

RESEARCH ARTICLE

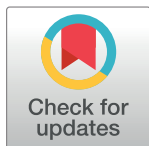
# Human online adaptation to changes in prior probability

Elyse H. Norton<sup>1</sup>, Luigi Acerbi<sup>1,2</sup>, Wei Ji Ma<sup>1,2</sup>, Michael S. Landy<sup>1,2\*</sup>

**1** Psychology Department, New York University, New York, New York, United States of America, **2** Center for Neural Science, New York University, New York, New York, United States of America

✉ Current Address: Département des neurosciences fondamentales, Université de Genève, CMU, Genève, Switzerland

\* [landy@nyu.edu](mailto:landy@nyu.edu)



**OPEN ACCESS**

**Citation:** Norton EH, Acerbi L, Ma WJ, Landy MS (2019) Human online adaptation to changes in prior probability. *PLoS Comput Biol* 15(7): e1006681. <https://doi.org/10.1371/journal.pcbi.1006681>

**Editor:** Ulrik R. Beierholm, Durham University, UNITED KINGDOM

**Received:** November 28, 2018

**Accepted:** June 20, 2019

**Published:** July 8, 2019

**Copyright:** © 2019 Norton et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** Code to perform the analyses described in this paper and all subjects' data are available at the following link: <https://github.com/lacerbi/ChangeProb>.

**Funding:** This work was supported by NIH grants EY08266 (to MSL) and EY020958 and EY026927 (to WJM). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

## Abstract

Optimal sensory decision-making requires the combination of uncertain sensory signals with prior expectations. The effect of prior probability is often described as a shift in the decision criterion. Can observers track sudden changes in probability? To answer this question, we used a change-point detection paradigm that is frequently used to examine behavior in changing environments. In a pair of orientation-categorization tasks, we investigated the effects of changing probabilities on decision-making. In both tasks, category probability was updated using a sample-and-hold procedure: probability was held constant for a period of time before jumping to another probability state that was randomly selected from a predetermined set of probability states. We developed an ideal Bayesian change-point detection model in which the observer marginalizes over both the current run length (i.e., time since last change) and the current category probability. We compared this model to various alternative models that correspond to different strategies—from approximately Bayesian to simple heuristics—that the observers may have adopted to update their beliefs about probabilities. While a number of models provided decent fits to the data, model comparison favored a model in which probability is estimated following an exponential averaging model with a bias towards equal priors, consistent with a conservative bias, and a flexible variant of the Bayesian change-point detection model with incorrect beliefs. We interpret the former as a simpler, more biologically plausible explanation suggesting that the mechanism underlying change of decision criterion is a combination of on-line estimation of prior probability and a stable, long-term equal-probability prior, thus operating at two very different timescales.

## Author summary

We demonstrate how people learn and adapt to changes to the probability of occurrence of one of two categories on decision-making under uncertainty. The study combined psychophysical behavioral tasks with computational modeling. We used two behavioral tasks: a typical forced-choice categorization task as well as one in which the observer specified the decision criterion to use on each trial before the stimulus was displayed. We formulated an ideal Bayesian change-point detection model and compared it to several

alternative models. We found that the data are explained best by a model that estimates category probability based on recently observed exemplars with a bias towards equal probability. Our results suggest that the brain takes multiple relevant time scales into account when setting category expectations.

## Introduction

Sensory decision-making involves making decisions under uncertainty. Furthermore, optimal sensory decision-making requires the combination of uncertain sensory signals with prior expectations. Perceptual models of decision-making often incorporate prior expectations to describe human behavior. In Bayesian models, priors are combined with likelihoods to compute a posterior [1]. In signal detection theory, the effect of unequal probabilities (signal present vs. absent) is a shift of the decision criterion [2].

The effects of prior probability on the decision criterion have been observed in detection [2–4], line tilt [5], numerosity estimation [6, 7], recognition memory [8], and perceptual categorization [9] tasks, among others. These studies generally use explicit priors, assume a fixed effect, and treat learning as additional noise. However, there are many everyday tasks in which the probability of a set of alternatives needs to be assessed based on one's past experience with the outcomes of the task. The importance of experience has been demonstrated in studies examining differences between experience-based and description-based decisions [10, 11] and in perceptual-categorization tasks with unequal probability, in which response feedback leads to performance that is closer to optimal than observational feedback [12, 13]. While these studies demonstrate the importance of experience on decision-making behavior, they do not describe how experience influences expectation formation. The influence of experience on the formation of expectations has been demonstrated for learning stimulus mean [14–17], variance [14, 18], higher-order statistics [19], likelihood distributions [20], the rate of change of the environment [15–17, 21–23], and prior probability [24, 25]. Here, we add to previous work by investigating how one's previous experience influences probability learning in a changing environment.

In the previous work on probability learning by Behrens et al. [24], participants tracked the probability of a reward to learn the value of two alternatives. This is a classic decision-making task that involves combining prior probability and reward. In contrast, we are interested in perceptual decision-making under uncertainty, in which prior probability is combined with uncertain sensory signals. We might expect differences in strategy between cognitive and perceptual tasks, as cognitive tasks can be susceptible to additional biases. For example, participants often exhibit base-rate neglect (i.e., they ignore prior probability when evaluating evidence) in cognitive tasks [26] but not in perceptual tasks [2]. On the other hand, Behrens et al. [24] found that participants' behavior was well described by an optimal Bayesian model, in that observed learning rates quantitatively matched those of a Bayesian decision-maker carrying out the same task. A more recent study by Zylberberg et al. [25] examined probability learning under uncertainty in a motion-discrimination task. In this study, probability was estimated from a confidence signal rather than explicit feedback. Additionally, probability was changed across blocks, allowing participants to infer a change had occurred. Here, we examine probability learning when feedback is explicit and changes are sudden.

To investigate probability learning in uncertain conditions, we designed a perceptual categorization task in which observers need to learn category probability through experience. Since our goal was to examine low-level perceptual and decision-making processes, we used a

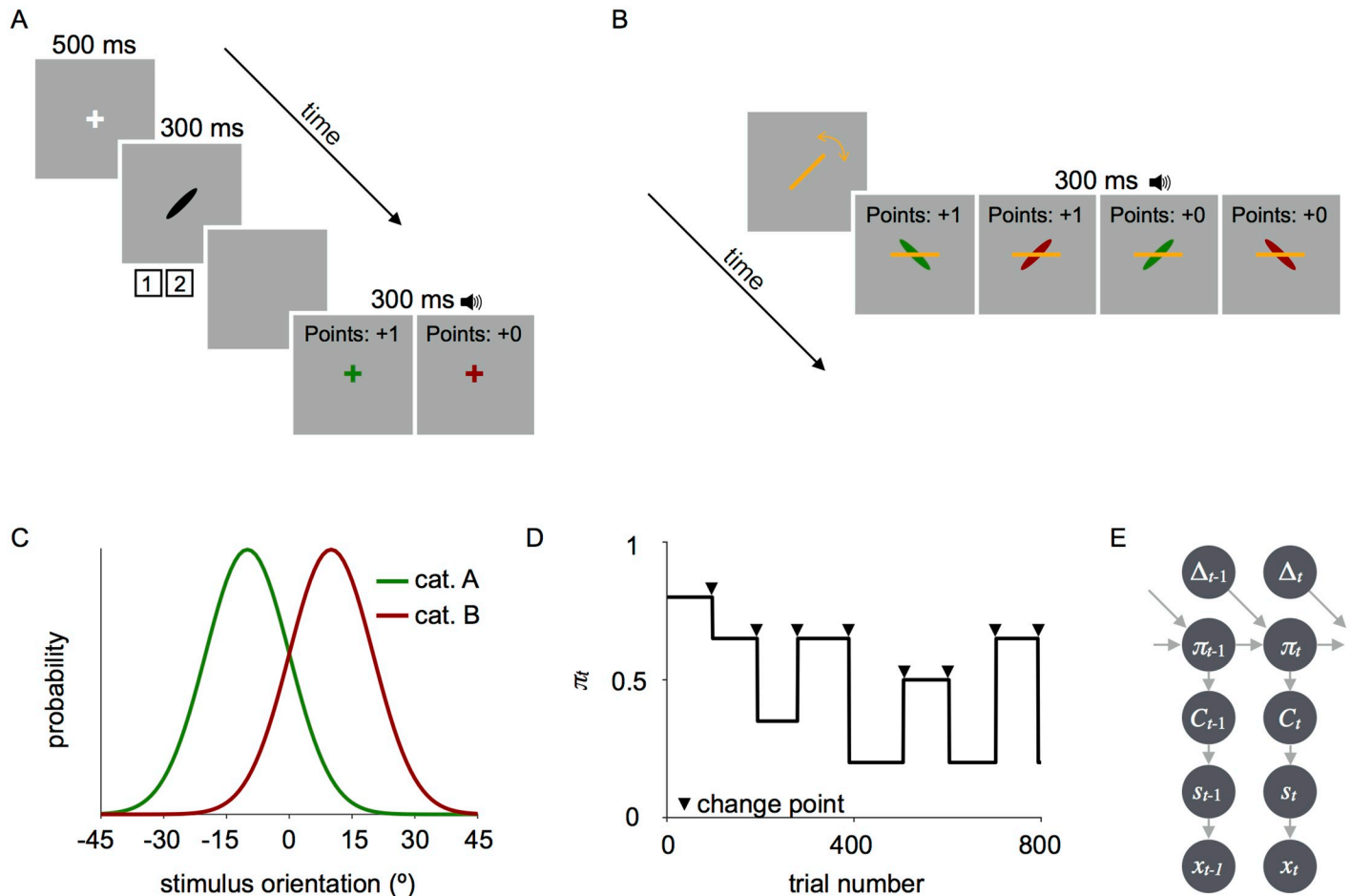
highly controlled experimental environment. To prevent the use of external cues (e.g., the start of a new block indicating a change in probability) probabilities were changed using a sample-and-hold procedure. This approach has been used in decision-making [21, 22, 24] and motor domains [18] to examine behavior in dynamic contexts. Observers completed both a covert- and overt-criterion task, in which the decision criterion was implicit or explicit, respectively. The overt task, previously developed by Norton et al. [16], provided a richer dataset upon which to compare models of human behavior. We determined how observers tracked category probability in a changing environment by comparing human behavior to both Bayesian and alternative models. We find that a number of models qualitatively describe human behavior, but that, quantitatively, model comparison favored an exponential averaging model with a bias towards equal priors and a flexible variant of the Bayesian change-point detection model with incorrect beliefs about the generative model. Although model comparison did not distinguish between these models, we interpret the exponential model with a conservative bias as a simpler, more biologically plausible explanation. Our results suggest that changes in the decision criterion are the result of both probability estimates computed on-line and a more stable, long-term prior.

## Results

### Experiment

During each session, observers completed one of two orientation-categorization tasks. On each trial in the *covert-criterion* task, observers categorized an ellipse as belonging to category A or B by key press (Fig 1A). On each trial in the *overt-criterion* task, observers used the mouse to rotate a line to indicate their criterion prior to the presentation of an ellipse (Fig 1B). The observer was correct if the ellipse belonged to category A and was clockwise of the criterion line or if the ellipse belonged to category B and was counter-clockwise of the criterion line. The overt-criterion task is an explicit version of the covert-criterion task developed by Norton et al. [16]. The overt-criterion task provides a richer dataset than the covert-criterion task in that it affords a continuous measure and allows us to see trial by trial changes in the reported decision criterion, at the expense of being a more cognitive task. In both tasks, the categories were defined by normal distributions on orientation with different means ( $\mu_B = \mu_A + \Delta\theta$ ) and equal standard deviation ( $\sigma_C$ ); the mean of category A was set clockwise of the the mean of category B and a neutral criterion would be located halfway between the category means (Fig 1C). The difficulty of the task was equated across participants by setting  $\Delta\theta$  to a value that predicted a  $d'$  value of 1.5 based on the data from the initial measurement session (see Methods). All data are reported relative to the neutral criterion,  $z_{\text{neutral}} = (\mu_A + \mu_B)/2$ .

Prior to testing, observers were trained on the categories. Only the covert-criterion task was used for training (see Section 6 in S1 Appendix). During training, category probability was equal ( $\pi = 0.5$ ) and observers received feedback on every trial that indicated both the correctness of the response (tone) and the generating category (visual). As a measure of category learning, we compute the probability of being correct in the training block and averaged across sessions. All observers learned the categories to the expected level of accuracy ( $p(\text{correct}) = 0.74 \pm 0.01$ ; mean  $\pm$  SEM across observers), given that the expected fraction of correct responses for an ideal observer with  $d' = 1.5$  and equal priors over categories is 0.77. As an additional test of category learning, immediately following training observers estimated the mean orientation of each category by rotating an ellipse. Each category mean was estimated once and no feedback was provided. There was a significant correlation between the true and estimated means for each category (category A:  $r = 0.82$ ,  $p < 0.0001$ ; category B:  $r = 0.97$ ,  $p < 0.0001$ ), suggesting that categories were learned. However, on average mean estimates



**Fig 1. Experimental design.** (A) Trial sequence in the covert-criterion task. After stimulus offset, observers reported the category by key press and received auditory feedback indicating the correctness of their response. In addition, the fixation cross was displayed in the color of the generating category. (B) Trial sequence in the overt-criterion task. Prior to stimulus onset, observers rotated a line to indicate their criterion by sliding the mouse side-to-side. When the observer clicked the mouse, a stimulus was displayed under the criterion line and auditory feedback was provided. (C) Example stimulus distributions for category A (green) and category B (red). Note that numbers on the  $x$ -axis are relative to the neutral criterion (i.e., 0 is the neutral criterion). (D) Example of random stepwise changes in probability across a block of trials. Change points occurred every 80-120 trials and are depicted above by the black arrows. Category A probabilities  $\pi_t$  were sampled from a discrete set,  $S_\pi = \{0.2, 0.35, 0.5, 0.65, 0.8\}$ . (E) Generative model for the task in which category probability  $\pi$  is updated following a sample-and-hold procedure, a category  $C$  is selected based on the category probability, a stimulus  $s$  is drawn from the category distribution and is corrupted by visual noise resulting in the noisy measurement  $x$ . Note that this diagram omits the dependency that leads to change points every 80-120 trials.

