Limitations of a proposed correction for slow drifts in decision criterion

Diksha Gupta[1], Carlos D. Brody[1,2]

[1]Princeton Neuroscience Institute, [2]Howard Hughes Medical Institute, Princeton, NJ 08544, USA

Abstract

Trial history biases in decision-making tasks are thought to reflect systematic updates of decision variables, therefore their precise nature informs conclusions about underlying heuristic strategies and learning processes. However, random drifts in decision variables can corrupt this inference by mimicking the signatures of systematic updates. Hence, identifying the trial-by-trial evolution of decision variables requires methods that can robustly account for such drifts. Recent studies (Lak'20, Mendonça'20) have made important advances in this direction, by proposing a convenient method to correct for the influence of slow drifts in decision criterion, a key decision variable. Here we apply this correction to a variety of updating scenarios, and evaluate its performance. We show that the correction fails for a wide range of commonly assumed systematic updating strategies, distorting one's inference away from the veridical strategies towards a narrow subset. To address these limitations, we propose a model-based approach for disambiguating systematic updates from random drifts, and demonstrate its success on real and synthetic datasets. We show that this approach accurately recovers the latent trajectory of drifts in decision criterion as well as the generative systematic updates from simulated data. Our results offer recommendations for methods to account for the interactions between history biases and slow drifts, and highlight the advantages of incorporating assumptions about the generative process directly into models of decision-making.

Introduction

Animals' choices in perceptual decision-making tasks often display a dependence on the recent history of stimuli, choices, and outcomes. This dependence is thought to arise from systematic updating of decision variables from trial to trial. These updates may reflect ongoing learning (Dayan and Daw 2008), for instance an agent learning to perform a perceptual categorization task might update its beliefs about the prior probabilities of the different categories (Yu and Cohen 2008; Zhang, Huang, and Yu 2014), the category boundary separating them (Drugowitsch et al. 2019; Mendonça et al. 2020), or the values of the available actions (Lak, Okun, et al. 2020; Lak, Hueske, et al. 2020; Pisupati et al. 2021). Alternatively, systematic updates may reflect heuristic strategies adopted by decision-makers due to resource constraints or mismatched objectives (Gigerenzer and Gaissmaier 2011; Abrahamyan et al. 2016; Gardner 2019). For example, a reward-seeking agent might be prone to repeating rewarded actions and avoid unrewarded

1

ones, i.e. follow a win-stay lose-switch strategy, similarly high costs of motor switching may encourage an agent to repeat previously chosen actions, yielding stay biases. These learning processes and heuristic strategies are often identified by the distinct patterns of choices they predict.

In addition to these systematic updates, decision variables may drift randomly from trial to trial (Dutilh et al. 2012; Lak, Hueske, et al. 2020; Ashwood et al. 2020; Cowley et al. 2020; Hennig et al. 2021; Roy et al. 2021; Lyamzin et al. 2021). These "unsystematic" drifts may arise from history-independent fluctuations in internal states such as attention, arousal and motivation, or from other unknown sources of noise (Renart and Machens 2014). However, despite their history independence, unsystematic drifts may nevertheless produce correlations in choices that obscure the effect of systematic updates and thereby complicate their inference (Lak, Hueske, et al. 2020; Sugawara and Katahira 2021).

An important recent study (Lak, Hueske, et al. 2020) considered the challenge posed by unsystematic drifts in one key decision variable, the decision criterion, i.e. the threshold for choosing one alternative over the other. The authors showed that a slow drift in criterion produces an apparent history dependence that mimics the signatures of updates from a systematic learning process. Hence, in order to remove the influence of these slow drifts, they proposed a model-free (MF) correction (Lak, Hueske, et al. 2020), similar in spirit to an approach employed by (Mendonça et al. 2020). They reasoned that slow drifts would produce correlations in choices across multiple successive trials. Thus behavior in a given trial would be correlated with both immediately previous and immediately future trials, giving the appearance that experience in the current trial would not just influence future choices (a causal process), but would also spuriously influence past choices (an *acausal*, therefore not biologically plausible process). Hence, they posited that the effect of slow drifts of decision criterion can be removed by s*ubtracting the influence of the current trial on past choice (the acausal component) from the influence of the current trial on future choice*. The simplicity of this correction is appealing and has already invited other authors to use it on their datasets. Indeed, recently the International Brain Laboratory reported that their dataset displayed win-stay lose-stay behavior, but when they accounted for possible slow drifts using the MF correction, they revealed instead a win-stay lose-switch strategy (The International Brain Laboratory et al. 2021).

