains a second multiplier—a learning rate parameter. However, we are not modeling differences in learning rate among individuals here, and we

will ignore it. The anticipated outcome, $a_{i,t}$ is simply the product of the cue presence or absence, and its weight:

$$\alpha_{i,t} = w_{i,t} c_{i,t}. \quad (2)$$

Learning under RW can be described as an process of iterative updating of w_i such that

$$w_{i,t+1} = w_{i,t} + \Delta w_{i,t}. \quad (3)$$

The RW equation captures many of the general properties of reinforcement learning, at the organismal and neuronal level (Schultz and Dickinson 2000); however, the anticipated outcome is described by a point estimate. This results in a number of well known problems, including how to understand blocking, overshadowing (Matzel et al. 1985), relapse after extinction (Bouton 2002; Dunsmoor et al. 2015), instrumental transfer (Cartoni et al. 2013), and facilitated relapse after extinction—several problems that are explicable if we model reinforcement in a Bayesian framework (Kruschke 2008; Maia 2009; Gershman 2015; Courville et al. 2006). Moreover, point estimate models of belief are inconsistent with several suggestive findings in psychology. For instance, primates are known to code uncertainty in specialised brain regions (Behrens et al. 2007), and to evaluate joint-probability in a manner consistent with Bayesian reasoning with uncertainty (Yang and Shadlen 2007). Additionally, many organisms are known to weight cues by their reliability (e.g. Vossel et al. 2014; Yu and Dayan 2005; Yang and Shadlen 2007; Fischer and Peña, 2011; Wystrach et al 2015), in a manner compatible with Bayesian reasoning.

While it may be possible to describe many of these phenomena using various point estimate models, they are natural outcomes of Bayesian cognition. An efficient Bayesian analogue of RW reinforcement may then be a useful tool for modelling learning and decision-making in an integrated likelihood framework.

2.3 Learning: Bayesian Reasoning

Similar to RW, Bayesian updating is a process of comparing prior beliefs to the results of ongoing experiences, and updating these beliefs. And also similar to RW, in Bayesian learning models, beliefs shift most rapidly when observed outcomes are very different from anticipated outcomes (Courville et al 2006). However, Bayesian approaches differ from RW in two major ways. Under Bayesian reasoning, we represent a belief as a *posterior probability distribution*, rather than a point estimate. Also, posterior distributions are estimated by updating explicit *prior* distributions, which reflect initial beliefs (Körding 2007; Knill and Pouget 2004; Gershman 2015)—these priors may simply be learned, or evolved preferences (McNamara et al. 2006). In practical terms, Bayesian models can be seen as recasting RW, with uncertainty playing the role of the term α_i in Eq. 1, and the prior taking the role of $w_{i,t=0}$ (Schultz and Dickinson 2000, Kruschke 2008, and see Table 1). We note that the formulation of the following recursive Bayesian analogue of RW is similar to the recursive updating component of the algorithm of Luttbegg (1996), but more general in scope.

Using Bayes' theorem, we are interested in estimating the probability distribution of a parameter of interest, θ_i —that is, the probability of observing values, θ , associated with some cue of type i (Table 1). We use a set of observations, x_i , which are taken to be sampled from the true distribution of θ_i to refine our estimate, $P(\theta_i/x_i)$. Formally, Bayes' theorem is written as:

$$\overbrace{P(\theta_i|x_i)}^{\text{posterior}} = \frac{\overbrace{P(x_i|\theta_i)P(\theta_i)}^{\text{likelihood prior}}}{P(x_i)}. \quad (4)$$

Here the term $P(\theta_i)$ is the prior distribution, that is, the initial believability of values of θ for cue i . For instance, in the context of our example, the distribution $P(\theta_i)$ would reflect the bee's initial belief about the likely value of geranium as a food source, where θ_i represents the unknown "true" reward provided by geraniums. After visiting a flower, and measuring a reward, the likelihood distribution, $P(x_i/\theta_i)$, represents the likelihood that we would observe the value x_i for a given value of θ_i ; that is, the likelihood that a reward of value x_i is detected if the flower with cue i has "quality" θ_i . There is an additional normalizing constant, $P(x_i)$, that ensures the probability distribution integrates to 1. When comparing the likelihood ratio of two models, this constant conveniently cancels out. If we need to solve for $P(x_i)$, we can do so by noting that the probability density must integrate to 1.

We are modeling learning as a process of recursive Bayesian estimation. Thus, the prior for the next time step, $P_{t+1}(\theta_i)$, is the posterior of the current time step. We update conditionally, if a cue is sampled, based on the reward, x , associated with cue i at time t , $x_{i,t}$:

$$P_{t+1}(\theta_i) = P_t(\theta_i|x_{i,t}) \propto \begin{cases} P(x_{i,t}|\theta_i)P_t(\theta_i), & \text{if } c_i=1 \\ P_{t-1}(\theta_i|x_{i,t-1}), & \text{if } c_i=0 \end{cases} \quad (5)$$

Here, c_i is 1 if cue i is sampled at time t ; it is 0 otherwise, when no updating of the posterior associated with cue i occurs—that is, beliefs regarding the value of θ_i stay constant when the flower is not visited, and are equal to their value at $t-1$. We ignore the normalizing constant here, because it will cancel out in the process of model comparison (below).

While beliefs and likelihoods can be modeled using any distribution function, for the purposes of our illustrative example we will assume that the prior and likelihood distributions are Gaussian. The Gaussian is a convenient choice for a probability density function in Bayesian applications for several reasons, including tractability. A Gaussian can be completely described by two terms—a mean, μ , and a variance, σ^2 . Then, we can describe the prior as:

$$P_0(\theta_i) = N(\mu_{i,0}, \sigma_{i,0}^2), \quad (6)$$

where $\mu_{i,0}$ represents the mean of our initial belief in the true value of θ_i and $\sigma_{i,0}^2$ is the initial variance around that estimate. If, also for convenience, we assume that the likelihood distribution is also a normal distribution, we can define it as a function of the observation associated with cue i at time t , $x_{i,t}$, and a variance, v_i^2 . The probability of observing a value of $x_{i,t}$ given a normal distribution denoted with mean θ_i and variance v_i^2 is:

$$P(x_{i,t}|\theta_i) = N(x_{i,t}|\theta_i, v_i^2). \quad (7)$$

Bayesian updating is in no way dependent on beliefs or likelihood distributions being represented as Gaussian. Other distributions may be modeled, for instance with a simple transformation (e.g. a log-normal distribution). Bayesian updating with other common distributions, such as the binomial, gamma, or exponential have well worked-out mathematical solutions (Gelman et al. 2013). Moreover, a large number of non-parametric solutions exist for nonlinear distributions—however, these may be computationally expensive.