<https://doi.org/10.1371/journal.pcbi.1006681.g001>

were repelled from the category boundary (average category A error of  $11.3^\circ \pm 6.3^\circ$  and average category B error of  $-8.0^\circ \pm 2.6^\circ$ ; mean  $\pm$  SEM across observers) suggesting a systematic repulsive bias.

To determine how category probability affects decision-making, during testing category A probability  $\pi_t$  was determined using a sample-and-hold procedure (Fig 1D; category B probability was  $1 - \pi_t$ ). For  $t = 1$ , category A probability was randomly chosen from a set of five probabilities  $S_\pi = \{0.2, 0.35, 0.5, 0.65, 0.8\}$ . On most trials, no change occurred ( $\Delta_t = 0$ ), so that  $\pi_{t+1} = \pi_t$ . Every 80-120 trials there was a change point ( $\Delta_t = 1$ ), with change point sampled uniformly. At each change point, category probability was randomly selected from the  $S_\pi$  excluding the current probability. On each trial  $t$ , a category  $C_t$  was randomly selected (with  $P(\text{category A}) = \pi_t$ ) and a stimulus  $s_t$  was drawn from the stimulus distribution corresponding to the selected category. We assume that the observer's internal representation of the stimulus

is a noisy measurement  $x_t$  drawn from a Gaussian distribution with mean  $s_t$  and standard deviation  $\sigma_v$ , which represents visual noise ( $v$ ). The generative model of the task is summarized in Fig 1E.

### Bayesian models

To understand how decision-making behavior is affected by changes in category probability, we compared observer performance to several Bayesian models. To compute the behavior of a Bayesian observer, we developed a Bayesian change-point detection algorithm, based on Adams and MacKay [27], but which also accounts for Markov dependencies in the transition distribution after a change. Specifically, the Bayesian observer estimates the posterior distribution over the current run length (time since the last change point), and the state (category probability) before the last change point, given the data so far observed (category labels until trial  $t$ ,  $C_t = (C_1, \dots, C_t)$ ). We denote the current run length at the end of trial  $t$  by  $r_t$ , the current state by  $\pi_t$ , and the state before the last change point by  $\xi_t$ , where both  $\pi_t, \xi_t \in S_\pi$ . That is, if a change point occurs after trial  $t$  (i.e.,  $r_t = 0$ ), then the new category A probability will be  $\pi_t$  and the previous run's category probability  $\xi_t = \pi_{t-1}$ . If no change point occurs, both  $\pi$  and  $\xi$  remain unchanged. We use the notation  $C_t^{(r)}$  to indicate the set of observations (category labels) associated with the run  $r_t$ , which is  $C_{t-r_t+1:t}$  for  $r_t > 0$ , and  $\emptyset$  for  $r_t = 0$ . The range of times with a colon, e.g.,  $C_{t':t}$  indicates the sub-vector of  $C$  with elements from  $t'$  to  $t$  included.

Both of our tasks provide category feedback, so that at the end of trial  $t$  the observer has been informed of  $C_{1:t}$ . In S1 Appendix we derive the iterative Bayesian ideal-observer model. After each trial, the model calculates a posterior distribution over possible run lengths and previous probability states,  $P(r_t, \xi_t | C_{1:t})$ . The generative model makes it easy to calculate the conditional probability of the current state for a given run length and previous state,  $P(\pi_t | r_t, \xi_t, C_{1:t})$ . These two distributions may be combined, marginalizing (summing) across the unknown run length and previous states to yield the predictive probability distribution of the current state,  $P(\pi_t | C_{1:t})$ . Given this distribution over states, in both tasks the observer needs to determine the probability of each category. In particular,

$$P(C_{t+1} = A | C_{1:t}) = \mathbb{E}[\pi_t] = \sum_{\pi_t \in S_\pi} \pi_t P(\pi_t | C_{1:t}). \tag{1}$$

In the overt task, the ideal observer sets the current criterion to the optimal value  $z_t^{\text{opt}}$  based on the known category orientation distributions and the current estimate of category probabilities. Further, in the ideal and all subsequent models of the overt task, in addition to early sensory noise ( $\sigma_v$ ) we assume the actual setting is perturbed by adjustment noise ( $z_t = z_t^{\text{opt}} + \varepsilon_t$ , where  $\varepsilon_t \sim \mathcal{N}(0, \sigma_a^2)$ ).

In the covert task, the observer views a stimulus and makes a noisy measurement  $x_t$  of its true orientation  $s_t$  with noise variance  $\sigma_v^2$ . The prior category probability is combined with the noisy measurement to compute category A's posterior probability  $P(C_{t+1} = A | x_{t+1}, C_{1:t})$ . The observer responds "A" if that probability is greater than 0.5.

We consider the ideal-observer model (Bayes<sub>ideal</sub>) and four (suboptimal) variants thereof, which deviate from the ideal observer in terms of their beliefs about specific features of the experiment (Bayes<sub>r</sub>, Bayes<sub>m</sub>, Bayes<sub>\beta</sub>, and Bayes<sub>r,\pi,\beta</sub>). Two further variants of the Bayesian model (Bayes<sub>bias</sub> and Bayes<sub>r,\beta</sub>) are described in S1 Appendix. Crucially, all these models are "Bayesian" in that they compute a posterior over run length and probability state, but they differ with respect to the observer's assumptions about the generative model. Note that these models differ from the model provided by Gallistel and colleagues [28], which was used to model a task in which participants explicitly indicated perceived probability and change points.



**Bayes<sub>ideal</sub>.** The ideal Bayesian observer uses knowledge of the generative model to maximize expected gain. This model assumes knowledge of sensory noise, the category distributions, the set of potential probability states, and the run length distribution. Because observers were trained on the categories prior to completing each categorization task, assuming knowledge of the category distributions seems reasonable. Further, it is reasonable to assume knowledge of sensory noise as this is a characteristic of the observer. However, since observers were not told how often probability would change, nor were they told the set of potential probability states, observers may have had incorrect beliefs about the generative model. Thus, we developed additional Bayesian models (described below), in which observers could have potentially incorrect beliefs about different aspects of the generative model.

**Bayes<sub>r</sub>.** The Bayes<sub>r</sub> model allows for an observer to have an incorrect belief about the run length distribution. For a given discrete  $r$ , the observer believes that the run length distribution is drawn from a discrete uniform distribution,  $\sim \text{Unif}[\lfloor \frac{2}{3}r \rfloor, r]$ . We chose this particular interval rather than a more general one, because it is simple and includes the true generative distribution. For the ideal observer,  $r = 120$ .

**Bayes <sub>$\pi$</sub> .** The Bayes <sub>$\pi$</sub>  model allows for an observer to have an incorrect belief about the set of probability states. The veridical set of experimental states is five values of  $\pi$  ranging from 0.2 to 0.8. The Bayes <sub>$\pi$</sub>  model observer also assumes five possible values of  $\pi$  evenly spaced, but ranging from  $\pi_{\min}$  to  $\pi_{\max} = 1 - \pi_{\min}$ , where  $\pi_{\min}$  is a free parameter.

**Bayes <sub>$\beta$</sub> .** While incorrect assumptions about parameters of the generative model result in suboptimal behavior, suboptimality can also arise from prior biases (that is, incorrect hyperparameters). The Bayes <sub>$\beta$</sub>  model, like the Bayes<sub>ideal</sub> model, assumes knowledge of the generative model, but also includes a hyperprior Beta( $\beta, \beta$ ) that is applied after a change point. Thus, as  $\beta$  increases, the posterior belief over category probabilities is biased toward equal probability. For the ideal observer,  $\beta = 1$  (a uniform distribution).

**Bayes <sub>$r, \pi, \beta$</sub> .** The ‘incorrect-belief’ Bayesian models described above vary one assumption at a time, with at most a single additional free parameter. To get a sense of model performance when multiple assumptions are varied, the Bayes <sub>$r, \pi, \beta$</sub>  model allows for incorrect beliefs about the run length and probability distributions and a prior bias towards equal probability. Specifically, the Bayes <sub>$r, \pi, \beta$</sub>  model observer assumes that run length is drawn from the discrete distribution  $[r, r + \Delta r]$  and that there are five possible values of  $\pi$  evenly spaced ranging from  $\pi_{\min}$  to  $\pi_{\max}$ , where  $r \in [1, 200]$ ,  $\Delta r \in [1, 200]$ ,  $\pi_{\min} \in (0, 0.5)$ , and  $\pi_{\max} \in (0.5, 1)$  are all free parameters. The model also includes the hyperprior Beta( $\beta, \beta$ ), with  $\beta \in (0, 100]$ , that is applied after a change point. Because of the model’s complexity, excessive flexibility, and differences in the fitting procedure (see [Methods](#)), we only compared it to the best-fit model as determined by a separate model-comparison analysis. For the same reasons, note that this model is not included in the following figures or parameter tables. However, the results are summarized in Fig S1 in [S1 Appendix](#).

## Alternative models

In addition to the Bayesian models described above, we tested the following alternative models that do not require the observer to compute a posterior over run length and probability state. In each of the following models, assumptions vary about whether and how probability is estimated. In the Fixed Criterion (Fixed) model the observer assumes fixed probabilities. In the Exponential-Averaging (Exp), Exponential-Averaging with Prior Bias (Exp<sub>bias</sub>), and the Wilson et al. (2013) models, probability is estimated based on the recent history of categories. In the Reinforcement-Learning (RL) model, the decision criterion is updated following an error-driven learning rule with no assumptions about probability. Finally, the Behrens et al. (2007)

model is an alternative Bayesian model with fewer assumptions and restrictions than the Bayesian change-point detection model described above. We also tested three additional models that are described in [S1 Appendix](#).

Each of the following models, except for the RL model, computes an estimate of category probability ( $\hat{\pi}_{A,t}$ ) on each trial and the estimated probability of the alternative is  $\hat{\pi}_{B,t} = 1 - \hat{\pi}_{A,t}$ . On each trial, the optimal criterion  $z_{opt}$  is computed based on these estimated probabilities in the identical manner as for the Bayesian models. To make a categorization decision in the covert-criterion task, the criterion is applied to the noisy observation of the stimulus. In the overt-criterion task, the observer reports the criterion, which we again assume is corrupted by adjustment noise.

**Fixed.** While incorporating category probability into the decision process maximizes expected gain, a simpler strategy is to ignore changes in probability and fix the decision criterion throughout the block. In the fixed-criterion model, we assume equal category probability and the criterion is set to the neutral criterion:

$$z = \frac{1}{2}(\mu_B + \mu_A). \tag{2}$$

This model is equivalent to a model in which the likelihood ratio,  $\frac{P(x_t|C_t=A)}{P(x_t|C_t=B)}$ , is used to make categorization decisions rather than the posterior odds ratio in the covert-criterion task. This is a reasonable strategy for an observer who wants to make an informed decision, but is unsure about the current probability state and its rate of change.

**Exp.** The exponential-averaging model computes smoothed estimates of category probability by taking a weighted average of previously experienced category labels, giving more weight to recently experienced labels:

$$\begin{aligned} \hat{\pi}_{A,1} &= 0.5 \\ \hat{\pi}_{A,t+1} &= \alpha_{Exp} C_t + (1 - \alpha_{Exp}) \hat{\pi}_{A,t}, \end{aligned} \tag{3}$$

where  $\alpha_{Exp}$  is the smoothing factor,  $0 < \alpha_{Exp} < 1$ , and  $C_t = 1$  if category A is observed on trial  $t$  and  $C_t = 0$  otherwise. The time constant of memory decay for this model is  $\tau = \frac{-1}{\log(1-\alpha_{Exp})}$ . Mathematically this model is equivalent to a delta-rule [29] model based on an “error” that is the difference between the current category and the current probability estimate:  $\hat{\pi}_{A,t+1} = \hat{\pi}_{A,t} + \alpha_{Exp}(C_t - \hat{\pi}_{A,t})$ . The criterion  $z$  is then set to the optimal value based on this category-probability estimate.