While Lak et al. 2020 demonstrated convincingly that the MF correction can remove the influence of random slow drifts, they did so in the absence of any systematic updates of decision-variables. It remains to be shown if the MF correction is robust to the presence of learning or other heuristic strategies. If the MF correction indiscriminately removes correlations across choices, including those expected from systematic trial-by-trial updating of decision-relevant variables, that would considerably undermine its usefulness.

Here, using simulations we show that the proposed model-free (MF) correction fails in the presence of systematic trial-by-trial updating, yielding potentially misleading conclusions about the nature of underlying behavioral strategies. We show that the correction erroneously removes the effect of trial-by-trial updates that produce correlations across choices, thereby misidentifying a variety of update strategies as belonging to a narrow subspace of win-stay lose-switch or win-switch lose-stay biases. We advocate for an alternate, model-based approach for disambiguating systematic updates from random drifts, and demonstrate the success of this approach by fitting synthetic data. Finally, we apply the two approaches to real data from rodents (Brunton et al. 2013) and demonstrate that whereas the model-based approach proposed here preserves individual variability in systematic history biases across animals (while removing the influence of slow drifts), the previous model-free approach collapses this variability into a small subset of *apparent* strategies. Our results highlight the importance of evaluating assumptions underlying model-free corrections, and offer arguments for incorporating assumptions about the generative process directly into model fitting.

Results

Correction in action

We begin by reproducing the results from Lak, Hueske, et al. 2020 that the proposed model-free (MF) correction removes spurious choice correlations that are introduced by slow drifts in decision criterion. We consider a generative model in which the stimulus affects the choice of an agent in accordance with signal detection theory (see Methods, Figure 1A). On any given trial t, the agent compares a noisy perceptual sample, centered around the true stimulus s_t with logistic noise (scale β_S) to a decision criterion b_t, and makes a rightward choice if the sample exceeds the criterion and leftward choice otherwise. Therefore, probability of making a rightward choice is given by the following logistic function:

$$p(c_t = 1|s_t, b_t) = \frac{1}{1 + e^{-(s_t - b_t)/\beta_S}}$$

The decision criterion b_t of this observer varies from trial-to-trial relative to a fixed baseline β_0 due to unsystematic variations arising from random drift z_t:

$$b_t = \beta_0 + z_t$$

This drift z in decision criterion b (Figure 1B) was simulated according to the following autoregressive process (discrete time Ornstein-Uhlenbeck process):

$$z_t = (1 - \lambda)z_{t-1} + \sigma_d \epsilon_t$$

where ϵ_t is an i.i.d sequence of standard Gaussian random variables and σ_d sets the standard deviation of the Gaussian noise. The decay rate λ was fixed to a small value ($\lambda = 5e - 4$) to slightly dampen the pure Brownian motion thereby reducing the odds of the criterion drifting too far from the true boundary.

We assume that there is no learning or systematic trial-by-trial updating of any decision-related variables. In spite of that, the slowly drifting criterion produces correlations in the agent's choices across trials, such that when the psychometric curve is conditioned on previous trial's outcome and choice, it appears that the agent is following a small win-stay larger lose-stay strategy (Figure 1C)-- in other words, it appears that subjects use the experience of a trial's outcome to affect future trials, when in fact reward outcomes do not affect the decision process at all. The psychometric curves following both correct and incorrect rightward choices are biased towards rightward choices and vice versa following leftward choices. We summarize these apparent history effects by taking the difference between conditioned psychometric curves following right and left wins/losses. Therefore, following correct trials the history bias is given according to:

$$\text{Bias}_{t|t-1}^{PC} = \frac{1}{2}\big(p(c_t = 1|s_t, o_{t-1} = 1, c_{t-1} = 1) - p(c_t = 1|s_t, o_{t-1} = 1, c_{t-1} = 0)\big)$$

where $o_{t-1} = 1$ if the trial $t-1$ is a win and is 0 otherwise (PC refers to post-correct). And to compute the bias following error trials (PE) we instead condition on trial $t-1$ being an incorrect trial or $o_{t-1} = 0$:

$$\text{Bias}_{t|t-1}^{PE} = \frac{1}{2}\big(p(c_t = 1|s_t, o_{t-1} = 0, c_{t-1} = 1) - p(c_t = 1|s_t, o_{t-1} = 0, c_{t-1} = 0)\big)$$

This metric is positive if the behavior has a stay bias and is negative when there is a tendency to switch. Even though the drift in criterion is independent of the previous trial's outcome, the stay bias appears previous outcome-dependent - its magnitude is smaller following correct choices than erroneous choices (Figure 1D, solid lines). This effect arises because the agent is more prone to errors and more likely to have a persistent choice bias when the criterion has drifted further away from the true boundary.