It is mathematically simple to calculate a posterior distribution, given normal prior and likelihood distributions. If we index time, defined as the number of visits, by t , (where the probability at time $t=0$ is the prior), we define the posterior at time $t>0$ by

$P_t(\theta_i|x_t) = N(\mu_{i,t}, \sigma_{i,t}^2)$. Given the definitions of the prior and likelihood distributions above, the mean and variance of the posterior will be:

$$\mu_{i,t} = \frac{\mu_{i,t-1}/\sigma_{i,t-1}^2 + x_{i,t}/v_i}{1/\sigma_{i,t-1}^2 + 1/v_i^2}, \quad \sigma_{i,t}^2 = (1/\sigma_{i,t-1}^2 + 1/v_i^2)^{-1}. \quad (8)$$

Belief in θ_i is updated as an inverse-weighted mean of the variances of the prior and the likelihood. Beliefs about the value of θ_i will be narrowed down quickly when the reliability of a cue is high (i.e. v_i^2 is small). In the case of the bees of Cnaani et al. (2006), learning was efficient when flower type was associated with changes in nectar *concentration*, but not nectar *volume*, to the point that bees nearly ignored volume changes when given nectar of different concentrations, even when energetic calculations suggested the high volume nectar was much more rewarding. This suggests that the bees may have been more certain of their estimates of flower quality when distinguishing on the basis of nectar concentration than volume (i.e. the v_i^2 associated with the cue of nectar concentration is lower). In reality, animals are known to jointly estimate the cue specific expected mean reward, and the variance (reliability) of the cues (Real 1981; Yang and Shadlen 2007; Selonen and Hanski 2010; Luttbegg 1996). This variance is likely to reflect a number of underlying factors, including empirical variance, sensory limitations, and memory limitations (Ma and Jazayeri 2014; Vossel et al. 2014). We will assume for simplicity, here, that the bees have already estimated v_i^2 , and that this estimate does not change appreciably over the course of the experiment.

According to Eq 5, an animal's uncertainty, ($\sigma_{i,t}^2$), associated with the posterior, $P_t(\theta_i/x_t)$, will decrease with each sampling event. In particular, in a stable environment, the mean of the posterior will converge to the sampled mean of θ_i . The variance of that belief will converge monotonically to 0 (certainty). Thus, a bee will be able to forage more efficiently with time, increasingly choosing the best available option. Note, of course, that $\sigma_{i,t}^2=0$ corresponds with a complete certainty in the value of the “expected reward”.

Evidence—including in Cnaani et al. (2006), and other results in foraging (Oster and Heinrich 1976), and the limits of conditioning (Kakade and Dayan 2002; Yi 2007; Gibbon et al. 1984)—suggests that animals rarely converge on complete certainty in their beliefs. Additional phenomena in animal learning, such as forgetting (Fusi et al. 2007) similarly suggest that cumulative uncertainty may be incorporated into animals' beliefs (Menzel 1993; Devenport et al. 1997). For example, in mice (Dunsmoor et al. 2015), fear responses to a conditioned stimulus return after extinction training, but in a more sporadic and attenuated manner. This is typically interpreted as a result of weakened inhibition (Rescorla 2004), but may also be represented as an increased uncertainty about whether the conditioned stimulus is, or is not, associated with an aversive cue after a long absence. Generally speaking, this makes behavioral sense, since the world is not static (Oster and Heinrich 1976). This residual “estimation uncertainty” is thought to play an important role in decision-making, including a role in exploratory behavior (Payzan-LeNestour and Bossaerts 2011). One point-estimate model used to describe this process is The Temporal Weighting Rule (TWR: Devenport 1998; Devenport et al. 1997), which describes the observable decay of the effect of the most recent conditioning events on belief. In the TWR framework, the effect of conditioning events on belief are weighted by their recency, with a value of $1/t$, where t is time. This is roughly equivalent to saying that animals are more certain of recent stimuli.

2.4 Learning: Residual Uncertainty

An asymptotic limit to learning is typically observed (e.g. Reger et al. 2012). Any realistic model of animal cognition needs to account for residual uncertainty, even after continued reinforcement (Dayan et al. 2000). Moreover, the processes of forgetting—i.e. a weakening of unreinforced associations with time (Kraemer and Golding 1997; White 2001; McNamara and Houston 1987; Luttbeg and Warner 1999)—and spontaneous de-extinction (Rescorla 2004), demonstrate that animals become more uncertain of their beliefs when they are not actively reinforced. We will call these processes collectively “uncertainty inflation”. Here, we will model these processes as if they are one process.

While some implementations of Bayesian inference, including the Kalman filter, can implicitly account for patterns of uncertainty inflation (Kruschke 2008), it is more common to model uncertainty inflation explicitly (Dayan et al. 2000). For instance, a likelihood distribution might be calculated on a sliding window of samples, and the original prior updated at each time step (rather than indefinite iterative updating; Kakade and Dayan 2002). Or, as above, we might use the window of samples from the current time t for some window size n , ($x_{i,t}$, $t \in (t - n, t)$), and calculate the uncertainty based on empirical variance among rewards. Here, we propose the inclusion of an explicit uncertainty inflation parameter that will set an asymptotic floor on the variance of a belief, and will account for “forgetting”

about options that have not been recently visited. We will assume that the uncertainty about belief increases by a fixed amount at each time step. We can accomplish this by simply adding a further Gaussian, uncertainty inflation distribution to the posterior at each time step. The sum of two, independent, normal distributions, $N(\mu_1, \sigma_1^2)$ and $N(\mu_2, \sigma_2^2)$, is another normal distribution with mean, μ_3 , and variance, σ_3^2 :

$$\mu_3 = \mu_1 + \mu_2, \quad \sigma_3^2 = \sigma_1^2 + \sigma_2^2. \quad (9)$$

The uncertainty inflation distribution will be taken to have a mean of 0, and a variance of u_i^2 : we model it as $N(0, u_i^2)$. Assuming that the posterior distribution, $P_t(\theta_i / x_{i,t})$, has a Gaussian distribution (as above), and following the notation of Eq 8, performing this summation will increase the variance of the posterior, but leave the mean unchanged. As such, we use it to model the process by which an individual slowly forgets what it has learned about a given parameter of interest (e.g. food quality at a particular flower).

This inflation factor will be added to the distribution describing the belief in each option, $P(\theta_{i,t})$, at each time step. In the case where a cue is not sampled for a given time step, the variance of the estimate of $P(\theta_i)$ will therefore linearly increase over time:

$$P(\theta_{i,t}) = N(\mu_{i,t} = \mu_{i,t-1}, \quad \sigma_{i,t}^2 = \sigma_{i,t-1}^2 + u_i^2). \quad (10)$$

In the alternate case, if an animal does sample an option, i , we will still include the variance inflation factor in updating the belief about $P(\theta_{i,t})$, to reflect the fact that animals do not reach perfect certainty in their beliefs (Dayan et al. 2000). We note that the choice to add the variance inflation term to beliefs about sampled and unsampled options does not qualitatively affect the overall performance of the probabilistic sampling model. The belief in the relative value of the two options depends on the variance of both (see below).