**Exp<sub>bias</sub>.** The Exp<sub>bias</sub> model is identical to the Exp model, while also incorporating a common finding in the literature [2] known as conservatism (i.e., observers are biased towards the neutral criterion). On each trial an estimate of probability is computed as described in [Eq \(3\)](#), and averaged with a prior probability of 0.5:

$$\tilde{\pi}_{A,t} = w\hat{\pi}_{A,t} + \frac{1}{2}(1 - w), \tag{4}$$

where  $w$  is the weight given to the probability estimate  $\hat{\pi}_{A,t}$  and  $(1 - w)$  is the weight given to  $\pi_A = 0.5$ . The criterion  $z$  is also set to the optimal value based on this category-probability estimate.

**Wilson et al. (2013).** Due to the complexity of the full Bayesian change-point detection model, Wilson et al. [30] developed an approximation to the full model using a mixture of delta rules. In their reduced model, the full run-length distribution is approximated by maintaining a subset of all possible run lengths. These are referred to as nodes  $\{l_1, \dots, l_j\}$ . On each

trial, an estimate of probability is computed for each node,

$$\begin{aligned} \hat{\pi}_{A,1}^{l_1} &= 0.5 \\ \hat{\pi}_{A,t+1}^{l_i} &= \hat{\pi}_{A,t}^{l_i} + \frac{1}{l_i + v_p} (C_t - \hat{\pi}_{A,t}^{l_i}), \end{aligned} \tag{5}$$

where  $v_p$  is a model parameter that represents the strength of the observer’s prior towards equal category probability (pseudocounts of a Beta prior; larger  $v_p$  corresponds to stronger conservatism) and  $C_t$  is the category label. The learning rate for each node is thus  $\alpha_{\text{Wilson},l_i} = \frac{1}{l_i + v_p}$ . For a single-node model, this is identical to the Exp model described above Eq (3). To obtain an overall estimate of probability in a multiple-node model, estimates (Eq (5)) are combined by taking a weighted average. The weights,  $p(l_i|C_{1:t})$ , are updated on every trial. The update is dependent on an approximation to the change-point prior, which in turn depends on the hazard rate  $h$  (for details see Wilson et al. [30], Eqs 25-31, along with the later correction [31]). In other words, when the probability of a change is high, more weight is given to  $l_1$ , which results in a greater change in the overall probability estimate. But when the probability of a change is low, more weight is given to nodes greater than  $l_1$  and the probability estimate is more stable. For a three-node model probability is estimated as

$$\hat{\pi}_{A,t+1} = p(l_1|C_{1:t})\hat{\pi}_{A,t+1}^{l_1} + p(l_2|C_{1:t})\hat{\pi}_{A,t+1}^{l_2} + p(l_3|C_{1:t})\hat{\pi}_{A,t+1}^{l_3}. \tag{6}$$

For the purpose of the current experiment, we used a three-node model in which the first node was fixed ( $l_1 = 1$ ) and  $l_2$  and  $l_3$  were allowed to vary. In addition, the prior strength parameter,  $v_p$ , which modulates the learning rate, was also free. By allowing  $v_p$  to vary, this model is equivalent to the three-node model described by Wilson and colleagues [30] in which all nodes were free and  $v_p = 2$  was fixed. The hazard rate was set to 0.01, corresponding to an expected block length of 100 trials (the experimental mean), and we assumed a change occurred at  $t = 1$ , so that all the weight was given to  $l_1$ . We also tested an alternative model with a fixed value of  $v_p = 2$ , but it resulted in a worse fit.

**RL.** While the models described above make assumptions about how probability is estimated, it is also possible that observers simply update the decision criterion without estimating the current probability state. A model-free reinforcement-learning approach assumes the observer does not use knowledge of the environmental structure. Instead, the decision criterion, rather than an estimate of category probability, is updated. The observer updates the internal criterion ( $z$ ) on each trial according to the following delta rule:

$$z_{t+1} = \begin{cases} z_t, & \text{if correct} \\ z_t + \alpha_{\text{RL}}(x_t - z_t), & \text{if incorrect.} \end{cases} \tag{7}$$

Thus, the criterion is updated when negative feedback is received by taking a small step in the direction of the difference between the noisy observation ( $x_t$ ) and current criterion ( $z_t$ ), where the step size is scaled by the learning rate  $\alpha_{\text{RL}}$ . Nothing is changed after a correct response. Assuming effective training, we initialize the RL model by setting  $z_1 = \frac{1}{2}(\mu_B + \mu_A)$ .

**Behrens et al. (2007).** Behrens and colleagues [24] developed a model that uses Bayesian learning to update one’s belief about the statistics of the environment to use on the next trial. Importantly, a Behrens et al. (2007) observer updates beliefs about the probability of category A ( $\pi$ ), volatility ( $v$ ) and the volatility’s rate of change ( $k$ ) without having to store the entire history of outcomes ( $C_{1:t}$ ) and reward statistics. The posterior probability over  $\pi$ ,  $v$ , and  $k$ , or



current belief about the environmental statistics, is computed as follows:

$$p(\pi_{t+1}, v_{t+1}, k | C_{1:t+1}) \propto p(C_{t+1} | \pi_{t+1}) \times \int \left[ \int p(\pi_t, v_t, k | C_{1:t}) p(v_{t+1} | v_t, k) dv_t \right] p(\pi_{t+1} | \pi_t, v_{t+1}) d\pi_t, \tag{8}$$

where  $p(\pi_{t+1} | \pi_t, v_{t+1}) \sim \beta(\pi_t, \exp(v_{t+1}))$  is the probability transition function,  $p(v_{t+1} | v_t, k) \sim \mathcal{N}(v_t, \exp(k))$  is the volatility transition function, and  $p(C_{t+1} | \pi_{t+1})$  is the likelihood of the most recently observed category. The likelihood is equal to  $\pi_{t+1}$  when  $C_{t+1} = 1$  and  $1 - \pi_{t+1}$  when  $C_{t+1} = 0$ . Integration was done numerically over a five-dimensional grid ( $\pi_t, v_t, k, \pi_{t+1}, v_{t+1}$ ). We assumed the observer adopted non-informative uniform priors over  $\pi, v$ , and  $k$ . Thus, each dimension was split into 30 equally spaced bins with  $v$  and  $k$  spaced evenly in log space (as per positive parameters of unknown scale). To obtain an estimate of the probability of category A, we marginalized over  $v_{t+1}$  and  $k$  and computed the mean of the probability distribution,

$$\hat{\pi}_{A, t+1} = \int \pi_{t+1} p(\pi_{t+1}) d\pi_{t+1}. \tag{9}$$

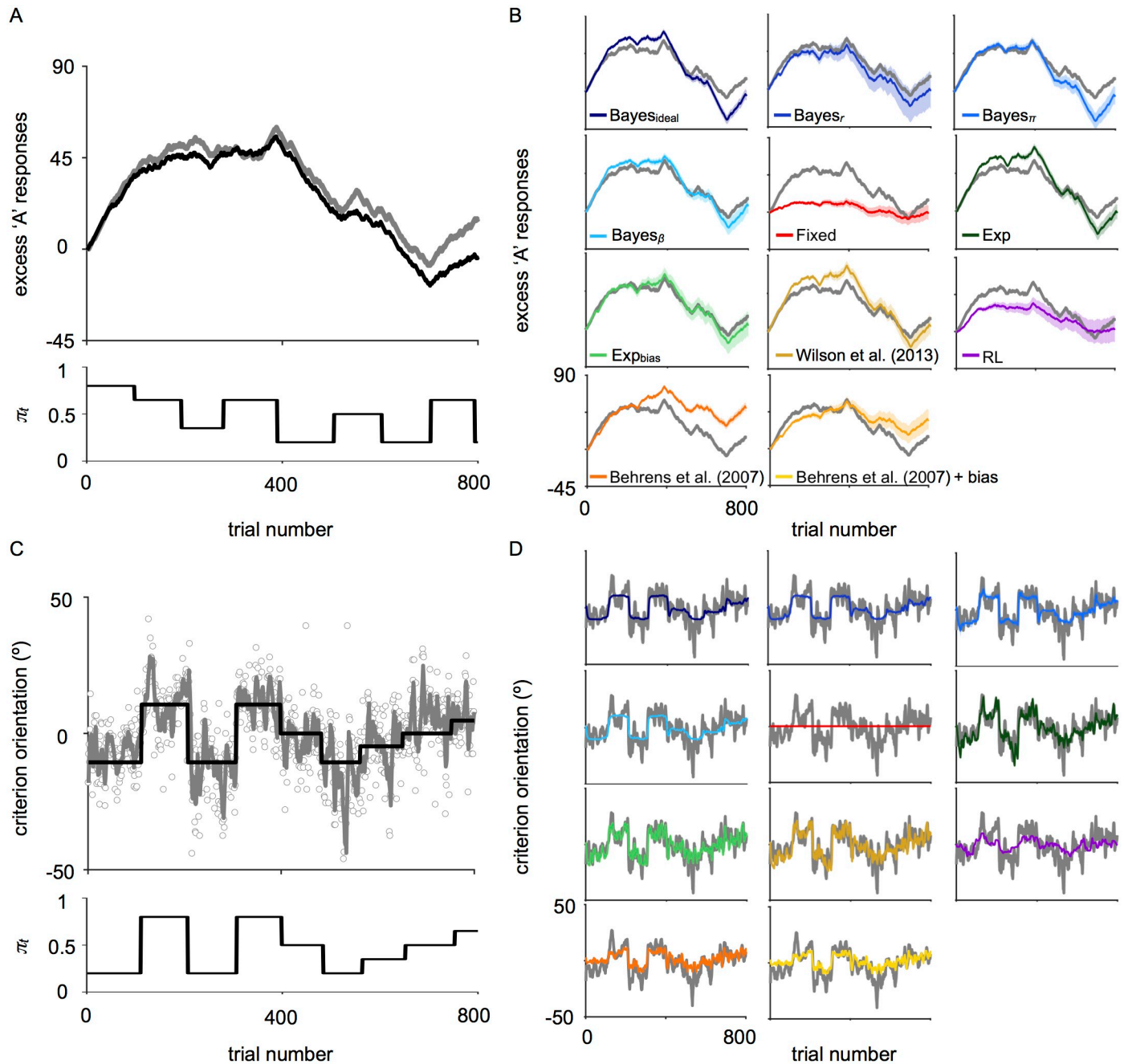
Other than sensory or adjustment noise, this model has no additional free parameters. Assuming perfect category knowledge, the criterion  $z$  is set to the optimal value based on this category-probability estimate.

**Behrens et al. (2007) with bias.** For the Behrens et al. (2007) model with a bias towards equal probability, category probability was estimated as in Eq 9 and a weighted average between  $\pi_A = 0.5$  and the model estimate was computed as in Eq 4. This model has one additional free parameter,  $w$ , which is the weight given to the probability estimate  $\hat{\pi}_{A,t}$  as in the Exp<sub>bias</sub> model. The criterion  $z$  is then set to the optimal value based on the weighted category-probability estimate.

### Raw data

Fig 2 shows raw data for a single observer in the covert (Fig 2A) and overt (Fig 2C) tasks. For visualization in the covert task, the ‘excess’ number of A responses is plotted as a function of trial (gray line in Fig 2A). To compute the ‘excess’ number of A responses, we subtracted  $\frac{1}{2}$  from the cumulative number of A responses. Thus, ‘excess’ A responses are constant for an observer who reported A and B equally often, increase when A is reported more, and decrease when A is reported less. To get a sense of how well the observer performed in the covert task, the number of ‘excess’ A trials (based on the actual category on each trial rather than the observer’s response) is shown in black (Fig 2A, top). For reference,  $\pi_A$  is shown as a function of trial (Fig 2A, bottom). From visual inspection, the observer reported A more often when  $\pi_{A,t} > 0.5$  and B more often when  $\pi_{A,t} < 0.5$ . Results for all observers in the covert task can be found in S1 Appendix (gray line in Figs S6A-S17A).

In the overt task, the orientation of the observer’s criterion setting, relative to the neutral criterion, is plotted as a function of trial (gray circles). For visualization, a running average was computed over a five-trial moving window (gray line). Here (Fig 2C, top), the black line represents the criterion on each trial, given perfect knowledge of the categories, sensory uncertainty, and category probability. While this is impossible for an observer to attain, we can see that the observer’s criterion follows the general trend. This suggests that observers update their criterion appropriately in response to changes in probability. That is, the criterion is set counter-clockwise from the neutral criterion when  $\pi_{A,t} > 0.5$ , and clockwise of neutral when  $\pi_{A,t} < 0.5$ .



**Fig 2. Behavioral data and model fits.** (A) Behavioral data from a representative observer (CWG) in the covert-criterion task. Top: The number of 'excess' A responses (i.e., the cumulative number of A responses minus  $\frac{1}{2}$ ) across trials. Gray line: Observer data. Black line: The true number of 'excess' A's. Bottom: The corresponding category probability. (B) Model fits (colored lines) for observer CWG in the covert-criterion task. Shaded regions: 68% CI for model fits. Gray: Observer data. (C) Behavioral data from a representative observer (GK) in the overt-criterion task. Top: The orientation of the criterion line relative to the neutral criterion as a function of trial number. Gray circles: raw settings. Gray line: running average over a 5-trial moving window. Black line: ideal criterion for an observer with perfect knowledge of the experimental parameters. Bottom: Category probability across trials. (D) Models fits (colored lines) for observer GK in the overt-criterion task. A running average was computed over a 5-trial window for visualization. Shaded regions: 68% CI on model fits. These are generally smaller than the model-fit line. Gray line: the running average computed from the observer's data.