Further, we examine how recent sensory history affects future choices under such a generative model and observe that again, the slow drifts in criterion produce the semblance of systematic trial-by-trial updating (Figure 1E, F; replication of Figure 2, figure supplement 1 from Lak et al., 2020). It appears as though the agent's choice on the current trial is modulated by the previous trial's stimulus difficulty (Figure 1E). Following correct trials the agent has a higher propensity to repeat the correct response to the previous trial's category if the previous trial was difficult. In contrast if the past trial was an error the bias is higher following an easy trial (not shown). We summarize these apparent stimulus history biases following rewarded trials for each pair of previous and current stimuli (Figure 1F, left panel). The bias is computed by subtracting the psychometric curve computed from all trials from the psychometric curves conditioned on the previous trial's stimulus strength (similar to Lak et al., 2020):

$$\text{Bias}_{s_t|s_{t-1}} = \frac{1}{2}\big(p(c_t = 1|s_t, s_{t-1}, o_{t-1} = 1) - p(c_t = 1|s_t)\big)$$

Next, we apply the MF correction to this synthetic data. The MF correction exploits the idea that the effect of slow drifts should be similar for at least a small stretch of trials (n>=2) and any influence of future rewards on past choices is acausal, likely due to slow drifts and needs to be removed. (Mendonça et al. 2020) proposed the variant of the correction for choice-outcome biases and (Lak, Hueske, et al. 2020) for stimulus history biases. We describe and test the efficacy of both these variants here. For post-correct (PC) history bias, the MF correction entails subtracting the (spurious) bias due to trial $t-1$ on the past trial $t-2$ from the bias due to trial $t-1$ on the future trial t:

$$\text{Corrected Bias}^{PC}_{t|t-1} = \text{Bias}^{PC}_{t|t-1} - \text{Bias}^{PC}_{t-2|t-1}$$

Post-error history bias is corrected for slow drifts in a similar manner. The MF correction for stimulus history effects stipulates that the effect of slow drifts in decision criterion can be removed by performing an equivalent operation on inferred stimulus history bias:

$$\text{Corrected Bias}_{s_t|s_{t-1}} = \text{Bias}_{s_t|s_{t-1}} - \text{Bias}_{s_{t-2}|s_{t-1}}$$