Under repeated sampling, the variance of $P(\theta_i)$ will reach a non-zero asymptote when $u_i^2 > 0$. This asymptote will depend on the variance of the likelihood distribution, v_i^2 , as well as the uncertainty parameter u_i^2 . The asymptotic variance of the posterior, v_i^{2*} , will be reached when $v_i^{2*} = (1/v_i^2 + 1/u_i^2)^{-1} + u_i^2$ (combining Eq. 8 and Eq. 9).

2.5 Decision-making

The process of learning and updating beliefs is distinct from making decisions based on those beliefs. An animal might simply choose the best of a set of alternatives. It might choose among these alternatives, but only when its belief crosses some threshold (Hutchinson and Gigerenzer 2005; Todd and Gigerenzer 2000; Gigerenzer and Gaissmaier 2011; Pelé and Sueur 2013). It might also choose probabilistically among a set of options.

We chose the example of the bee data of Cnaani et al. (2006), because it clearly demonstrates this latter scheme, which we will refer to as probabilistic foraging—and

clearly demonstrates the role of uncertainty in belief in decision-making. For the purposes of illustration, we will describe models that assume that animals are choosing among options directly according to their belief that a sample from option 1 will be greater than a sample from option 2.

We can model probabilistic decision-making in the RW model, if we scale the probability of choice by the difference in expected values of the options. In this case, we are modelling binomial choice, thus we will set between options 1 and 2, we use a simplified logistic regression framework, with a regression parameter β . This describes the effect that the difference in expected rewards $w_{1,t} - w_{2,t}$ at time t , has on choice. We evaluate the probability of choosing option $i = 1$ at time t , denoted by $C_{1,t}$ as:

$$P_t(C_{1,t}|w_{1,t}, w_{2,t}) = \frac{1}{1 + e^{-(\beta(w_{1,t} - w_{2,t}))}}. \quad (11)$$

The term β effectively imposes a scale for adding variance, or uncertainty, to our point-estimates of belief.

When beliefs are represented as probability distributions, as in the Bayesian framework, the decision-making process is more straightforward. In the simplest form, we can choose probabilistically among options 1 and 2 at time t , by evaluating the probability that the value of θ_1 is greater than θ_2 , given our beliefs after the most recent observation, $x_{i,t}$ (which may be either $x_{1,t}$ or $x_{2,t}$):

$$P(\theta_1 > \theta_2 | x_{i,t}). \quad (12)$$

We do so by constructing a distribution for $P(\theta_{1,t} | x_{1,t}) - P(\theta_{2,t} | x_{2,t})$. Remembering Eq 9, this is a normal distribution with mean $\mu_{3,t} = \mu_{2,t} - \mu_{1,t}$ and variance $\sigma_{3,t}^2 = \sigma_{2,t}^2 + \sigma_{1,t}^2$. Evaluating the cumulative density function of this Gaussian, for $x = 0$, gives us the probability that a draw from $P(\theta_{1,t} | x_{1,t})$ will be higher than a draw from $P(\theta_{2,t} | x_{2,t})$. If an animal is choosing strictly probabilistically based on its beliefs, this is also the proportion of times that it will choose option 1 rather than option 2. This framework generalizes naturally to cases in which there are more than two options to choose from.

3 Model evaluation

We wanted to evaluate whether we can obtain naturalistic behavior from our models. We are not interested in estimating precise parameters, but in testing whether combinations of parameters can reproduce the range of observed behaviour. Thus, we assessed qualitative fit—whether we could change the rate of learning, and the lower asymptote of the learning curve, using a variety of parameter values—through simulation. We examined behavior of both the RW model, and the Bayesian model. In the experiment of Cnaani et al. (2006), cues of flower quality included odor, nectar volume, and nectar concentration. For simplicity, we modeled only two kinds of flower in each simulation, distinguished by a single composite

cue, i —such that flower odor (geranium or rose) was perfectly correlated with the differences in either (high or low) volume or (high or low) sugar concentration. We varied either volume or concentration, but not both, within a given simulation (i.e. we did not model the reciprocal treatment, in which flowers differed in both nectar concentration and volume); thus values of α_i and v_i^2 did not vary among flower types within a simulation.

In Cnaani et al. (2006), the rewards were tightly controlled, so, again for simplicity, we did not model variation around the rewards the bees received, ($\theta_i, \bar{x}_i$ —Table 1). That is, empirical sampling variance was set at 0. As noted above, this does not entail that the likelihood distribution has a variance, v_i^2 , of 0. There may be multiple contributions to v_i^2 , in addition to sampling variance—including sensory and memory limitations, and evolved biases (see also Discussion). We have found that including empirical sampling variance in rewards, $x_{i,b}$ has little qualitative effect on model behavior, apart from some stochastic fluctuation in preference around the lower asymptote.

At each time step of a simulation, a bee chose probabilistically among two flower types, i , according to its belief about the relative value of the two cues. The bee updated its beliefs, based on the reward it received, as described above, and then it chose again.

For the RW model, at each time, t , bees chose among flowers according to Eq 11—defined by their current beliefs about each option, $w_{i,b}$ and the regression coefficient β (Table 1). The belief about the value of the chosen flower type was updated according to Eq 3, given the reward of the flower $x_{i,b}$ and the associability parameter, α_i . Among simulations, we varied α_i for values between 0.03 (rapid learning) and 0.003 (slow learning). We also varied the uncertainty parameter, β , between a high ($\beta=3$) and low ($\beta=1.3$) value.

For the Bayesian model, decisions were made according to Eq 12, defined by the posterior distributions at time t , for each option, $P_t(\theta_i/x)$. The posterior distribution, $P_t(\theta_i/x_{i,b})$, was updated according to Eq 5 for the chosen flower type. The likelihood—of observing the reward, $x_{i,b}$ for values of θ_i —was described by the likelihood distribution, Eq 7, which was a normal distribution with mean μ_i (the true reward value), and variance v_i^2 . The variance of the posterior distributions for both flower types was updated at each time step, according to Eq 10, with the uncertainty inflation factor, u_i^2 . Among simulations we varied the likelihood-distribution variance, v_i^2 , for values between 2.5 (highly believable cues) and 10 (less believable cues). We also varied the uncertainty inflation parameter, u_i^2 , between a low ($u_i^2=0.0008$) and high ($u_i^2=0.013$) value.

The simulation ran in three stages, to mimic the conditions experienced by the bees used in Cnaani et al. (2006)—a burn-in period (pre-training), a first equilibrium period, and a learning period. For the first 270 time steps (or flower visits) of the simulation we set the rewards equal to each other ($\theta_1=\theta_2=2.0$, $x_1=x_2=2.0$ for the Bayesian and RW simulations respectively); and let the bees' beliefs distribution for each reward stabilize. This was the burn-in. We ran the simulation for another 150 visits, recording the bees' choices, for the first equilibrium period. We numbered these time steps -150:0. At time 0, we changed the flower rewards. Flower 1, the low-reward flower, was given a value of $\theta_1=x_1=1.6$; and

flower 2, $\theta_2=x_2=2.2$. These differences mimicked the change in expected nectar quality (net energy gain) for one of the manipulations of (?). The bees were allowed to forage for another 480 visits, and all visits were recorded. This was the learning period. Each set of parameters was replicated 10 times (by 10 different bees) in a single experiment.