<https://doi.org/10.1371/journal.pcbi.1006681.g002>

Fig 2C (bottom) shows  $\pi_A$  as a function of trial. Results for all observers in the overt task can be found in S1 Appendix (gray line in Figs S6B-S17B).

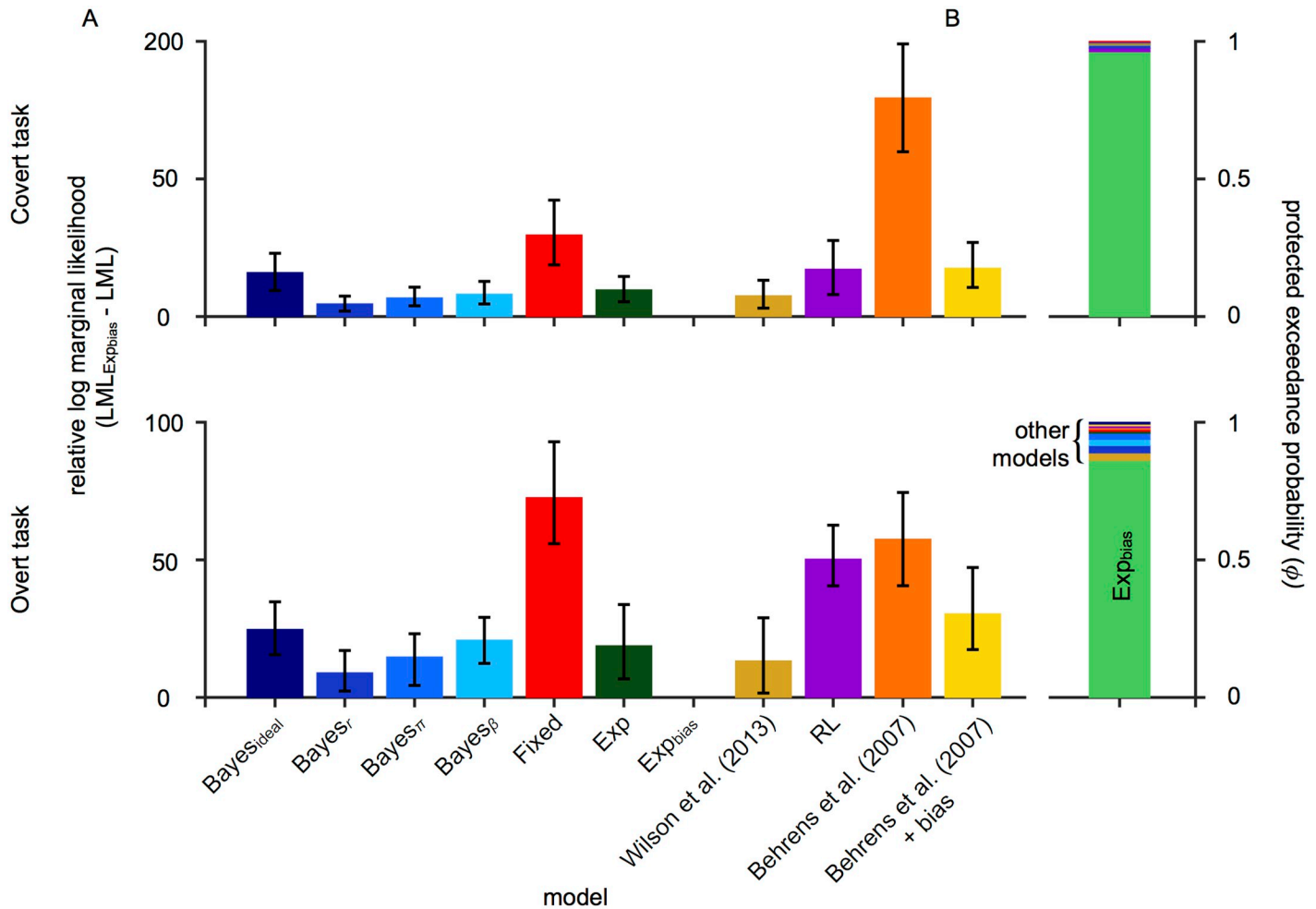
## Modeling results

**Model predictions.** A qualitative comparison of the behavioral data to the ground truth suggests that observers updated their criterion in response to changes in probability. However, it does not tell us how. To explore the mechanism underlying these changes, we compared the observers' data to multiple models. For each task and model, the mean model response for a representative subject is plotted in Fig 2B (covert task) and Fig 2D (overt task). Shaded regions indicate 68% CIs computed from the posterior over model parameters. Specifically, we sampled parameters from the posterior over model parameters and computed the model response for the given set of parameters. We then computed the standard deviation across model responses for a large sample (see Model visualization). Occasionally, shaded regions computed on fits are narrower than the data line. Qualitatively, most models captured observers' changing criteria, with the fixed model being much worse. Differences across models are especially pronounced in the overt task. Specifically, we see that the Exp, Exp<sub>bias</sub>, Wilson et al. (2013), RL, and Behrens et al. (2007) models capture changes in criterion that occur between change points that the ideal Bayesian change-point model generally fails to capture. The MAP model fits for each observer, model, and task can be found in S1 Appendix (colored lines in Figs S6-S17).

**Model comparison.** We quantitatively compared models by computing the log marginal likelihood (LML) for each subject, model, and task. The marginal likelihood is the probability of the data given the model, marginalized over model parameters (see Methods). Here, we report differences in log marginal likelihood ( $\Delta$ LML, also known as log Bayes factor) from the best model, so that larger  $\Delta$ LML correspond to worse models. We compare models both by looking at average performance across subjects (a fixed-effects analysis), and also via a Bayesian Model Selection approach (BMS; [32]) in which subjects are treated as random variables. With BMS, we estimate for each model its posterior frequency  $f$  in the population and its protected exceedance probability  $\phi$ , which is the probability that a given model is the most frequent model in the population, above and beyond chance [33].

Model comparison (Fig 3) favored the Exp<sub>bias</sub> model, which outperformed the second best model Bayes, (covert task:  $\Delta$ LML =  $9.27 \pm 2.86$ ; overt task:  $\Delta$ LML =  $8.96 \pm 4.01$ ; mean and SEM across observers) in the two tasks (covert task:  $t(10) = 3.99$ ,  $p = 0.003$ ; overt task:  $t(10) = 2.37$ ,  $p = 0.04$ ). Similarly, Bayesian model comparison performed at the group level also favored the Exp<sub>bias</sub> model (covert task:  $f = 0.42$  and  $\phi = 0.96$ ; overt task:  $f = 0.34$  and  $\phi = 0.86$ ). These results suggest that observers estimate probability by taking a weighted average of recently experienced categories with a bias towards  $\pi = 0.5$ .

To evaluate the performance of the more complex Bayes <sub>$r, \pi, \beta$</sub>  model, we computed both the Bayesian Information Criterion (BIC), obtained via maximum-likelihood estimation, and an approximate expected lower bound (ELBO) of the log marginal likelihood, obtained via Variational Bayesian Monte Carlo [34], a recent variational inference technique (see Methods for details). For a fair comparison, we also computed the same metrics for the best-fitting model from the analysis above (Exp<sub>bias</sub>), using the same fitting procedures. In this analysis, negative  $\Delta$ BIC and positive  $\Delta$ ELBO correspond to evidence in favor of the Exp<sub>bias</sub> model. For the covert task, the two models were indistinguishable in terms of both BIC ( $\Delta$ BIC =  $-2.75 \pm 6.07$ ;  $t(10) = -0.45$ ,  $p = 0.66$ ) and ELBO ( $\Delta$ ELBO =  $0.93 \pm 2.81$ ;  $t(10) = 0.33$ ,  $p = 0.75$ ). The same finding held for the overt task, in that models performed comparably in terms of both BIC ( $\Delta$ BIC =  $-2.14 \pm 10.58$ ;  $t(10) = -0.20$ ,  $p = 0.84$ ) and ELBO ( $\Delta$ ELBO =  $1.54 \pm 4.89$ ;  $t(10) = 0.32$ ,  $p = 0.76$ ).



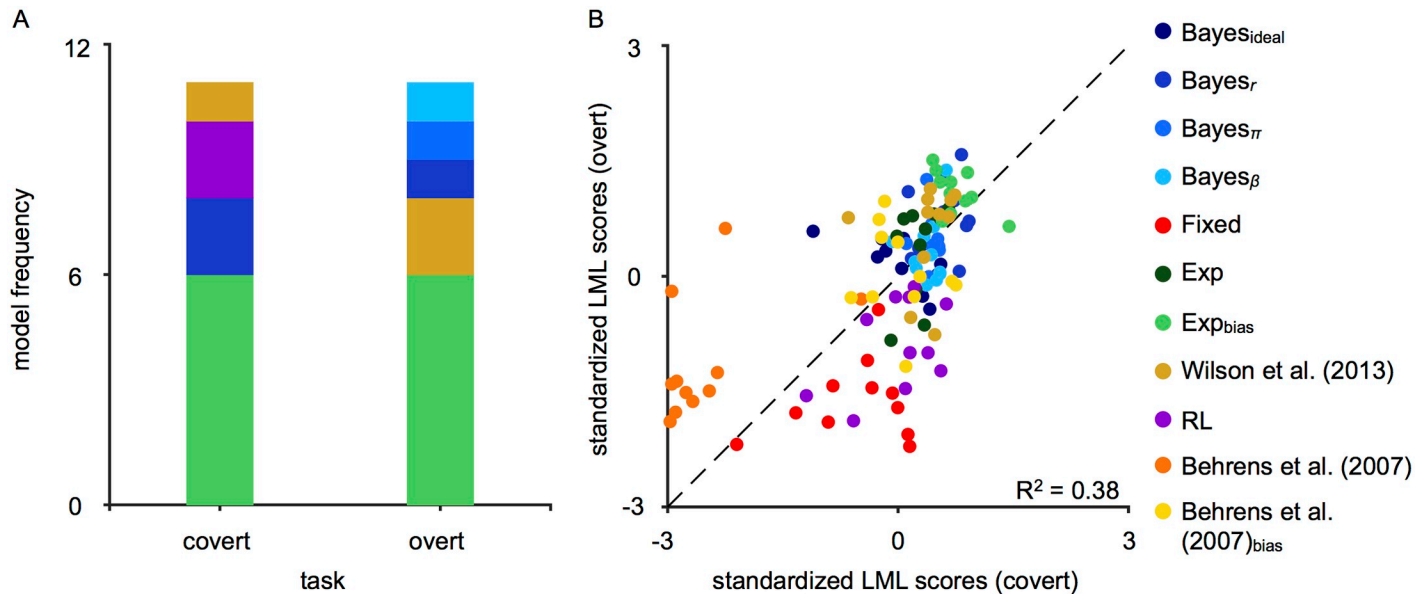
**Fig 3. Model comparison.** (A) Average log marginal likelihood (LML) scores relative to the best-fitting model, Exp<sub>bias</sub> (top: covert task; bottom: overt task). Lower scores indicate a better fit. Error bars: 95% CI (bootstrapped). (B) Bayesian model selection at the group level. The protected exceedance probability ( $\phi$ ) is plotted for each model. Models are stacked in order of decreasing probability.

<https://doi.org/10.1371/journal.pcbi.1006681.g003>

In short, according to both metrics the Exp<sub>bias</sub> model describes the data at least as well as the much more complex Bayes<sub>r,π,β</sub> model. The model fits and results of this comparison can be found in Fig S1 in S1 Appendix.

Fig 4A shows the number of observers that were best fit by each model for each task. To compare LML scores across tasks, and for the purpose of this analysis only, we standardized model scores for each observer and task. Standardized LML scores in the overt task are plotted as a function of standardized LML scores in the covert task in Fig 4B. We found a significant positive correlation,  $r = 0.60$ ,  $p < 0.01$ , indicating that models with higher LML scores in the covert task were also higher in the overt task. This result suggests that strategy was fairly consistent across tasks at the group level. In addition, there was more variance in model scores for worse-fitting models.

**Model parameters.** We examine here the parameter estimates of the Exp<sub>bias</sub> model, recalling that the two model parameters  $\alpha_{Exp}$  and  $w$  represent, respectively, the exponential smoothing factor and the degree to which observers exhibit conservatism. The maximum a posteriori (MAP) parameter estimates are plotted in Fig 5. Converting the smoothing factor to a time constant over trials, we found that the time constant in both tasks was well below the true rate



**Fig 4. Modeling results for individual observers.** (A) Model frequency. The number of observers best fit by each model plotted for each task. Models are stacked in order of decreasing frequency. (B) Comparison of LML scores across tasks. LML scores were standardized for each observer and task. Standardized LML scores in the overt task are plotted as a function of standardized LML scores in the covert task (colored data points). Black dashed line: identity line.