When applied to synthetic data, the MF correction removes these apparent choice history effects caused by the drifting criterion and successfully recovers the systematic component of past choice's influence on subsequent choice i.e. no influence (Figure 1D, dashed lines). Similarly, specious stimulus history effects could be successfully removed by applying the MF correction (Figure 1F, right panel).
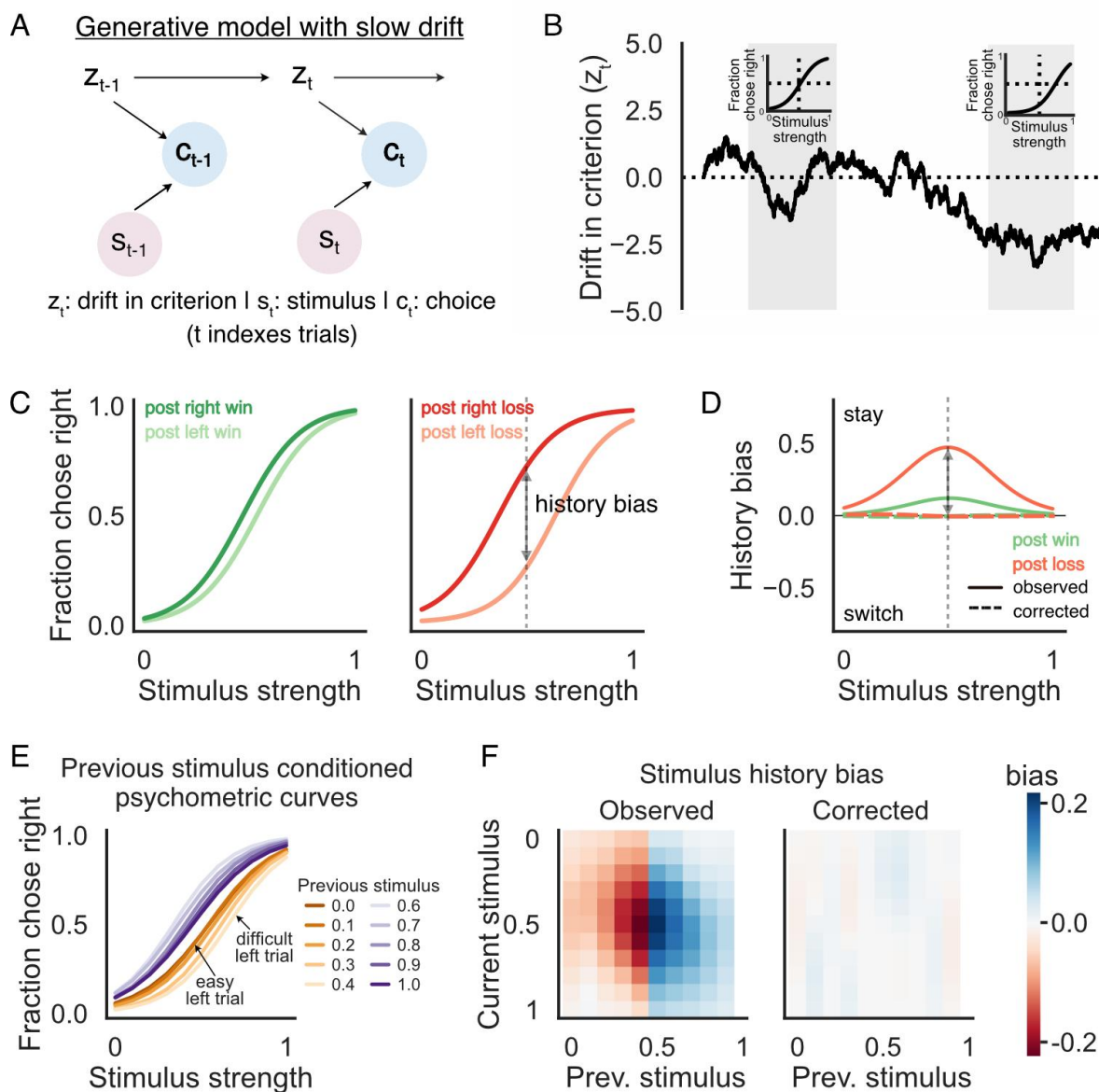
Figure 1: Correction for removing influences of slow drift in psychometric parameters on choice data (as proposed in Lak '20)

A) Generative model of choices based on signal detection theory: Each trial's choice **c** is made by comparing a noisy sample of that trial's stimulus **s** to a slowly drifting criterion (drift in the criterion represented by **z**, trials indexed by **t**)

B) Trajectory of the slow drift in criterion over a period of 5000 trials revealing extended periods of consistent choice bias. Insets: psychometric curves for the trials highlighted in grey when choices are on average unbiased (left) and biased (right)

C) Psychometric curves (fraction of rightward choices as a function of the current trial's stimulus strength) conditioned on the previous trial's choice and outcome, displaying an apparent win-stay lose-stay bias. Colors indicate previous outcome (green - win, red - loss) and lightness indicates previous choice (dark - right, light - left)

D) History bias, measured as the difference between conditioned psychometric curves following right and left wins (green) or losses (red). Solid lines indicate observed history bias, displaying an apparent small win-stay, larger lose-stay effect. Dashed lines indicate history bias after applying the MF correction, which removes the spurious choice correlation created by slow drift.

E) Psychometric curves conditioned on the previous trial's stimulus when the previous trial was rewarded. Curves display an apparent dependence on past stimulus strength, and have bigger shifts following hard trials compared to easy trials. Colors indicate previous stimulus category (violet - rightward, amber - leftward) and lightness indicates strength of the stimulus (dark - high stimulus strength or easy trials, light - low stimulus strength or hard trials)

F) (Left) Stimulus history bias, measured as the difference between psychometric curves conditioned on previous stimulus and average psychometric curve. Biases are larger following hard trials (color gradient along x-axis) and disproportionately affect the performance on hard trials (color gradient along y-axis). Color indicates the direction of induced bias (blue - towards right, red - towards left) (Right) Stimulus history biases after applying the MF correction, which removes the spurious dependence due to slow drifts in decision criterion.