The change in foraging preference was modeled by examining the change in the probability of visiting the lowest reward flower across time, in 30-visit increments, to mimic the protocol of Cnaani et al. (2006). We modeled the change in foraging with 4 parameter sigmoid regression (nplr, in R), as they did. We were particularly interested in the inflection point, (X_0), which indicates speed of learning, and the final asymptote, (Y_0), which under our model reflects the bees' equilibrium certainty in their preferences.

4 Model Results

In simulations using the RW and Bayesian models, we were able to obtain patterns of observed choice that spanned the range of values reported in Cnaani et al. (2006), shown in Fig. 1. In their experiments, the final asymptote, Y_0 , ranged between 0.033 and 0.225. In our RW simulations (Fig. 2), we obtained final asymptotes ranging between 0.018 ($\alpha_j=0.03$, $\beta=3.0$) and 0.247 ($\alpha_j=0.003$, $\beta=1.3$). In the Bayesian simulations, our final asymptotes ranged between 0.018 ($v_i^2=2.5$, $u_i^2=0.0008$), and 0.227 ($v_i^2=10.0$, $u_i^2=0.013$). Similarly, we were able to obtain rates of learning (defined by the midpoint, X_0), that spanned the range of those observed by Cnaani et al. (2006)—between 74.70, and 153.02. In the RW simulations, we obtained values of X_0 between 36.0 ($\alpha_j=0.03$, $\beta=3.0$) and 205.92 ($\alpha_j=0.003$, $\beta=1.3$). In the Bayesian simulations we obtained values between 28.89 ($v_i^2=2.5$, $u_i^2=0.013$), and 168.79 ($v_i^2=10.0$, $u_i^2=0.0008$).

The RW results show that, as expected, learning rate is primarily affected by the associability parameter, α_j , whereas the final asymptote is most affected by the regression parameter, β (Fig. 2). The Bayes simulation results are similar. The signal reliability term, v_i^2 , has a strong relationship with the speed at which the bees acquire new information. Learning is almost immediate when v_i^2 is low, and more moderate at high values of v_i^2 . The uncertainty inflation parameter, u_i^2 , has a strong effect on the final asymptote.

It is striking that values of v_i^2 need to be so high, in order to produce realistic rates of bee learning. Note that in Cnaani et al. (2006), flower types were invariant in the rewards they provided. Bees were clearly updating their priors less strongly than would result from purely the empirical standard deviation in flower rewards. That is, the bees were acting as if any one visit to a flower of a given type is almost completely uninformative. This is surprising, since bumble bees are known for having excellent memories for patterns and places. However, they seem require hundreds of visits to flowers to converge on strong preferences for certain cues—and they never stop sampling alternatives. Conversely, the value of uncertainty inflation needed to produce realistic asymptotic rates of foraging was a small fraction of the actual difference in reward produced by the two flower types. This suggests even low rates of “forgetting” may be enough to ensure probabilistic sampling of alternatives.

5 Discussion

We argue that any minimal model of learning and choice needs to account not only for learning rates, but also patterns of variation in choice. We were able to model naturalistic behavior with a simple, recursive, model of Bayesian updating with probabilistic choice. The Bayesian model resembles models supported by several recent results in animal cognition, and describes learning in an efficient way. The likelihood-distribution representation of beliefs allowed us to model choice in a straightforward manner. Our Bayesian model has a direct analogue with a standard Rescorla-Wagner model, but has the advantage that it explicitly accounts for uncertainty at every stage of learning and choice, instead of adding it in after the fact.

5.1 Parsimony

Contra Bowers and Davis (2012a), who claim that Bayesian models are overparameterized and thus non-parsimonious, modeling learning and choice with our Bayesian formulation is very nearly as simple as the Rescorla-Wagner model. Both models account for asymptotic variation in flower choice, with analogous terms. In the RW model, this is the logistic regression, β , term. In the Bayesian model, it is the uncertainty inflation parameter, u_i^2 . While uncertainty inflation is not inherent in a purely-Bayesian approach, it is included because it reflects commonly observed patterns of animal cognition—namely limits to reinforcement, and forgetting (Dayan et al. 2000; Kraemer and Golding 1997; White 2001; Rescorla 2004), and is similar to the Temporal Weighting Rule (Devenport 1998; Devenport et al. 1997), which describes the observable decay of the effect of the most recent conditioning events on belief. The major difference between the two models of uncertainty inflation is that in the TWR model, beliefs converge on unbiased means over time, while in the Bayesian model, it is the variance of the belief distribution which changes. Both the Bayesian and RW model incorporated another term that affected learning rates. In the RW model learning required a cue-specific salience term, α ; the Bayesian model updates at a rate proportional to the believability, v_i^2 , of the cue.

In both the Bayesian and Rescorla Wagner models, an explicit prior belief was required. While Bayesian priors are often misunderstood, and criticized (Bowers and Davis 2012a), an accounting of prior beliefs seems unavoidable in any model of learning. Even the choice of an initial belief of 0, with a variance of 0, represents some choice of prior. The single degree of additional complexity in our Bayesian model, over the RW model, came from explicitly defining the variance of this prior (that is the uncertainty inherent in the naïve belief). The RW model has no analogue to the term $\sigma_{i,t}^2$ (Table 1). As a practical matter, this means that under a Bayesian model an animal can become increasingly confident in its beliefs through experience, even without changing its estimate of the value of the rewards associated with cues; or less confident in its estimates, through forgetting. The term $\sigma_{i,t}^2$ may also include explicit evaluations of cue reliability or reward variance, which we will not consider here (but see below). There is no equivalent process under the RW model, which assumes that the error around belief estimates will be constant. It is worth noting that the problem of changing error variance, resulting from experience, is often dealt with in animal experiments

by extensive training periods, or burn in times (e.g. Cnaani et al. 2006; Bröder 2000). This removes the need to model the additional parameter, but sacrifices the ability to model learning in a naïve individual.

These models are based on the simplest general models of learning and choice. It might be possible to posit simpler models for restricted scenarios, that require one or two fewer degrees of freedom than the models outlined above. But it is not clear that these “simpler” models will, in fact, be any more parsimonious. First, as we have already noted, some of the apparent complexity in Bayesian models is due to explicitly defining terms that are implicit in other models. A prior of 0 is still a prior. More problematically, model selection, among an arbitrary number of competing models (Kensinger and Luttbegg 2014), introduces a potentially large number of free parameters. Model-selection degrees of freedom are typically unaccounted for. In contrast, model-fitting, using a well-supported underlying model, is more transparent, and can be more parsimonious.