<https://doi.org/10.1371/journal.pcbi.1006681.g004>

of change (covert:  $\tau = [4.24, 7.18]$ ; overt:  $\tau = [3.48, 4.75]$ ;  $\tau_{\text{true}} = 100$  trials on average). We conducted paired-sample *t*-tests to compare the raw parameter estimates in the covert and overt tasks. We found a significant difference in  $w$  ( $t(10) = -2.55, p = 0.03$ ), suggesting that observers were more conservative in the covert than the overt task. No significant difference was found for  $\alpha_{\text{Exp}}$  ( $t(10) = -0.98, p = 0.35$ ).

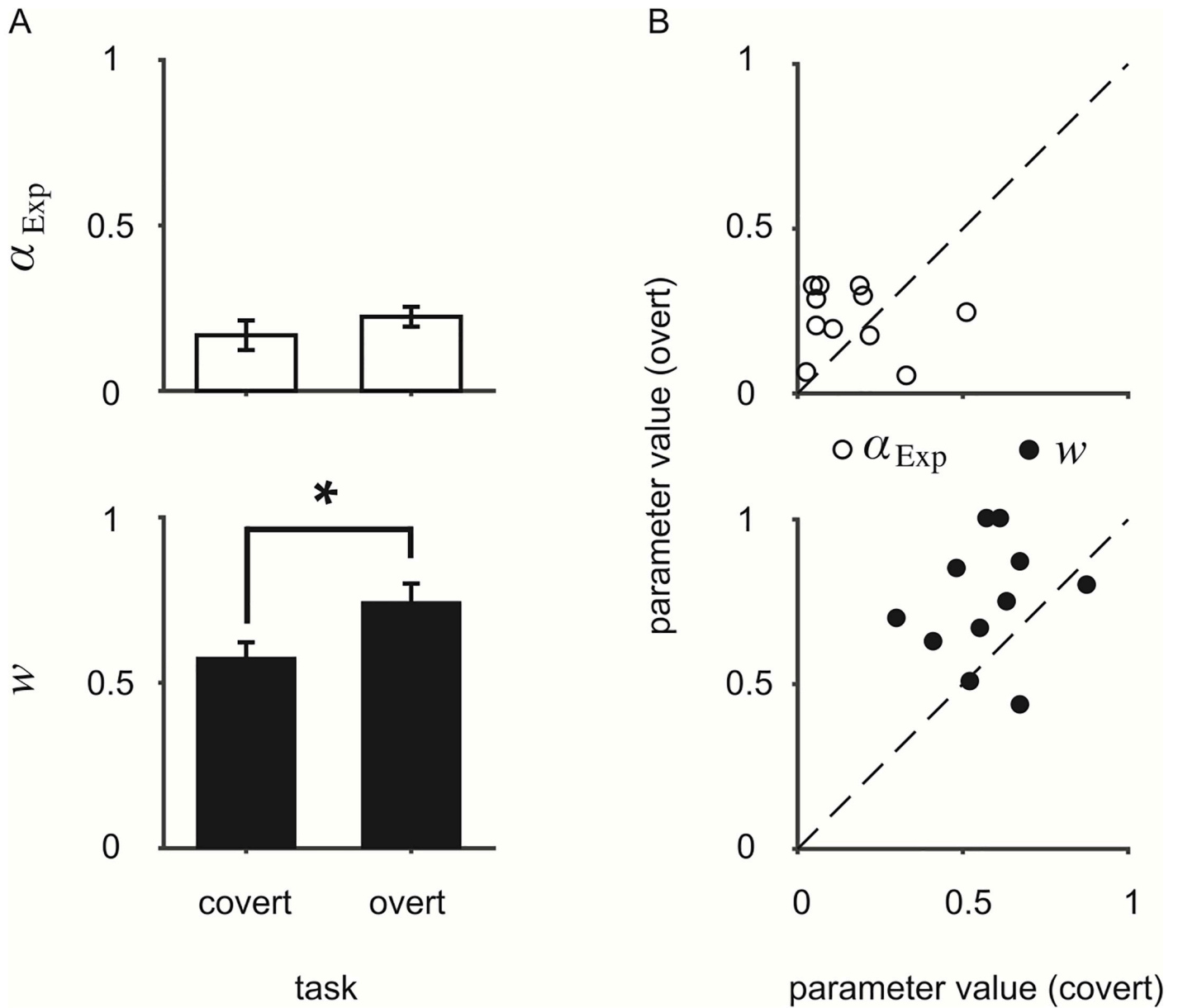
To investigate whether there was bias in the parameter-estimation procedure when fitting the Exp<sub>bias</sub> model, we also conducted a parameter-recovery analysis. Most parameters could be recovered correctly, except for adjustment variability ( $\sigma_a$ ) in the overt task, which we found to be overestimated on average (see Section 4 in S1 Appendix for details). Note that this bias in estimating  $\sigma_a$  in the overt task does not affect our model comparison, which is based on LMLs and not on point estimates.

While we might expect performance to be similar across tasks and observers (i.e., a correlation between the parameter fits in each task), no significant correlations were found ( $\alpha_{\text{Exp}}$ :  $r = -0.14, p = 0.67$ ;  $w$ :  $r = 0.16, p = 0.64$ ). Parameter estimates for all models, except the Bayes<sub>r,π,β</sub> model, are shown in Tables 1 and 2.

As the Bayes<sub>r,π,β</sub> contained a number of parameters that were unique to the model, parameter estimates for this model were not included in the tables below (covert:  $\sigma_v = 9.6 \pm 0.7$ ,  $r = 56.9 \pm 19.1$ ,  $\Delta r = 37.0 \pm 14.8$ ,  $\pi = .15 \pm .04$ ,  $\Delta\pi = .35 \pm .04$ , and  $\beta = 1.68 \pm .6$ ; overt:  $\sigma_a = 15.5 \pm 1.1$ ,  $r = 32.3 \pm 10.2$ ,  $\Delta r = 30.7 \pm 9.5$ ,  $\pi = .13 \pm .04$ ,  $\Delta\pi = .42 \pm .03$ , and  $\beta = 0.42 \pm .2$ ).

## Discussion

Although we know that people update decision criteria in response to explicit changes in prior probability, the effects of implicit changes in prior probability on decision-making behavior are less well known. In the present study, we used model comparison to investigate the mechanisms underlying decision-making behavior in an orientation-categorization task as prior probability changed. We tested a set of models that varied in both computational and memory



**Fig 5. Parameter estimates for the  $Exp_{bias}$  model.** (A) Mean  $\alpha_{exp}$  (top) and  $w$  (bottom) MAP values across observers. The \* denotes significance at the  $p < 0.05$  level. Error bars:  $\pm$ SEM. (B) Individual  $\alpha_{exp}$  (top) and  $w$  (bottom) MAP values from fits in the overt task as a function of MAP values from fits in the covert task. Black dashed lines: identity line.

<https://doi.org/10.1371/journal.pcbi.1006681.g005>

demands. Models were tested on data from both a covert- and overt-criterion task. A comprehensive approach, consisting of both qualitative and quantitative analysis, was performed to determine the best fitting model. We found that observers updated their decision criterion following changes in probability. Additionally, we observed systematic changes in the decision criterion during periods of stability, which was clearly evident in the overt-criterion data. While most models fit the data reasonably well qualitatively, model comparison slightly favored an exponential-averaging model with a bias towards equal probability, which was indistinguishable from a flexible variant of the Bayesian change-point detection model with



**Table 1. Maximum a posteriori parameter estimates  $\pm$ S.E. in the covert-criterion task.**

Model	$\sigma_v$	$r$	$\pi$	$\beta$	$\alpha$	$w$	$l_2$	$l_3$	$v_p$
Bayes <sub>ideal</sub>	9.7 $\pm$ 0.6	-	-	-	-	-	-	-	-
Bayes <sub>r</sub>	9.9 $\pm$ 0.8	51.5 $\pm$ 11.8	-	-	-	-	-	-	-
Bayes <sub><math>\pi</math></sub>	10.5 $\pm$ 0.8	-	0.32 $\pm$ 0.02	-	-	-	-	-	-
Bayes <sub><math>\beta</math></sub>	10.2 $\pm$ 0.7	-	-	18.4 $\pm$ 8.2	-	-	-	-	-
Fixed	10.9 $\pm$ 0.8	-	-	-	-	-	-	-	-
Exp	9.4 $\pm$ 0.7	-	-	-	0.04 $\pm$ 0.01	-	-	-	-
Exp <sub>bias</sub>	10.0 $\pm$ 0.8	-	-	-	0.17 $\pm$ 0.04	0.58 $\pm$ 0.05	-	-	-
Wilson et al. (2013)	9.4 $\pm$ 0.7	-	-	-	-	-	12.2 $\pm$ 2.8	118.5 $\pm$ 23.3	43.3 $\pm$ 14.4
RL	9.2 $\pm$ 0.7	-	-	-	0.26 $\pm$ 0.03	-	-	-	-
Behrens et al. (2007)	8.5 $\pm$ 0.3	-	-	-	-	-	-	-	-
Behrens et al. (2007) + bias	10.4 $\pm$ 0.7	-	-	-	-	0.31 $\pm$ 0.04	-	-	-

<https://doi.org/10.1371/journal.pcbi.1006681.t001>

incorrect beliefs and a bias towards equal probability. We can thus interpret the Exp<sub>bias</sub> model as a simpler explanation, in which observers update the decision criterion by combining on-line estimation of probability with an equal-probability prior. Ultimately, our results help explain decision-making behavior in situations in which people need to assess the probability of an outcome based on previous experience.

### Criterion updates in response to implicit changes in category probability

To determine the influence of prior probability on decision-making behavior, we examined changes in the decision criterion. First, we found that no participant was best fit by a fixed-criterion model. This finding suggests that observers update decision criteria in response to implicit changes in probability. This result is consistent with previous studies in which prior probability was explicit [2–9]. Further, this finding complements recent studies suggesting that individuals can learn and adapt to statistical regularities in changing environments [14–17, 24, 25, 35, 36]. Although this finding suggests that observers dynamically adjust decision criteria in response to changes in prior probability, it does not tell us how they do this (e.g., do observers compute on-line estimates of probability?). To uncover the mechanisms underlying changes in decision-making behavior, we compared multiple models ranging from the full Bayesian change-point detection model to a model-free reinforcement-learning (RL) model.

**Table 2. Maximum a posteriori parameter estimates  $\pm$ S.E. in the overt-criterion task.**

Model	$\sigma_a$	$r$	$\pi$	$\beta$	$\alpha$	$w$	$l_2$	$l_3$	$v_p$
Bayes <sub>ideal</sub>	17.2 $\pm$ 1.1	-	-	-	-	-	-	-	-
Bayes <sub>r</sub>	16.4 $\pm$ 1.1	70 $\pm$ 15	-	-	-	-	-	-	-
Bayes <sub><math>\pi</math></sub>	16.7 $\pm$ 1.1	-	0.17 $\pm$ 0.03	-	-	-	-	-	-
Bayes <sub><math>\beta</math></sub>	17.0 $\pm$ 1.1	-	-	13.2 $\pm$ 8.2	-	-	-	-	-
Fixed	19.4 $\pm$ 1.1	-	-	-	-	-	-	-	-
Exp	16.7 $\pm$ 1.1	-	-	-	0.09 $\pm$ 0.02	-	-	-	-
Exp <sub>bias</sub>	16.0 $\pm$ 1.0	-	-	-	0.22 $\pm$ 0.03	0.74 $\pm$ 0.06	-	-	-
Wilson et al. (2013)	16.4 $\pm$ 1.1	-	-	-	-	-	10.8 $\pm$ 2.0	46.4 $\pm$ 18.4	9.2 $\pm$ 5.1
RL	18.1 $\pm$ 1.2	-	-	-	0.21 $\pm$ 0.03	-	-	-	-
Behrens et al. (2007)	18.6 $\pm$ 1.0	-	-	-	-	-	-	-	-
Behrens et al. (2007) + bias	17.4 $\pm$ 1.0	-	-	-	-	0.66 $\pm$ 0.06	-	-	-

<https://doi.org/10.1371/journal.pcbi.1006681.t002>

## Systematic criterion fluctuations

How is the decision criterion set? Qualitatively, most models appear to fit the data reasonably well in the covert task. However, when we look at data from the overt task, while the Bayesian change-point detection models captured the overall trend, some variants failed to capture local fluctuations in the decision criterion observed during periods of stability (i.e., time intervals between change points). In other words, the criterion predicted by these models stabilized whereas the observers' behavior did not. This was less true when the Bayesian change-point detection model was allowed to vary more freely. In contrast, the exponential-averaging models continually update the observer's estimate of probability based on recently experienced categories. It can be difficult to discriminate a hierarchical model (e.g., the Bayesian change-point detection models) from flat models (e.g., our exponential models). For example, in a similar paradigm, Heilbron and Meyniel [37] found that predictions of change points could be similar for hierarchical and flat models. The addition of confidence judgments allowed them to discriminate the two model forms more readily.