Correction in misaction

In the past section, the generative model lacked active trial-by-trial updating and hence the correlations in choice were entirely due to the slow drift in decision criterion. Next we consider a more complex scenario, one in which the decision variables are systematically updated from trial-to-trial due to ongoing learning or other heuristic strategies in addition to the random drifts in decision criterion, and we determine how the MF correction performs under these circumstances.

Previous studies have shown that a large subset of possible learning strategies can be approximated by logistic regression models that directly represent the influence of reward and choice history on future choices (Katahira 2015; Miller et al. 2019). For this reason, in this study we explore the space of possible choice and outcome history effects instead of examining the correction's effect in the presence of individual learning algorithms. We parametrically vary the bias induced by past choice (c) and outcome (o) and study the effect of MF correction both in the presence and absence of slow drifts (z) of decision criterion (Figure 2A). We find that while the MF correction can correctly recover a symmetric win-stay lose-switch history dependence, it fails to recover the generative history-dependence in other scenarios.

In order to simulate choices from this family of generative models, we update the randomly drifting decision criterion b_t every trial based on the recently observed choice and outcome from the previous trial. Therefore, on any trial the decision criterion is given by:

$$b_t = z_t + b_0 + \beta_C(\mathbb{I}_{t-1}^{RC} - \mathbb{I}_{t-1}^{LC}) + \beta_E(\mathbb{I}_{t-1}^{RE} - \mathbb{I}_{t-1}^{LE})$$

where $\mathbb{I}_{t-1}^{RC}$, $\mathbb{I}_{t-1}^{LC}$, $\mathbb{I}_{t-1}^{RE}$ and $\mathbb{I}_{t-1}^{LE}$ are indicator variables denoting successes or failures on the previous trial ($t-1$) following rightward and leftward choices (RC refers to right correct, LC left correct, RE right error and LE left error). β_C and β_E determine how much previous trial's outcome modulates the criterion (β_C following correct trials and β_E following error trials). As before, β_0 is the fixed baseline and z_t denotes unsystematic variations arising from random drift.

First, we simulate data with a symmetric win-stay lose-switch rule, such that positive outcomes or correct trials bias the agent towards repeating the choice that led to them, and negative outcomes do the opposite i.e. $\beta_C = 1$, $\beta_E = -1$ and $z_t = 0$ $\forall t$ (Figure 2B, C, first column, solid lines). When, in addition to win-stay lose-switch, the decision criterion slowly drifts over trials ($z_t \neq 0$), the systematic lose-switch rule is obscured and the choice behavior seems to have no updating following error trials (Figure 2B, C, first column, dotted lines). Moreover, the slope of the psychometric curve (also referred to as perceptual sensitivity) appears shallower.

We apply the MF correction to this data and examine the observed and corrected history biases with and without an underlying slow random drift in the decision criterion of the generative process. In the absence of slow drifts, as expected, the MF correction (dotted lines) does not change our estimate of history biases and we recover the generative win-stay lose-switch bias (Figure 2D, first column). But even in the presence of slow drifts, the MF correction successfully nullifies its contaminating influence on history bias and we approximately recover the generative win-stay lose-switch bias (Figure 2E, first column).

Next, we consider a scenario in which the magnitude of biases following wins (i.e.. rewarded trials) versus losses (unrewarded trials) are asymmetric. We eliminate the bias induced by negative outcomes and study win-stay lose-nothing ($\beta_C = 1, \beta_E = 0$; Figure 2B,C second column). Here the slow drift masks the true post-error biases and instead accentuates stay biases. In the absence of slow drift, we can directly observe the generative bias, therefore in principle the MF correction should not affect our estimate of history biases. But instead we observe that the MF correction distorts our estimate of history biases, skewing them towards win-stay lose-switch like updating (Figure 2D second column). In the presence of slow drift too, the MF correction is unable to recover the generative systematic history bias (Figure 2E, second column solid vs dashed) and instead inaccurately estimates a stay bias following wins and a switch bias following losses.

When there is no dependence on previous outcomes, and the bias following both correct and error trials are towards the previous choice i.e. a win-stay lose-stay bias ($\beta_C = 1, \beta_E = 1$; Figure 2B, C third column), slow drifts in decision criterion exacerbate the stay effects following error trials. In synthetic data generated without any drifts in decision criterion but with stay bias in effect, the MF correction spuriously infers a win-stay lose-switch strategy (Figure 2D, third column). A similar failure in recovering the generative parameters is observed in the presence of slow drifts (Figure 2E third column).