5.2 Interpretation of the likelihood distribution

In a Bayesian framework, the variance of the likelihood distribution, v_i^2 is typically assumed known, since the focus is usually on estimating the mean of θ . It is possible, however, to estimate v_i^2 from a set of observations, and update this estimate as new observations are collected. Animals are known to evaluate cue reliability in this manner, and to update their weighting of cues in multi-cue choice as a consequence (Vossel et al. 2014). This contrasts with the cue salience term, α_i of the RW model, which is usually taken as a constant, presumably reflecting innate biases. We have primarily interpreted v_i^2 from the Bayesian perspective, above; and described it in the context of innate sensory limits, which could contribute to perceived variance among sampling events. From this perspective, the higher v_i^2 associated with nectar volume, compared with nectar concentration, might indicate that bees are more accurate in their estimates of sweetness than volume.

In the framework of the current model, animals did not explicitly estimate the variance of the likelihood distribution because we assumed animals did not retain explicit memories of prior visits. This assumption is, in part, a simplification of convenience. Animals as complex as bees might retain memories of individual visits, and update their beliefs about v_i^2 accordingly, based on empirical reward variance. In practice, however, v_i^2 may not simply reflect empirical cue reliability; and may be constrained, as in our model. First, v_i^2 may include the effects of evolved preferences, or beliefs, about reward reliability. In this way, we might interpret v_i^2 in a manner consistent with the usual interpretation of α_i —that is, as an associability term. Additionally, estimates of signal reliability may be plastic in early development, but fixed in adults, which would also produce a constant v_i^2 , independent of the current empirical variance in rewards. And finally, animals with limited memories may be constrained to have a fixed estimate of v_i^2 .

The high estimates of v_i^2 associated with realistic patterns of foraging in bees were somewhat surprising, given that Cnaani et al. (2006) had manipulated the nectar cues such that they

were completely predictive of the rewards (as best as possible, the variance in rewards associated with cues was set to 0). Still, the bees required many visits to become confident of the relative values of the odor cues. This might imply that bees were unable to consistently assess the rewards, or to remember the value of rewards; in this view, the high value of v_i^2 reflects sensory or cognitive limits on the part of the bees, or possibly neural noise. Another contributor to this high v_i^2 , however, may be an evolved tendency to place little reliability in certain cues when estimating mean expectation of reward from a given flower type. In particular, nectar volume may be unreliable in nature, particularly when multiple nectar feeders are foraging in the same flowers; thus there may be some adaptive benefit to not quickly forming a belief in mean flower reward based on nectar volume cues.

The role of uncertainty in the exploration of novel and familiar cues can be seen in the way it can be employed in the description of several common heuristic models.

5.3 Comparison with “heuristics”

Heuristics are simple decision rules that can be surprisingly efficient in arriving at correct decisions. Heuristics are often offered as alternatives to Bayesian models (Bowers and Davis 2012a,b; Hutchinson and Gigerenzer 2005). But modeling the underlying belief state does not preclude heuristics. Heuristics might instead be described as the *behavioral algorithms* employed in the process of belief updating, and decision-making. Specific heuristics, then, should naturally emerge from the underlying cognitive model, and modeling belief explicitly will provide a better description of actual (as opposed to theoretical) patterns of choice. To illustrate, let us consider two heuristics—the “descending-threshold”, and “take-the-best” (Todd and Gigerenzer 2000; Hutchinson and Gigerenzer 2005). We will describe these heuristics in terms of what they imply about the underlying belief updating process. We will also show how our minimal model of recursive Bayesian updating and probabilistic choice can account for how these heuristics are implemented in practice by humans, and other animals.

The “descending-threshold” heuristic has been proposed as a simple rule of explaining extended, careful search among n options, before settling on the best. The typical scenario given is a female searching for a mate in a lek (Todd and Gigerenzer 2000; Real 1990). It posits an initially high threshold for accepting any option. As the individual (the female in the lek) evaluates a series of these unacceptably poor options, her acceptance threshold gradually gets lower, and she will mate with the first male whose quality exceeds this lowered threshold. It is expected that pursuing this strategy will give a female a very good chance of identifying, and mating with, the best quality male available.

We can represent this in our Bayesian model, as was hinted at in Real (1990). The high initial threshold described in the heuristic, corresponds to a high prior expectation of overall (lek-level) male value (Eq 4). Let i_t be an observed male at time t , and $x_{i,t}$ be his measured quality. For the sake of expediency, we will assume no observation error around this measurement, but the illustration can generalize to a more realistic scenario. The female will evaluate her belief about the general quality of males in the lek in the standard Bayesian manner (Eq 5), as a function of her prior belief in male quality, and the likelihood of her

observation. The likelihood distribution, in this case, will be described by $x_{i,b}$ and the support, v_i^2 , that this observation gives to the female's belief in overall male quality (Eq 7).

After each successive evaluation, we can model probabilistic female mate choice. In the spirit of the simplest form of the descending threshold heuristic, we will assume that a female does not remember males individually. She will choose to either mate with the male she has just evaluated, or will select another male from the lek to evaluate. Her choice rule, similar to the relation described in Eq 12, will simply be the probability that the male i_b , given his quality, $x_{i,b}$ is better than a randomly drawn male from the population, given her belief in the quality of males overall, $P_i(\theta/x_{i,b})$. Note that this relationship will be a point mass (again for simplicity), but we can easily evaluate the distribution, if we assume uncertainty around the estimate of male quality.

The descending-threshold heuristic will emerge naturally, given: a prior estimate of θ higher than actual male quality, updating, and probabilistic choice. Under these conditions, the perceived values of all males will be very low initially (females will be choosy). But as the female's belief about male value converges on the actual mean of the males in the lek, the probability that a female will choose a male will increase. This probability will be especially high for large values of $x_{i,b}$ —that is, females will be most likely to mate with the best males. This model also predicts some of the properties of real mate choice in leks. For instance, probabilistic choice, and a gradient of male quality, ensures a gradient of male mating success, as is typically observed (Alatalo et al. 1992; Mackenzie et al. 1995).

We are not claiming that this formulation is a completely accurate description of the cognition of females in leks. This would need further empirical work. Luttbeg (1996) has a different Bayesian solution, his Comparative Bayes tactic, for the same problem, which has garnered some strong empirical support (Luttbeg and Langen 2004; Luttbeg 2002). We want to emphasize, however, that modelling updating of uncertain beliefs simplifies the description of the heuristic, and better fits observed behaviour.

Finally, the “take-the-best” heuristic is at the other extreme of information gathering, and has been one of the heuristics most studied. It produces fast, and computationally frugal, choice. As it is usually described (Hutchinson and Gigerenzer 2005; Todd and Gigerenzer 2000; Newell et al. 2003), under take-the-best, individuals choose among a finite set of n objects, that can differ in observed value for a number of cues, $x_{i,n}$. The cues are ranked in order of their reliability; and are assessed in order, from most to least reliable. Imagine a squirrel examining two figs. The squirrel can only take one, and must do it quickly. It will evaluate the figs in turn by their size (the best cue), then smell, and finally color. As soon as the squirrel recognizes that one of the figs is better, it will choose. It will first attempt to take the largest fig. But if the figs are too close in size to differentiate, it will smell them, and take the most sweet smelling.