How quickly observers updated this estimate is determined in the model by the decay-rate parameter. From our model fits, we found that observers had an average decay rate that was substantially smaller than the true run length distribution (on average 4.5 vs. 100 trials, respectively), leading to frequent, systematic fluctuations in decision criteria. Although we cannot directly observe these fluctuations in the covert task, because the estimated decay rate was not significantly different across tasks we can assume the fluctuations occurred in a similar manner. Like the exponential models, the RL model was also able to capture local fluctuations in the decision criterion. However, the amplitude of the changes in criterion predicted by the RL model was generally too low compared to the data. This discrepancy was especially clear in the overt task; no participant was best fit by the RL model. As a more fair comparison to the exponential models, we also fit the Bayesian model developed by Behrens and colleagues [24] with and without a bias towards equal probability. These models are more fair in that they do not require specifying a run-length distribution and allowed us to increase the number of possible probability states. While this allowed us to capture local fluctuations, model comparison favored the  $\text{Exp}_{\text{bias}}$  model. These results have two important implications. (1) It is important to test alternatives to Bayesian models: observers' behavior might be explained without requiring an internal representation of probability. (2) Using multiple tasks together with rigorous model comparison can provide additional insight into behavior (see also [38]). Here, the fluctuations in decision criteria between change points led to suboptimal behavior. Overall, our findings suggest that suboptimality arose from an incorrect, possibly heuristic inference process, that goes beyond mere sensory noise [39–41].

## A dual mechanism for criterion updating

While the  $\text{Bayes}_{r,\pi,\beta}$  and  $\text{Exp}_{\text{bias}}$  models fit the data equally well, the  $\text{Exp}_{\text{bias}}$  model provides a simpler explanation—considering that our experiment did not necessarily require subjects to build a hierarchical model of the task (see [37]). This suggests that observers compute on-line estimates of category probability based on recent experience. Further, the bias component of the model suggests that observers are conservative, as reflected in a long-term prior that categories are equally likely. The degree to which observers weight this prior varied across individuals and tasks. Taken together, these results suggest a dual mechanism for learning and incorporating prior probability into decisions. That is, there are (at least) two components to decision making that are acquired and updated at very different timescales.

Multiple-mechanism models have been used to describe behavior in decision-making [30] and motor behavior [42]. A model that combines delta rules predicts motor behavior better

than either delta rule alone [42]. Using a combination of delta rules [30], we were able to capture the local fluctuations in criterion that the ideal Bayesian model missed. However, we found that a constant weight on  $\pi = 0.5$  fit better than the multiple-node model described by Wilson and colleagues [30]. Temporal differences between their task and ours might explain some of the differences we observed, as changes occurred much more slowly in our experiment. Additionally, while fitting Wilson et al.'s model we set the hazard rate to 0.01 (the average rate of change), but observers had to learn this value throughout the experiment and may have had incorrect assumptions about the rate of change [21, 22].

### Explanations of conservatism

Conservatism was an important feature in our models, as it improved model fits each time it was incorporated. While we observed conservatism in both the covert- and overt-criterion tasks, we found that, on average, observers were significantly more conservative in the covert task. To understand why conservatism differs across tasks, we need to understand the differences between the tasks. While the generative model was identical across tasks, the observer's response differed. In the covert task, observers chose between two alternatives. In the overt task, observers selected a decision criterion. This is an important difference because it allows us to potentially rule out previous explanations of conservatism, such as the use of subjective probability [4], misestimation of the relative frequency of events [43, 44], and incorrect assumptions about the sensory distributions [45, 46]; these explanations predict similar levels of conservatism across tasks. On the other hand, conservatism may be due to the use of suboptimal decision rules. Probability matching is a strategy in which participants select alternatives proportional to their probability, and has been used to explain suboptimal behavior in forced-choice tasks in which observers choose between two or more alternatives [6, 47–50]. Thus, the higher levels of conservatism in the covert task may have been due to the use of a suboptimal decision rule like probability matching, which would effectively smooth the observer's response probability across trials. Probability matching is not applicable to responses in the overt task. Thus, the use of different decision rules may result in different levels of conservatism. These differences may also arise from an increase in uncertainty in the covert task due to less explicit feedback. An observer with greater uncertainty will rely more on the prior. Thus, conservatism may be the result of having a prior over criteria that interacts with task uncertainty. This can be tested by manipulating uncertainty over the generative model and measuring changes in conservatism. It is also possible that our training protocol introduced a bias towards equal probability and, due to the greater similarity between the covert and training tasks, the bias was stronger in the covert task. Finally, it is also possible that conservatism is the result of both the use of suboptimal decision rules and one or more of the previously proposed explanations.

### Incorrect assumptions about the generative model

While we tested a number of Bayesian change-point detection models that explored an array of assumptions about the generative model, clearly one could propose even more variants (e.g., a model with incorrect assumptions about category means and variance). Here, we analyzed one such assumption at a time. A simple way to expand the model space is via a factorial comparison [40, 51], which we did not consider here due to computational intractability and the combinatorial explosion of models. We did however, fit the Bayes<sub>r,π,β</sub> model, which simultaneously accounted for incorrect assumptions about the run length and probability-state distributions and a bias towards equal probability. We compared the fit to the Exp<sub>bias</sub> model. Notably these two models explained the data equally well, despite the higher flexibility of the

Bayesian model. We can thus interpret the  $\text{Exp}_{\text{bias}}$  model as a simpler explanation for Bayesian change-point detection behavior with largely erroneous beliefs. For all models except the RL model we assumed knowledge of the category distributions. However, Norton et al. [16] found that for the same orientation-categorization task, category means were estimated dynamically, even after prolonged training. Similarly, Gifford et al. [52] observed suboptimality in an auditory-categorization task and found that the data were best explained by a model with non-stationary categories and prior probability that was updated using the recent history of category exemplars. This occurred despite holding categories and probability constant within a block. In fact, similar effects of non-stationarity have been observed in several other studies [53–55]. In addition to non-stationary category means, observers may also have misestimated category variance [56], especially since learning category variance takes longer than learning category means [14].

## Conclusion

In sum, our results provide a computational model for how decision-making behavior changes in response to implicit changes in prior probability. Specifically, they suggest a dual mechanism for learning and incorporating prior probability that operate at different timescales. Importantly, this helps explain behavior in situations in which assessment of probability is learned through experience. Further, our results demonstrate the need to compare multiple models and the benefit of using tasks that provide a richer, more informative dataset.

## Methods

### Ethics statement

The Institutional Review Board at New York University approved the experimental procedure and observers gave written informed consent prior to participation.

### Participants

Eleven observers participated in the experiment (mean age 26.6, range 20–31, 8 females). All observers had normal or corrected-to-normal vision. One of the observers (EHN) was also an author.

### Apparatus and stimuli

Stimuli were presented on a gamma-corrected Dell Trinitron P780 CRT monitor with a 31.3 x 23.8° display, a resolution of 1024 x 768 pixels, a refresh rate of 85 Hz, and a mean luminance of 40 cd/m<sup>2</sup>. Observers viewed the display from a distance of 54.6 cm. The experiment was programmed in MATLAB [57] using the Psychophysics Toolbox [58, 59].

Stimuli were 4.0 x 1.0° ellipses presented at the center of the display on a mid-gray background. In both the orientation-discrimination and covert-criterion tasks, trials began with a central white fixation cross (1.2°). In the overt-criterion task, a yellow line with random orientation was presented at the center of the display (5.0 x 0.5°).

### Procedure

**Categories.** In the ‘categorization’ sessions described below, stimulus orientations were drawn from one of two categories (A or B). Category distributions were Gaussian with different means ( $\mu_B > \mu_A$ ) and equal variance ( $\sigma_A^2 = \sigma_B^2 = \sigma_s^2 = 100 \text{ deg}^2$ ). The mean of category A was chosen randomly from all possible orientations at the beginning of each session and the mean of category B was set so that  $d'$  was approximately 1.5, which was determined for each

observer based on the estimates of sensory uncertainty obtained during a ‘measurement session’ (Fig 1C).

**Sessions.** All observers participated in three 1-hour sessions. Observers completed a ‘measurement’ session first followed by two ‘categorization’ sessions. Observers completed the covert-criterion task in the first ‘categorization’ session followed by the overt-criterion task in the second, or vice versa (chosen randomly). At the beginning of each ‘categorization’ session, observers completed 200 trials of category training followed by 800 experimental trials. Prior to training, observers were provided with a detailed explanation of the category distributions and training task. After training, observers were provided with additional instructions about the subsequent ‘categorization’ task and told that the categories would remain constant for the remainder of the session but that category probability may change. Observers were not told how often category probability would change or the range of probability states.

### Measurement task

During the ‘measurement’ session, sensory uncertainty ( $\sigma_v$ ) was estimated using a two-interval, forced-choice, orientation-discrimination task in which two black ellipses were presented sequentially on a mid-gray background. The observer reported the interval containing the ellipse that was more clockwise by keypress. Once the response was recorded, auditory feedback was provided and the next trial began. An example trial sequence is shown in Fig S3A in [S1 Appendix](#).

The orientation of the ellipse in the first interval was chosen randomly on every trial from a uniform distribution ranging from  $-90^\circ$  to  $90^\circ$ . The orientation of the second ellipse was randomly oriented clockwise or counter-clockwise of the first. The difference in orientation between the two ellipses was selected using an adaptive staircase procedure. The minimum step-size was  $1^\circ$  and the maximum step-size was  $32^\circ$ . Each observer ran two blocks. In each block, four staircases (65 trials each) were interleaved (two 1-up, 2-down and two 1-up, 3-down staircases) and randomly selected on each trial. For analyses and results see [S1 Appendix](#).

### Category training

Each training trial was identical to a covert-criterion trial (Fig 1A). During training there was an equal chance that a stimulus was drawn from either category. To assess learning of category distributions, observers were asked to estimate the mean orientation of each category following training. The mean of each category was estimated exactly once. The order in which category means were estimated was randomized. For estimation, a black ellipse with random orientation was displayed in the center of the display. Observers slid the mouse to the right and left to rotate the ellipse clockwise and counterclockwise, respectively and clicked the mouse to indicate they were satisfied with the setting. No feedback was provided. We computed the proportion correct for each observer to ensure category learning by comparing it to the expected proportion correct ( $p(\text{correct}) = 0.77$ ) for  $d' = 1.5$ . Mean estimates are plotted in Fig S4C in [S1 Appendix](#) as a function of the true category means. We computed the average estimation error for each category and observer by subtracting the estimate from the true mean. From visual inspection, it appears that training was effective with the exception of one outlier, which we assume was a lapse.

### Categorization tasks

**Covert-criterion task.** In the covert-criterion task, observers categorized ellipses based on their orientation. The start of each trial ( $N_{\text{trials}} = 800$ ) was signaled by the appearance of a

central white fixation cross (500 ms). A black oriented ellipse was then displayed at the center of the screen (300 ms). Observers categorized the ellipse as A or B by keypress. Observers received feedback as to whether they were correct on every trial. Observers received a point for every correct response and aggregate points were displayed at the top of the screen to motivate observers. In addition, the fixation cross was displayed at the center of the screen in the color corresponding to the true category (category A: green; category B: red). The next trial began immediately. An example trial sequence is depicted in Fig 1A.

**Overt-criterion task.** In the overt-criterion task, observers completed an explicit version of the categorization task described above that was developed by Norton et al. [16]. At the beginning of each trial ( $N_{\text{trials}} = 800$ ), a line was displayed at the center of the screen. The orientation of the line was randomly selected from a uniform distribution ranging from  $-90$  to  $90^\circ$ . The observers' task was to rotate the line to indicate the criterion for that trial. Observers were explicitly instructed to set the criterion so that a subsequent category A stimulus would fall clockwise of the line and category B stimuli would fall counter-clockwise of it. Observers rotated the line clockwise or counterclockwise by sliding the mouse to the right or left and clicked the mouse to indicate their setting. Next, an ellipse was displayed under the criterion line in the color corresponding to the true category for 300 ms. Auditory feedback indicated whether the set criterion correctly categorized the ellipse. That is, observers were correct when a category A stimulus was clockwise of the criterion line or a category B stimulus was counter-clockwise of the line. Observers received a point for a correct response and aggregate points were displayed at the top of the screen. The next trial began immediately. An example trial sequence is depicted in Fig 1B.

## Model fitting

For fitting, all models had one free noise parameter. In the covert-criterion task, this was sensory noise ( $\sigma_v$ ). In the overt-criterion task, sensory noise was fixed and set to the value obtained in the 'measurement' session, but we included a noise parameter for the adjustment of the criterion line ( $\sigma_a$ ). Fixing one noise parameter in the overt-criterion task ameliorated potential issues of lack of parameter identifiability [60], and ensured that models had the same complexity across tasks. The Bayes<sub>ideal</sub>, Fixed, and Behrens et al. (2007) models had no additional parameters. The following suboptimal Bayesian models had one additional parameter: Bayes<sub>r</sub> ( $r$ ); Bayes <sub>$\pi$</sub>  ( $\pi_{\text{min}}$ ); Bayes <sub>$\beta$</sub>  ( $\beta$ ). The Bayes <sub>$r,\pi,\beta$</sub>  model had 5 additional parameters ( $r$ ,  $\Delta r$ ,  $\pi_{\text{min}}$ ,  $\Delta\pi$ ,  $\beta$ ). The Exp and RL models also only had one additional parameter ( $\alpha$ ), as did the Behrens et al. (2007) model with a bias towards equal probability ( $w$ ). The Exp<sub>bias</sub> model had two additional parameters ( $\alpha_{\text{exp}}$  and  $w$ ), and the Wilson et al. (2013) model had three ( $l_2$ ,  $l_3$ , and  $v_p$ ).