Next we examine the effect of the MF correction when the data is generated from an outcome independent switch rule i.e. win-switch lose-switch ($\beta_C = -1, \beta_E = -1$; Figure 2B,C last column). In this case too, slow drift obscures the true post-error effects and overlays a stay bias on it. Again, we find that the MF

8

correction fails to recover generative biases both in the presence and absence of slow drifts - it instead returns a win-switch lose-stay like bias (Figure 2D,E last column).


**A** Generative model with slow drift and history-dependent updating

z_t: drift in criterion
s_t: stimulus
c_t: choice
o_t: outcome
(t indexes trials)
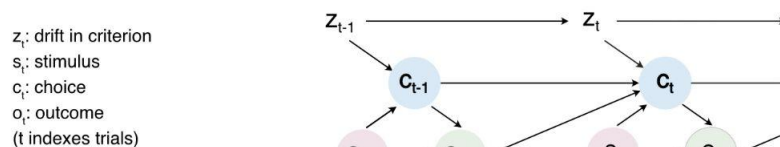
| WIN-STAY LOSE-SWITCH | WIN-STAY LOSE-NOTHING | WIN-STAY LOSE-STAY | WIN-SWITCH LOSE-SWITCH |

— Choices simulated **without** slow drifts in criterion   ⋯⋯ Choices simulated **with** slow drifts in criterion
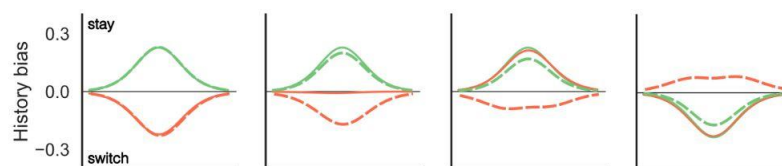
B Post win psychometrics   post right win   post left win
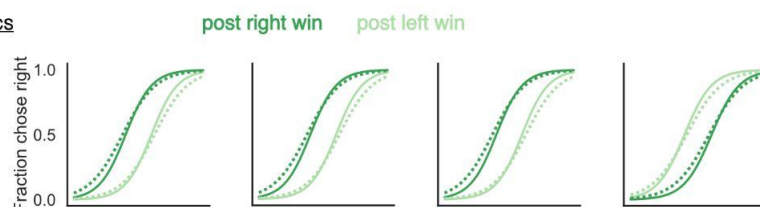
C Post loss psychometrics   post right loss   post left loss
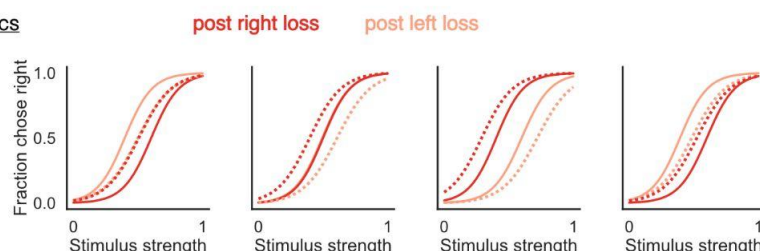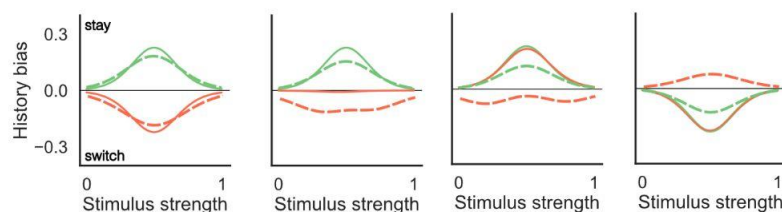
— Generative systematic history bias   - - - Predicted bias from correction

post win   post loss

D Choices simulated **without** slow drifts in criterion

E Choices simulated **with** slow drifts in criterion

Figure 2: Correction fails to recover generative parameters when decision variables are actively updated from trial-to-trial

A) Generative model in which the current trial's choice c_t is influenced by the previous choice c_{t-1} and previous outcome o_{t-1}, in addition to the current trial's stimulus s_t and a slow drift in criterion z_t. In some scenarios, the previous outcome does not influence the choice on the next trial e.g. win-stay lose-stay or win-switch lose-switch.