Under the recursive Bayesian model, this is equivalent to saying that the squirrel has different beliefs regarding the reliability of the available cues, v_i^2 . There is an increasing body of evidence from studies of selective attention, that animals often evaluate the world in precisely this way. That is, in many situations they estimate the reliability of cues, and attend

to the most reliable cues in ways that are most consistent with Bayesian inference (Vossel et al. 2014; Yu and Dayan 2005)—although this does not preclude point-estimate models of selective attention in other cases (Le Pelley et al. 2016).

Rigorously developing a full model reconciling the heuristic and the Bayesian approach is beyond the scope of this discussion. However, the outlines of the probabilistic analogue of take-the-best are simple. If the squirrel samples the cues, one-at-a-time, and conducts Bayesian updating of its belief regarding the quality of each option after sampling each cue, then a natural analogue of “take the best” arises. If the squirrel decides to stop and choose one of the two options with probability proportional to that given in Eq. 12, we will see that the squirrel sometimes “guesses” (before it evaluates enough cues, according to the take-the-best heuristic). Sometimes it will evaluate “too many cues” (after a difference has been noticed—because the quality of the two cues, while different, is not different enough to support one option over the other, with high probability). And sometimes it will make a “mistake” (that is, it will take an option not predicted to be best, by the most reliable cues).

This is precisely what the empirical evaluation of take-the-best has shown. There have been extensive attempts to experimentally verify that individuals use take-the-best, in real decision-making contexts (e.g. Newell et al. 2003; Bröder 2000). These experiments (in humans) have been carefully designed to maximize the chances of observing take-the-best behavior. However, these experiments demonstrate that there is a great deal of inter- and intra-individual variation in the use of the heuristic. There was a great deal of guessing (insufficient evaluation of cues), over-evaluation, and “wrong” choices (Newell et al. 2003).

Consistent with the idea that the brain is a “Bayesian inference machine” (Vossel et al. 2014), these results suggest that humans, and animals, often try to *maximise their confidence* in their choices, in the process of learning while decision-making (Hausmann and Läge 2008). Rather than trying to optimize decision-making based on simple point-estimate comparisons, humans and non-human animals instead seem to be attempting to discriminate among the attractiveness of the available options (Harvey and Bolger 2010; Newell and Shanks 2003). Thus, modeling their belief states, including uncertainty, is critical to understanding their behavior and choice. It is here that we see the most important practical distinction between the RW model and the Bayesian model.

5.4 Summary

Recursive algorithms, such as those based on the Rescorla-Wagner model of reinforcement learning, and Bayesian updating, are computationally tractable. They are based on decades of extensive empirical results in neurobiology and animal behavior. In many ways, the RW model and the Bayesian model are extremely similar—and for some analytical purposes may be interchangeable. The Bayesian model, however, represents certainty as a probability distribution, rather than adding probability post hoc.

A number of experimental results in studies of cognition suggest that the Bayesian model is a more accurate description of how animals learn, and make decisions. Accordingly, our model makes several testable predictions about the way in which animals may learn, given novel cues; or forget, when familiar cues are withheld. For instance, assessment of novel

cues (maximum uncertainty) should be less sensitive to known high-value cues in a Bayesian model, than predicted by a point-estimate model. That is, the presence of known high-value cues should not suppress exploration. Forgetting should also take a different course. In point estimate models, the belief in the value associated with withheld cues typically converges to a common, presumably low, “neutral” value, thus any attractive cue will become less attractive when withheld. In the case of the Bayesian model, we would expect very attractive cues to become less attractive when they are withheld, but less attractive cues to become more attractive, since the mean of the belief does not change, but the uncertainty around it increases. And finally, we would expect to observe a very different pattern of choice under the Bayesian model and point estimate models, when there are multiple (>2) options. Under most point estimate models, relative preferences are proportional, and transitive. In a Bayesian framework, if there is high uncertainty about a cue, an animal will “oversample” it, relative to a well-known high-mean cue, and “undersample” it relative to a well-known low-mean cue; thus relative preferences among cues will not fall along a simple proportional scale.

An explicit accounting of uncertainty may also inform our understanding of heuristics. Far from being a competing description of animal behavior, heuristics may result directly from the cognitive processes described by the recursive Bayesian algorithm. This level of description may unify diverse heuristics within a single framework, and explain otherwise puzzling “mistakes”. This common cognitive framework may be important for studying behavioral evolution, and genetic variation in decision-making. While recursive Bayesian models have been known for some time, they have not yet been applied to animal behavior. Their use may enable us to explain that ubiquitous property of animal behavior: why animals are probabilistic, not deterministic.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

National Institute of Health R01MH100879, National Science Foundation DMS 1101060

We would like to thank Stephen Hamblin for the germ of the idea that became this paper, and for extensive discussion on the heuristics and foraging literature.

Funding: This study was funded by the National Institute of Health (grant number R01MH100879) and the National Science Foundation (DMS 1101060)

References

- Abrahams MV. Foraging guppies and the ideal free distribution: the influence of information on patch choice. *Ethology*. 1989; 82(2):116126.
- Alatalo RV, Hgund J, Lundberg A, Sutherland WJ. Evolution of black grouse leks: female preferences benefit males in larger leks. *Behav Ecol*. 1992; 3(1):5359.
- Behrens TE, Woolrich MW, Walton ME, Rushworth MF. Learning the value of information in an uncertain world. *Nat Neurosci*. 2007; 10(9):121421.
- Biernaskie JM, Walker SC, Gegear RJ. Bumblebees learn to forage like Bayesians. *Am Nat*. 2009; 174(3):41323.