To fit each model, for each subject and task we computed the logarithm of the *unnormalized* posterior probability of the parameters,

$$\log p^*(\theta|\text{data}) = \log p(\text{data}|\theta, M) + \log p(\theta|M), \quad (10)$$

where data are category decisions in the covert-criterion task and criterion orientation in the overt-criterion task,  $M$  is a specific model, and  $\theta$  represents the model parameters (generally, a vector). The first term of Eq 10 is the log likelihood, while the second term is the prior over parameters (see below).

For each model (except Bayes <sub>$r,\pi,\beta$</sub> ) we evaluated Eq 10 on a cubic grid, with bounds chosen to contain almost all posterior probability mass. The grid for all models, except Wilson et al. (2013), consisted of 100 equally spaced values for each parameter. Due to the computational demands of the Wilson et al. (2013) model, we reduced the grid to 50 equally spaced values. The grid allowed us to approximate the full posterior distribution over parameters, and also to



evaluate the normalization constant for the posterior, which corresponds to the evidence or marginal likelihood, used as a metric of model comparison (see Model comparison). We reported as parameter estimates the best-fitting model parameters on the grid, that is the maximum-a-posteriori (MAP) values (see Tables 1 and 2). We used the full posterior distributions to compute posterior predictive distributions, that is, model predictions for visualization (see Model visualization), and to generate plausible parameter values for our model-recovery analysis. The Bayes<sub>r,π,β</sub> model had too many parameters to compute the marginal likelihood via brute force on a grid, so we adopted a different procedure for model comparison (see Model comparison).

**Priors over parameters.** We chose all priors to be uninformative. For noise parameters, we used uniform priors over a reasonably large range ([1°, 30°]). For α<sub>exp</sub>, α<sub>RL</sub>, and w we used a uniform prior from 0 to 1. For r we used a uniform prior from 2 to 200 trials. For π<sub>min</sub> we used a uniform prior from 0 to 0.5. For β, we used a uniform prior on the square root of the parameter value, ranging from 0 to 10. Instead of fitting the individual nodes in the Wilson et al. (2013) model, we fit the difference between nodes, i.e., δ<sub>1</sub> = l<sub>2</sub> - l<sub>1</sub> and δ<sub>2</sub> = l<sub>3</sub> - l<sub>2</sub>. We used a uniform prior on the square root of δ<sub>1</sub> ranging from 1.01 to 5 and on the square root of δ<sub>2</sub> ranging from 1.01 to 14. Finally, for v<sub>p</sub> we used a uniform prior in log space from 0 to 5.

### Response probability

**Covert-criterion task.** For each model, parameter combination, observer, and trial in the covert-criterion task, we computed the probability of choosing category A on each trial given a stimulus, s<sub>t</sub>, and all previously experienced categories, C<sub>1:t-1</sub>. In all models, the observer's current decision depends on the noisy measurement, x<sub>t</sub>, so the probability of responding A for a given stimulus s<sub>t</sub> is

$$p(\hat{C}_A | s_t, \mathbf{C}_{1:t-1}, \mathbf{M}, \theta) = \int p(\hat{C}_A | x_t, \mathbf{C}_{1:t-1}, \mathbf{M}, \theta) \mathcal{N}(x_t | s_t, \sigma_v^2) dx_t. \tag{11}$$

Because the current criterion setting in the RL model depends on the vector of all previous stimulus measurements, x<sub>1:t</sub>, the probability could not be computed analytically for this model. As an approximation, we used Monte Carlo simulations with 5000 sample measurement vectors. For each measurement vector, we applied the model's decision rule and approximated the probability by computing the proportion of times the model chose A out of all the simulations. For all models, we included a fixed lapse rate, λ = 10<sup>-4</sup>, that is the probability of a completely random response. The probability of choosing category A,  $\hat{C}_A$ , in the presence of lapses was then

$$p(\hat{C}_A | s_t, \mathbf{C}_{1:t-1}, \mathbf{M}, \theta) = (1 - \lambda)p(\hat{C}_A | s_t, \mathbf{C}_{1:t-1}, \mathbf{M}, \theta) + \frac{\lambda}{2}. \tag{12}$$

Effectively, the lapse rate acts as a regularization term that avoids excessive penalties to the likelihood of a model for outlier trials.

Next, assuming conditional independence between trials, we computed the log likelihood across all of the observer's choices, given each model and parameter combination

$$\log p(\text{data} | \mathbf{M}, \theta) = \sum_{t=1}^{N_{\text{trials}}} \log p(\hat{C}_A | s_t, \mathbf{C}_{1:t-1}, \mathbf{M}, \theta), \tag{13}$$

where t is the trial index and N<sub>trials</sub> is the total number of trials.

**Overt-criterion task.** For each model, parameter combination, observer, and trial in the overt-criterion task, we computed the decision criterion on each trial. For all models except

the RL model, the criterion was computed as in [S1 Appendix](#). For the RL model, the criterion was computed as in [Eq 7](#) for 5000 sample measurement vectors. For all models in the overt-criterion task, the criterion was corrupted by adjustment noise with variance  $\sigma_a^2$ , so that  $\hat{z}_t \sim \mathcal{N}(z_t, \sigma_a^2)$ , where  $z_t$  was the observer's chosen criterion at trial  $t$ , and  $\hat{z}_t$  was the actual reported criterion after adjustment noise. In addition, the observer had a chance of lapsing (e.g., a misclick), in which case the response was uniformly distributed in the range. Therefore, the probability that the observer reports the criterion  $\hat{z}_t$  was

$$p(\hat{z}_t | \mathbf{C}_{1:t-1}, \mathbf{M}, \theta) = (1 - \lambda)p(\hat{z}_t | \mathbf{C}_{1:t-1}, \mathbf{M}, \theta) + \frac{\lambda}{180}, \quad (14)$$

with  $\lambda = 5 \times 10^{-5}$ . As in the covert-criterion task, we computed the log likelihood across all trials by summing the log probability

$$\log p(\text{data} | \mathbf{M}, \theta) = \sum_{t=1}^T \log p(\hat{z}_t | \mathbf{C}_{1:t-1}, \mathbf{M}, \theta). \quad (15)$$

### Model comparison

To obtain a quantitative measure of model fit, for each observer, model (except  $\text{Bayes}_{r,\pi,\beta}$ ), and task we computed the log marginal likelihood (LML) by integrating over the parameters in [Eq 10](#),

$$\log p(\text{data} | \mathbf{M}) = \log \int p(\text{data} | \theta, \mathbf{M}) p(\theta | \mathbf{M}) d\theta. \quad (16)$$

To approximate the integral in [Eq 16](#), we marginalized across each parameter dimension using the trapezoidal method. Assuming equal probability across models, the marginal likelihood is proportional to the posterior probability of a given model and thus represents a principled metric for comparison that automatically accounts for both goodness of fit and model complexity via Bayesian Occam's razor [61]. Penalizing for model complexity is a desirable feature of a model-comparison metric to reduce overfitting.

In addition to the Bayesian model-comparison metric described above, we computed the Akaike Information Criterion (AIC) [62] for each of our models. AIC is one of many information criteria that penalize the maximum log likelihood by a term that increases with the number of parameters. LML and AIC results were consistent (see [S1 Appendix](#) for model comparison using AIC scores).

For comparison purposes, we report relative model-comparison scores,  $\Delta\text{LML}$  and  $\Delta\text{AIC}$ . We used bootstrapping to compute confidence intervals on the mean difference scores. Specifically, we simulated 10,000 sample data sets. For each simulated dataset we sampled, with replacement, 11 difference scores (the same number of difference scores as observers) and calculated the mean. To determine the 95% CI, we sorted the mean difference scores and determined the scores that corresponded to the 2.5 and 97.5 percentiles.

For an additional analysis at the group level, we used the random-effects Bayesian model selection analysis (BMS) developed by Stephan et al. [32] and expanded on by Rigoux et al. [33]. Specifically, using observers' LML scores we computed the protected exceedance probability  $\phi$  and the posterior model frequency for each model. Exceedance probability represents the probability that one model is the most frequent decision-making strategy in the population, given the group data, above and beyond chance. This analysis was conducted using the open-source software package Statistical Parametric Mapping (SPM12; <http://www.fil.ion.ucl.ac.uk/spm>).

To compare the Bayes <sub>$r, \pi, \beta$</sub>  and Exp<sub>bias</sub> models, we first fitted the data by maximum likelihood. For each of the two models, subject, and task we minimized the negative log likelihood of the data using Bayesian Adaptive Direct Search (BADS; <https://github.com/lacerbi/bads> [63]), taking the best result of 20 optimization runs with randomized starting points. Using the negative log likelihood from the fits, we computed the Bayesian Information Criterion (BIC), which penalizes the maximum log likelihood by a term that increases with the number of parameters in a similar way to AIC. We then estimated approximate lower bounds (ELBO) to the log marginal likelihood (and approximate posterior distributions, unused in our current analysis) via Variational Bayesian Monte Carlo (VBMC; <https://github.com/lacerbi/vbmc> [34]), taking the ‘best’ out of 10 variational optimization runs. VBMC combines variational inference and active-sampling Bayesian quadrature to perform approximate Bayesian inference in a sample-efficient manner. We validated the technique on the Exp<sub>bias</sub> model, by comparing the ELBO obtained via VBMC and the log marginal likelihood calculated via numerical integration. For all subjects and task, the two metrics differed by  $< 0.2$  points. As a diagnostic of convergence for VBMC, for each problem instance we also verified that the majority of variational optimization runs returned an ELBO less than 1 point away from the ‘best’ solution we used in the analysis. For the few datasets that exhibited a slightly larger variability in the ELBOs across optimization runs, such variability was still much lower than the standard error across subjects of the  $\Delta$ ELBO, thus not affecting the results of the model comparison.

### Model visualization

For each model, observer, and task, we randomly sampled 1000 parameter combinations from the joint posterior distribution with replacement. For each parameter combination, we simulated model responses using the same stimuli that were presented to the observer. Because the model output in the covert task was the probability of reporting category A, for each trial in a simulated dataset we simulated 10,000 model responses (i.e., category decisions), calculated the cumulative number of A’s for each simulated dataset, and averaged the results. The mean and standard deviation were computed across all simulated datasets in both tasks. Model fit plots show the mean response (colored line) with shaded regions representing one standard deviation from the mean. Thus, shaded regions represent a 68% confidence interval on model fits.

### Model recovery

To ensure that our models were discriminable, we performed a model-recovery analysis, details of which can be found in [S1 Appendix](#). In addition to the model-recovery analysis, we also performed a parameter-recovery analysis for the Exp<sub>bias</sub> model. This was done to determine whether our parameter estimation procedure was biased for each parameter and task.

### Supporting information

**S1 Appendix. Supplementary information.** Ideal observer model derivation; additional models; comparison of the Bayes <sub>$r, \pi, \beta$</sub>  and Exp<sub>bias</sub> models, model comparison with AIC; recovery analysis; measurement task; category training; individual model fits. (PDF)

### Acknowledgments

We would like to thank Chris Grimmick for helping with data collection. This work utilized the NYU IT High Performance Computing resources and services.