B) Psychometric curves post rewarded trials, conditioned on previous trial's choice for a model with win-stay, lose-switch (column 1), win-stay lose-nothing (column 2), win-stay lose-stay (column 3) and win-switch lose-switch (column 4) biases. Solid lines represent psychometric curves without slow drift and dotted lines with slow drift. Addition of slow drift gives rise to worse perceptual sensitivity.

C) Same as B but for trials following losses or unrewarded trials. Addition of slow drift masks lose-switch effects and overlays a stay bias across all updating strategies.

D) History biases following wins (green) and losses (red) *without* slow drifts in criterion. Solid: generative systematic history biases, dashed: estimates from MF correction. MF correction nearly recovers the true generative post win history biases (dashed vs solid green line) across different updating strategies. For post loss biases (dashed vs solid red line), the MF correction has no effect for win-stay lose-switch bias (column 1), but produces spurious lose-switch effects for win-stay lose-nothing (column 2) and win-stay lose-stay (column 3) biases. For win-switch lose-switch (column 4) bias the correction returns a lose-stay bias.

E) Same as D but for history biases following wins (green) and losses (red) *with* slow drifts in criterion. The correction fails to recover the generative history biases (dashed vs solid lines) in this case as well.

A model-based solution

The foregoing analysis demonstrates that the proposed MF correction produces inaccurate results when the systematic trial-to-trial updating of history bias deviates from a symmetric win-stay lose-switch strategy. The correction assumes that correlations in choice arise from processes unrelated to deterministic trial-by-trial updating - an assumption that is untrue for many learning or heuristic algorithms. This could be remedied by explicitly estimating the contributions of both slow drifts and systematic updates to choice behavior.

Here we describe one such approach for the simulations discussed in the previous section. We infer the parameters governing behavior by fitting the choices to the generative model at hand: a signal detection theory observer with logistic noise, systematic trial-history dependent biases and drifts in criterion (Figure 3A, see Methods). We fit the parameters $\Theta = \{\sigma_d, b_0, \beta_S, \beta_C, \beta_E\}$ of the model to choices using the Expectation Maximization algorithm with a Laplace approximation of the posterior over the latent state (drifting variable; Macke 2011, Macke et al. 2015) as implemented in the *ssm library*.

We simulate data from an agent that follows a win-stay lose-stay strategy and has random drifts in its decision criterion (similar to Figure 2, third column). The MF correction is ineffective in recovering the generative parameters in this case. The model successfully captures trends in the data (Figure 3B: Fits (dashed-dotted lines) plotted alongside observed choice and outcome conditioned psychometric curves from the simulated data (solid lines)). We next examine if the model is able to tease apart the relative influences of slow drift and trial history bias on choice behavior and indeed, the inferred most likely trajectory