- Bouton ME. Context, ambiguity, and unlearning: sources of relapse after behavioral extinction. *Biol Psychiat*. 2002; 52(10):976986.
- Bowers JS, Davis CJ. Bayesian just-so stories in psychology and neuroscience. *Psychol Bull*. 2012a; 138(3):389414.
- Bowers JS, Davis CJ. Is that what Bayesians believe? reply to Griffiths, Chater, Norris, and Pouget (2012). *Psychol Bull*. 2012b; 138(3):4236.
- Bröder A. Assessing the empirical validity of the take-the-best heuristic as a model of human probabilistic inference. *J Exp Psychol Learn*. 2000; 26(1):13321346.
- Cartoni E, Puglisi-Allegra S, Baldassarre G. The three principles of action: a Pavlovian-instrumental transfer hypothesis. *Front Behav Neurosci*. 2013; 7:153. [PubMed: 24312025]
- Chater N, Oaksford M, Hahn U, Heit E. Bayesian models of cognition. *Wires Cogn Sci*. 2010; 1(6): 81123.
- Cnaani J, Thomson JD, Papaj DR. Flower choice and learning in foraging bumblebees: effects of variation in nectar volume and concentration. *Ethology*. 2006; 112(3):278285.
- Courville AC, Daw ND, Touretzky DS. Bayesian theories of conditioning in a changing world. *Trends Cogn Sci*. 2006; 10(7):294300.
- Dayan P, Kakade S, Montague PR. Learning and selective attention. *Nat Neurosci*. 2000; 3 Suppl(11): 121823.
- Devenport L, Hill T, Wilson M, Ogden E. Tracking and averaging in variable environments: A transition rule. *J Exp Psychol: Anim B*. 1997; 23(4):450460.
- Devenport LD. Spontaneous recovery without interference: Why remembering is adaptive. *Anim Learn Behav*. 1998; 26(2):172181.
- Dunsmoor J, Niv Y, Daw N, Phelps E. Rethinking extinction. *Neuron*. 2015; 88(1):4763.
- Fischer BJ, Pea JL. Owls behavior and neural representation predicted by Bayesian inference. *Nat Neurosci*. 2011; 14(8):10616.
- Foley BR, Saltz JB, Nuzhdin SV, Marjoram P. A Bayesian approach to social structure uncovers cryptic regulation of group dynamics in *Drosophila melanogaster*. *Am Nat*. 2015; 185(6):797808.
- Fusi S, Asaad WF, Miller EK, Wang XJ. A neural circuit model of flexible sensorimotor mapping: learning and forgetting on multiple timescales. *Neuron*. 2007; 54(2):31933.
- Gelman, A., Carlin, JB., Stern, HS., Dunson, DB., Vehtari, A., Rubin, DB. Bayesian data analysis. 3. Chapman and Hall; 2013.
- Gershman SJ. A unifying probabilistic view of associative learning. *PLoS Comput Biol*. 2015; 11(11):e1004, 567.
- Gershman SJ, Blei DM, Niv Y. Context, learning, and extinction. *Psychol Rev*. 2010; 117(1):197209.
- Gibbon J, Church RM, Meck WH. Scalar timing in memory. *Ann NY Acad Sci*. 1984; 423:5277.
- Gigerenzer G, Gaissmaier W. Heuristic decision making. *Annu Rev Psychol*. 2011; 62:45182.
- Green RF. A simpler, more general method of finding the optimal foraging strategy for Bayesian birds. *Oikos*. 2006; 112(2):274284.
- Greggers U, Menzel R. Memory dynamics and foraging strategies of honeybees. *Behav Ecol and Sociobiol*. 1993; 32(1):1729.
- Griffiths TL, Chater N, Norris D, Pouget A. How the Bayesians got their beliefs (and what those beliefs actually are): comment on Bowers and Davis (2012). *Psychol Bull*. 2012; 138(3):41522.
- Harvey N, Bolger F. Collecting information: optimizing outcomes, screening options, or facilitating discrimination? *Q J Exp Psychol-A*. 2010
- Hausmann D, Läge D. Sequential evidence accumulation in decision making: The individual desired level of confidence can explain the extent of information acquisition. *Judgment and Decision Making*. 2008; 3(3):229243.
- Hutchinson JMC, Gigerenzer G. Simple heuristics and rules of thumb: where psychologists and behavioural biologists might meet. *Behav Process*. 2005; 69(2):97124.
- Iwasa Y, Higashi M, Yamamura N. Prey distribution as a factor determining the choice of optimal foraging strategy. *Am Nat*. 1981; 117(5):710723.

- Janmaat KR, Ban SD, Boesch C. Ta chimpanzees use botanical skills to discover fruit: what we can learn from their mistakes. *Anim Cogn.* 2013; 16(6):85160.
- Jones M, Love BC. Bayesian fundamentalism or enlightenment? On the explanatory status and theoretical contributions of Bayesian models of cognition. *Behav Brain Sci.* 2011; 34(4):16988. discussion 188231.
- Jovani R, Mavor R. Group size versus individual group size frequency distributions: a nontrivial distinction. *Anim Behav.* 2011; 82:10271036.
- Kaelbling L, Littman M, Moore A. Reinforcement learning: a survey. *J Artif Intell Res.* 1996; 4:237285.
- Kakade S, Dayan P. Acquisition and extinction in autoshaping. *Psychol Rev.* 2002; 109(3):533544.
- Kauffman A, Parsons L, Stein G, Wills A, Kaletsky R, Murphy C. *C. elegans* positive butanone learning, short-term, and long-term associative memory assays. *JoVE.* 2011; (49)
- Kensinger BJ, Luttbegg B. The limitations of inferring decision rule use from individuals sampling behaviour: a computational test of old and new algorithms. *Evol Ecol Res.* 2014; 16(2):179194.
- Knill DC, Pouget A. The Bayesian brain: the role of uncertainty in neural coding and computation. *Trends Neurosci.* 2004; 27(12):712719.
- Körding K. Decision theory: what should the nervous system do? *Science.* 2007; 318(5850):60610.
- Kraemer PJ, Golding JM. Adaptive forgetting in animals. *Psychon B Rev.* 1997; 4(4):480491.
- Krakauer DC, Rodríguez-Girones MA. Searching and learning in a random environment. *J of Theor Biol.* 1995; 177(4):417429.
- Kruschke J. Bayesian approaches to associative learning: From passive to active learning. *Learn Behav.* 2008; 36(3):210226.
- Le Pelley ME, Mitchell CJ, Beesley T, George DN, Wills AJ. Attention and associative learning in humans: an integrative review. *Psychol Bull.* 2016; 142(10):131.
- Liu CC, Watanabe T. Accounting for speed-accuracy tradeoff in perceptual learning. *Vision Res.* 2012; 61:10714.
- Luttbegg B. A Comparative Bayes tactic for mate assessment and choice. *Behav Ecol.* 1996; 7(4): 451460.
- Luttbegg B. Assessing the robustness and optimality of alternative decision rules with varying assumptions. *Anim Behav.* 2002; 63(4):805814.
- Luttbegg B, Langen TA. Comparing alternative models to empirical data: Cognitive models of western scrub-jay foraging behavior. *Am Nat.* 2004; 163(2):263276.
- Luttbegg B, Warner RR. Reproductive decision-making by female peacock wrasses: flexible versus fixed behavioral rules in variable environments. *Behav Ecol.* 1999; 10(6):666674.
- Ma WJ, Jazayeri M. Neural coding of uncertainty and probability. *Annu Rev Neurosci.* 2014; 37(1): 205220.
- Mackenzie A, Reynolds J, Brown V, Sutherland W. Variation in male mating success on leks. *Am Nat.* 1995; 145(4):633652.
- Maia TV. Reinforcement learning, conditioning, and the brain: successes and challenges. *Cogn Affect Behav Neurosci.* 2009; 9(4):34364.
- Matzel LD, Schachtman TR, Miller RR. Recovery of an overshadowed association achieved by extinction of the overshadowing stimulus. *Learn and Motiv.* 1985; 16(4):398412.
- McNamara JM, Houston AI. Memory and the efficient use of information. *J Theor Biol.* 1987; 125(4): 385395.
- McNamara JM, Green RF, Olsson O. Bayes theorem and its applications in animal behaviour. *Oikos.* 2006; 112(2):243251.
- Menzel R. Associative learning in honey bees. *Apidologie.* 1993; 24(3):157168.
- Michelena P, Sibbald AM, Erhard HW, McLeod JE. Effects of group size and personality on social foraging: the distribution of sheep across patches. *Behav Ecol.* 2009; 20(1):145152.
- Newell BR, Shanks DR. Take the best or look at the rest? Factors influencing one-reason decision making. *J Exp Psychol Learn.* 2003; 29(1):5365.