## Author Contributions

**Conceptualization:** Elyse H. Norton, Luigi Acerbi, Wei Ji Ma, Michael S. Landy.

**Data curation:** Elyse H. Norton.

**Formal analysis:** Elyse H. Norton, Luigi Acerbi, Wei Ji Ma, Michael S. Landy.

**Funding acquisition:** Michael S. Landy.

**Methodology:** Elyse H. Norton, Michael S. Landy.

**Project administration:** Michael S. Landy.

**Software:** Luigi Acerbi.

**Supervision:** Michael S. Landy.

**Validation:** Elyse H. Norton.

**Visualization:** Elyse H. Norton.

**Writing – original draft:** Elyse H. Norton, Luigi Acerbi.

**Writing – review & editing:** Elyse H. Norton, Luigi Acerbi, Wei Ji Ma, Michael S. Landy.

## References

1. Bernardo J, Smith A. Bayesian theory. New York: Wiley; 1994.
2. Green D, Swets JA. Signal detection theory and psychophysics. New York: Wiley; 1966.
3. Tanner WP, Jr. Theory of recognition. *J Acoust Soc Am*. 1956; 28:882–888. <https://doi.org/10.1121/1.1908504>
4. Ackermann JF, Landy MS. Suboptimal decision criteria are predicted by subjectively weighted probabilities and rewards. *Atten Percept Psychophys*. 2015; 77:638–658. <https://doi.org/10.3758/s13414-014-0779-z> PMID: 25366822
5. Ulehla ZJ. Optimality of perceptual decision criteria. *J Exp Psychol*. 1966; 71:564. <https://doi.org/10.1037/h0023007> PMID: 5909083
6. Healy AF, Kubovy M. Probability matching and the formation of conservative decision rules in a numerical analog of signal detection. *J Exp Psychol Hum Learn*. 1981; 7:344. <https://doi.org/10.1037/0278-7393.7.5.344>
7. Kubovy M, Healy AF. The decision rule in probabilistic categorization: What it is and how it is learned. *J Exp Psychol Gen*. 1977; 106:427. <https://doi.org/10.1037/0096-3445.106.4.427>
8. Healy AF, Kubovy M. The effects of payoffs and prior probabilities on indices of performance and cutoff location in recognition memory. *Mem Cognit*. 1978; 6:544–553. <https://doi.org/10.3758/BF03198243> PMID: 24203388
9. Maddox WT. Toward a unified theory of decision criterion learning in perceptual categorization. *J Exp Anal Behav*. 2002; 78:567–595. <https://doi.org/10.1901/jeab.2002.78-567> PMID: 12507020
10. Barron G, Erev I. Small feedback-based decisions and their limited correspondence to description-based decisions. *J Behav Decis Mak*. 2003; 16:215–233. <https://doi.org/10.1002/bdm.443>
11. Hertwig R, Barron G, Weber EU, Erev I. Decisions from experience and the effect of rare events in risky choice. *Psychol Sci*. 2004; 15:534–539. <https://doi.org/10.1111/j.0956-7976.2004.00715.x> PMID: 15270998
12. Bohil CJ, Wismer AJ. Implicit learning mediates base rate acquisition in perceptual categorization. *Psychon Bull Rev*. 2015; 22:586–593. <https://doi.org/10.3758/s13423-014-0694-2> PMID: 25037267
13. Wismer AJ, Bohil CJ. Base-rate sensitivity through implicit learning. *PLoS One*. 2017; 12(6):e0179256. <https://doi.org/10.1371/journal.pone.0179256> PMID: 28632779
14. Berniker M, Voss M, Körding K. Learning priors for Bayesian computations in the nervous system. *PLoS One*. 2010; 5(9):e12686. <https://doi.org/10.1371/journal.pone.0012686> PMID: 20844766
15. Summerfield C, Behrens TE, Koehlin E. Perceptual classification in a rapidly changing environment. *Neuron*. 2011; 71:725–736. <https://doi.org/10.1016/j.neuron.2011.06.022> PMID: 21867887

16. Norton EH, Fleming SM, Daw ND, Landy MS. Suboptimal criterion learning in static and dynamic environments. *PLoS Comput Biol.* 2017; 13(1):e1005304. <https://doi.org/10.1371/journal.pcbi.1005304> PMID: 28046006
17. Nassar MR, Wilson RC, Heasley B, Gold JI. An approximately Bayesian delta-rule model explains the dynamics of belief updating in a changing environment. *J Neurosci.* 2010; 30:12366–12378. <https://doi.org/10.1523/JNEUROSCI.0822-10.2010> PMID: 20844132
18. Landy MS, Trommershäuser J, Daw ND. Dynamic estimation of task-relevant variance in movement under risk. *J Neurosci.* 2012; 32:12702–12711. <https://doi.org/10.1523/JNEUROSCI.6160-11.2012> PMID: 22972994
19. Acerbi L, Wolpert DM, Vijayakumar S. Internal representations of temporal statistics and feedback calibrate motor-sensory interval timing. *PLoS Comput Biol.* 2012; 8(11):e1002771. <https://doi.org/10.1371/journal.pcbi.1002771> PMID: 23209386
20. Sato Y, Körding KP. How much to trust the senses: Likelihood learning. *J Vis.* 2014; 14(13):13. <https://doi.org/10.1167/14.13.13> PMID: 25398975
21. Glaze CM, Kable JW, Gold JI. Normative evidence accumulation in unpredictable environments. *eLife.* 2015; 4. <https://doi.org/10.7554/eLife.08825> PMID: 26322383
22. Wilson RC, Nassar MR, Gold JI. Bayesian online learning of the hazard rate in change-point problems. *Neural Comput.* 2010; 22:2452–2476. [https://doi.org/10.1162/NECO\\_a\\_00007](https://doi.org/10.1162/NECO_a_00007) PMID: 20569174
23. Meyniel F, Schlunegger D, Dehaene S. The sense of confidence during probabilistic learning: A normative account. *PLoS Comput Biol.* 2015; 11(6):e1004305. <https://doi.org/10.1371/journal.pcbi.1004305> PMID: 26076466
24. Behrens TE, Woolrich MW, Walton ME, Rushworth MF. Learning the value of information in an uncertain world. *Nat Neurosci.* 2007; 10:1214. <https://doi.org/10.1038/nn1954> PMID: 17676057
25. Zylberberg A, Wolpert DM, Shadlen MN. Counterfactual reasoning underlies the learning of priors in decision making. *Neuron.* 2018; 99(5):1083–1097. <https://doi.org/10.1016/j.neuron.2018.07.035> PMID: 30122376
26. Tversky A, Kahneman D. Judgment under uncertainty: Heuristics and biases. *Science.* 1974; 185:1124–1131. <https://doi.org/10.1126/science.185.4157.1124> PMID: 17835457
27. Adams RP, MacKay DJ. Bayesian online changepoint detection. *arXiv preprint arXiv:07103742.* 2007.
28. Gallistel CR, Krishan M, Liu Y, Miller RR, Latham PE. The perception of probability. *Psychol Rev.* 2014; 121:96–123. <https://doi.org/10.1037/a0035232> PMID: 24490790
29. Rescorla RA, Wagner AR. A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In: Black AH, Prokasy WF, editors. *Classical conditioning II: Current research and theory.* New York: Appleton-Century-Crofts; 1972. p. 64–99.
30. Wilson RC, Nassar MR, Gold JI. A mixture of delta-rules approximation to Bayesian inference in change-point problems. *PLoS Comput Biol.* 2018; 14(6):e1006210. <https://doi.org/10.1371/journal.pcbi.1003150> PMID: 23935472
31. Wilson RC, Nassar MR, Gold JI. Correction: A mixture of delta-rules approximation to Bayesian inference in change-point problems. *PLoS Comput Biol.* 2013; 9(7):e1003150. <https://doi.org/10.1371/journal.pcbi.1006210> PMID: 29944654
32. Stephan KE, Penny WD, Daunizeau J, Moran RJ, Friston KJ. Bayesian model selection for group studies. *Neuroimage.* 2009; 46:1004–1017. <https://doi.org/10.1016/j.neuroimage.2009.03.025> PMID: 19306932
33. Rigoux L, Stephan KE, Friston KJ, Daunizeau J. Bayesian model selection for group studies—revisited. *Neuroimage.* 2014; 84:971–985. <https://doi.org/10.1016/j.neuroimage.2013.08.065> <https://www.ncbi.nlm.nih.gov/pubmed/24018303>
34. Acerbi L. Variational Bayesian Monte Carlo. In: *Advances in Neural Information Processing Systems.* vol. 31; 2018. p. 8213–8223.
35. Burge J, Ernst MO, Banks MS. The statistical determinants of adaptation rate in human reaching. *J Vis.* 2008; 8(4):20. <https://doi.org/10.1167/8.4.20> PMID: 18484859
36. Qamar AT, Cotton RJ, George RG, Beck JM, Prezhdo E, Laudano A, et al. Trial-to-trial, uncertainty-based adjustment of decision boundaries in visual categorization. *Proc Natl Acad Sci USA.* 2013; 110:20332–20337. <https://doi.org/10.1073/pnas.1219756110> PMID: 24272938
37. Heilbron M, Meyniel F. Confidence resets reveal hierarchical adaptive learning in humans. *PLoS Comput Biol.* 2019; 15(4). <https://doi.org/10.1371/journal.pcbi.1006972> PMID: 30964861
38. Acerbi L, Dokka K, Angelaki DE, Ma WJ. Bayesian comparison of explicit and implicit causal inference strategies in multisensory heading perception. *PLoS Comput Biol.* 2018; 14(7):e1006110. <https://doi.org/10.1371/journal.pcbi.1006110> PMID: 30052625



39. Beck JM, Ma WJ, Pitkow X, Latham PE, Pouget A. Not noisy, just wrong: The role of suboptimal inference in behavioral variability. *Neuron*. 2012; 74(1):30–39. <https://doi.org/10.1016/j.neuron.2012.03.016> PMID: 22500627
40. Acerbi L, Vijayakumar S, Wolpert DM. On the origins of suboptimality in human probabilistic inference. *PLoS Comput Biol*. 2014; 10(6):e1003661. <https://doi.org/10.1371/journal.pcbi.1003661> PMID: 24945142
41. Drugowitsch J, DeAngelis GC, Angelaki DE, Pouget A. Tuning the speed-accuracy trade-off to maximize reward rate in multisensory decision-making. *eLife*. 2015; 4:e06678. <https://doi.org/10.7554/eLife.06678> PMID: 26090907
42. Izawa J, Shadmehr R. Learning from sensory and reward prediction errors during motor adaptation. *PLoS Comput Biol*. 2011; 7(3):e1002012. <https://doi.org/10.1371/journal.pcbi.1002012> PMID: 21423711
43. Attneave F. Psychological probability as a function of experienced frequency. *J Exp Psychol*. 1953; 46:81. <https://doi.org/10.1037/h0057955> PMID: 13084849
44. Varey CA, Mellers BA, Birnbaum MH. Judgments of proportions. *J Exp Psychol Hum Percept Perform*. 1990; 16:613. <https://doi.org/10.1037/0096-1523.16.3.613> PMID: 2144575
45. Maloney LT, Thomas EA. Distributional assumptions and observed conservatism in the theory of signal detectability. *J Math Psychol*. 1991; 35:443–470. [https://doi.org/10.1016/0022-2496\(91\)90043-S](https://doi.org/10.1016/0022-2496(91)90043-S)
46. Kubovy M. A possible basis for conservatism in signal detection and probabilistic categorization tasks. *Percept Psychophys*. 1977; 22:277–281. <https://doi.org/10.3758/BF03199690>
47. Lee W, Janke M. Categorizing externally distributed stimulus samples for unequal molar probabilities. *Psychol Rep*. 1965; 17:79–90. <https://doi.org/10.2466/pr0.1965.17.1.79> PMID: 5826504
48. Murray RF, Patel K, Yee A. Posterior probability matching and human perceptual decision making. *PLoS Comput Biol*. 2015; 11(6):e1004342. <https://doi.org/10.1371/journal.pcbi.1004342> PMID: 26079134
49. Thomas EA, Legge D. Probability matching as a basis for detection and recognition decisions. *Psychol Rev*. 1970; 77:65. <https://doi.org/10.1037/h0028579>
50. Wozny DR, Beierholm UR, Shams L. Probability matching as a computational strategy used in perception. *PLoS Comput Biol*. 2010; 6(8):e1000871. <https://doi.org/10.1371/journal.pcbi.1000871> PMID: 20700493
51. van den Berg R, Awh E, Ma WJ. Factorial comparison of working memory models. *Psychol Rev*. 2014; 121:124. <https://doi.org/10.1037/a0035234> PMID: 24490791
52. Gifford AM, Cohen YE, Stocker AA. Characterizing the impact of category uncertainty on human auditory categorization behavior. *PLoS Comput Biol*. 2014; 10(7):e1003715. <https://doi.org/10.1371/journal.pcbi.1003715> PMID: 25032683
53. Petzschner FH, Glasauer S. Iterative Bayesian estimation as an explanation for range and regression effects: A study on human path integration. *J Neurosci*. 2011; 31:17220–17229. <https://doi.org/10.1523/JNEUROSCI.2028-11.2011> PMID: 22114288
54. Raviv O, Ahissar M, Loewenstein Y. How recent history affects perception: The normative approach and its heuristic approximation. *PLoS Comput Biol*. 2012; 8(10):e1002731. <https://doi.org/10.1371/journal.pcbi.1002731> PMID: 23133343
55. Fischer J, Whitney D. Serial dependence in visual perception. *Nat Neurosci*. 2014; 17:738–743. <https://doi.org/10.1038/nn.3689> PMID: 24686785
56. Zylberberg A, Roelfsema PR, Sigman M. Variance misperception explains illusions of confidence in simple perceptual decisions. *Conscious Cogn*. 2014; 27:246–253. <https://doi.org/10.1016/j.concog.2014.05.012> PMID: 24951943
57. MATLAB. version 7.10.0 (R2010a). Natick, Massachusetts: The MathWorks Inc.; 2010.
58. Brainard DH, Vision S. The psychophysics toolbox. *Spat Vis*. 1997; 10:433–436. <https://doi.org/10.1163/156856897X00357> PMID: 9176952
59. Pelli DG. The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spat Vis*. 1997; 10:437–442. <https://doi.org/10.1163/156856897X00366> PMID: 9176953
60. Acerbi L, Ma WJ, Vijayakumar S. A framework for testing identifiability of Bayesian models of perception. In: *Advances in Neural Information Processing Systems*; 2014. p. 1026–1034.
61. MacKay DJ. *Information Theory, Inference and Learning Algorithms*. Cambridge, England: Cambridge University Press; 2003.
62. Akaike H. Information theory and an extension of the maximum likelihood principle. In: *Proceedings of the 2nd International Symposium on Information*; 1973. p. 267–281.
63. Acerbi L, Ma WJ. Practical Bayesian optimization for model fitting with Bayesian Adaptive Direct Search. In: *Advances in Neural Information Processing Systems*. vol. 30; 2017. p. 1836–1846.