- Newell BR, Weston NJ, Shanks DR. Empirical tests of a fast-and-frugal heuristic: not everyone takes-the-best. *Organ Behav Hum Decision*. 2003; 91(1):8296.
- Olsson O, Brown JS. The foraging benefits of information and the penalty of ignorance. *Oikos*. 2006; 112(2):260273.
- Olsson O, Holmgren NMA. The survival-rate-maximizing policy for Bayesian foragers: wait for good news. *Behav Ecol*. 1998; 9(4):345353.
- Oster G, Heinrich B. Why do bumblebees major? A mathematical model. *Ecol Monogr*. 1976; 46(2): 129.
- Payzan-LeNestour E, Bossaerts P. Risk, unexpected uncertainty, and estimation uncertainty: Bayesian learning in unstable settings. *PLoS Comput Biol*. 2011; 7(1):e1001048.
- Pelé M, Sueur C. Decision-making theories: linking the disparate research areas of individual and collective cognition. *Anim Cogn*. 2013; 16(4):54356.
- Pouget A, Beck JM, Ma WJ, Latham PE. Probabilistic brains: knowns and unknowns. *Nat Neurosci*. 2013; 16(9):11701178.
- Real L. Search theory and mate choice. *Am Nat*. 1990; 136(3):376404.
- Real L, Ellner S, Harder LD. Short-term energy maximization and risk-aversion in bumble bees: a reply to Possingham et al. *Ecology*. 1990; 71(4):16251628.
- Real LA. Uncertainty and pollinator-plant interactions: the foraging behavior of bees and wasps on artificial flowers. *Ecology*. 1981; 62(1):2026.
- Reger ML, Poulos AM, Buen F, Giza CC, Hovda DA, Fanselow MS. Concussive brain injury enhances fear learning and excitatory processes in the amygdala. *Biol Psychiat*. 2012; 71(4):33543.
- Rescorla RA. Spontaneous recovery. *Learn Memory*. 2004; 11(5):5019.
- Roche J, Stubbs D, Glanz W. Assessment and choice: an operant simulation of foraging in patches. *J Exp Anal Behav*. 1996; 66(3):327347.
- Roesch MR, Esber GR, Li J, Daw ND, Schoenbaum G. Surprise! Neural correlates of Pearce-Hall and Rescorla-Wagner coexist within the brain. *Eur J Neurosci*. 2012; 35(7):1190200.
- Rushworth M, Behrens T. Choice, uncertainty and value in prefrontal and cingulate cortex. *Nat Neurosci*. 2008; 11(4):38997.
- Saltz J. Natural genetic variation in social environment choice: context-dependent gene-environment correlation in *Drosophila melanogaster*. *Evolution*. 2011; 65:2325–2334. [PubMed: 21790578]
- Schultz W, Dickinson A. Neuronal coding of prediction errors. *Annu Rev Neurosci*. 2000; 23:473500.
- Selonen V, Hanski IK. Decision making in dispersing Siberian flying squirrels. *Behav Ecol*. 2010; 21(2):219225.
- Stephens D. How important are partial preferences? *Anim Behav*. 1985; 33(2):667669.
- Todd PM, Gigerenzer G. Précis of Simple heuristics that make us smart. *Behav Brain Sci*. 2000; 23(5): 727741.
- Valone T. Are animals capable of Bayesian updating? An empirical review. *Oikos*. 2006; 112(2): 252259.
- Vossel S, Mathys C, Daunizeau J, Bauer M, Driver J, Friston KJ, Stephan KE. Spatial attention, precision, and Bayesian inference: a study of saccadic response speed. *Cereb Cortex*. 2014; 24(6): 143650.
- White KG. Forgetting functions. *Anim Learn Behav*. 2001; 29(3):193207.
- Wilson RC, Nassar MR, Gold JJ. Bayesian online learning of the hazard rate in change-point problems. *Neural Comput*. 2010; 22(9):245276.
- Wystrach A, Mangan M, Webb B. Optimal cue integration in ants. *Proc R Soc B*. 2015; 282(1816): 20151484.
- Yang T, Shadlen MN. Probabilistic reasoning by neurons. *Nature*. 2007; 447(7148):10751080.
- Yi L. Applications of timing theories to a peak procedure. *Behav Process*. 2007; 75(2):188198.
- Yu AJ, Dayan P. Uncertainty, neuromodulation, and attention. *Neuron*. 2005; 46(4):68192.

Fig. 1.

The results of Chanaani et al. 2006, used with permission. The proportion of visits bees pay to the less-rewarding flower type is shown. Before visit 0, rewards are equal. After visit 0, rewards are changed, either via changes in nectar volume (filled circles), or concentration (filled triangles). A third, reciprocal, manipulation varied both cues (open diamonds). Symbols represent mean values for 30 successive events, and vertical lines indicate standard error. Curves are fitted using 4-parameter sigmoid regression. Inflection points are shown with the horizontal bar.

Fig. 2.

Simulations of bee behavior, using Rescorla Wagner (RW), and Bayes-based updating. A range of values were tested. In the RW, the cue associability, α , and logistic regression coefficient, β , parameters were varied. For the recursive Bayes simulations, the cue believability, v_i^2 , and uncertainty inflation, u_i^2 terms were varied. Shown are the mean proportion of flower visits (symbols), for windows of 30 successive visits, and $n=10$ replicates. Vertical bars represent standard error. Before time=0, rewards among flower types are identical (2.0 units); after the bar, the high reward is set to 2.2, and the low reward is set to 0.61.

Table 1

Free parameters in the Bayes and RW models.

Bayes	Description	RW	Description
μ_{i0}	the mean of the prior associated with cue i	w_{i1}	the initial weight on cue i
$\sigma_{i,0}^2$	the variance of the prior associated with cue i	—	—
θ_i	the “true” value of the distribution of rewards, x_p , associated with cue i	$\bar{x}_i$	The “true” value of the reward
v_i^2	the variance of the likelihood for cue i (the “believability”)	a_i	the associability of cue i and a stimulus
u_i^2	the uncertainty inflation parameter	β	logistic regression coefficient