of the latent drifting variable (Figure 3C, grey line) closely tracks the true simulated drift (Figure 3C, black line). Furthermore, the model correctly infers the true parameters governing the systematic component of the behavior. We demonstrate this by simulating choices in the absence of slow drifts with both the generative parameters and estimated parameters. The previous choice and outcome conditioned psychometrics from the generative and estimated parameters show good correspondence (Figure 3D). Additionally, this method helps recover the true perceptual sensitivity, which is otherwise confounded by the drifting biases (as in Figure 2B,C).
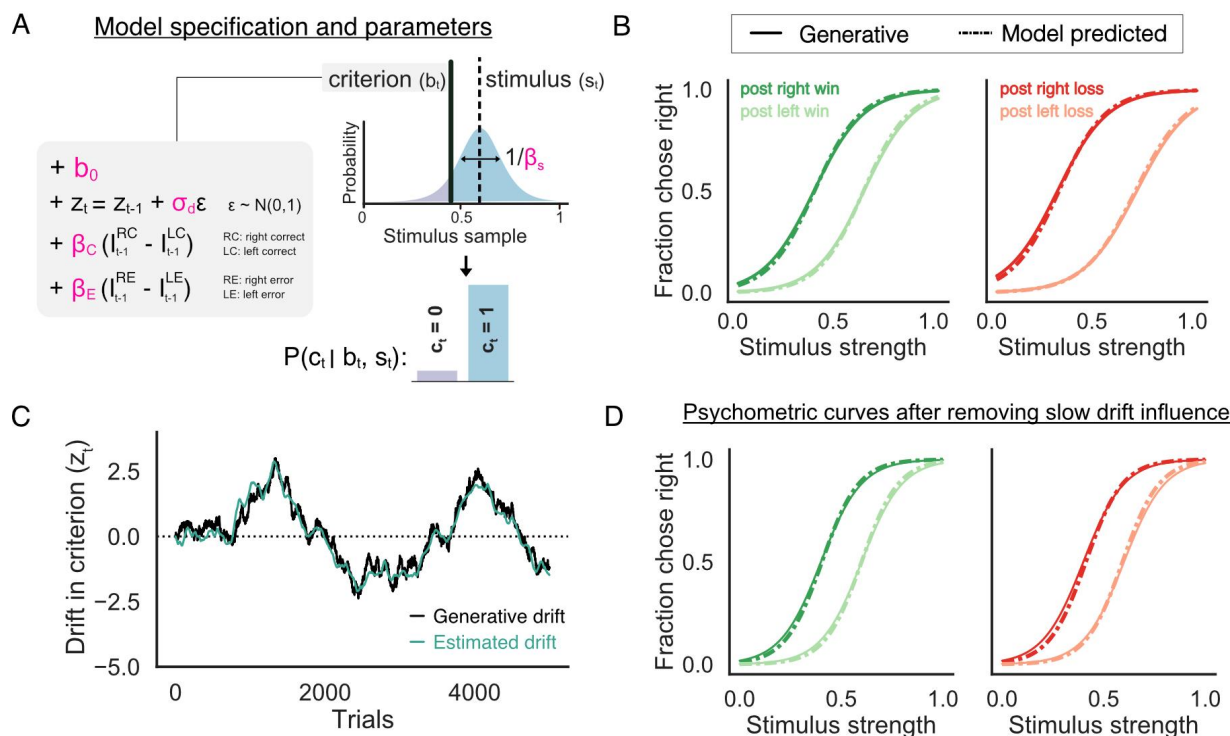

Figure 3: Fitting the model helps disambiguate systematic history biases from effects of slow drift

A) Schematic of the model with parameters highlighted in pink. (Top right) Logistic distribution of noisy samples perceived by a subject for a given true stimulus s_t (dotted line). The probability of the current trial's choice c_t being rightward/leftward (shaded blue, bottom right) is given by the mass to the right/left of the criterion b_t (solid black line), which evolves according to the equations shown in the box (left - z_t is drift in criterion, $\mathbb{I}_{t-1}$ is an indicator variable for the previous trial's choice, outcome).

B) Psychometric curves from simulated data (solid lines) shown alongside fits (dashed-dotted lines) to a model containing both history biases and slow drifts in criterion

C) Inferred slow drifts from model fits (grey) compared to the true generative slow drift (black) showing good correspondence

D) Inferred history biases from model fits (dashed-dotted lines) compared to the true history biases (solid lines) showing good correspondence. Note: these are hypothetical psychometric curves that would have been observed in the absence of slow drift, and hence not directly observable from the data as in B)

Comparing the model-free correction and the model-based approach

Next, we compare the history biases estimated by the two approaches with synthetic as well as real data. We simulate choices for a variety of possible post-correct and post-error biases (Figure 4A-C, squares) in addition to random drifts in decision criteria and evaluate the results obtained by applying the MF correction and by model-fitting. We summarize these history effects by plotting the choice bias observed post-correct trials against post-error trials at the point of indifference. First, we show that consistent with our previous simulations - presence of slow drifts exaggerates stay biases in the data (Figure 4A, dots) especially following error trials, and obscures the true underlying update strategy (Figure 4A, squares). Application of the MF correction does not selectively remove the influence of slow drifts, rather it warps all the history effects to lie along the diagonal represented by win-switch lose-stay and win-stay lose-switch strategies (Figure 4B, dots). In contrast, the model-based inference of systematic updates successfully removes the influence of slow drifts and reveals the underlying strategy for all considered settings (Figure 4C, dots vs squares).

Next, we analyze choice data from 19 rats participating in a two-alternative auditory evidence-accumulation task (data from Brunton et al, 2013; see Methods for more details). In this study, even though successive trials were independent of each other, the rat subjects showed varying extents of win-stay and lose-switch biases (Figure 4D). The application of MF correction to this dataset, squashes the variability in history biases across all rats and returns a homogenous win-stay lose-switch bias (Figure 4E). Whereas, the model fits retain the individual variability while alleviating the stay effects in post-error bias that are caused by slow drifts (Figure 4F).
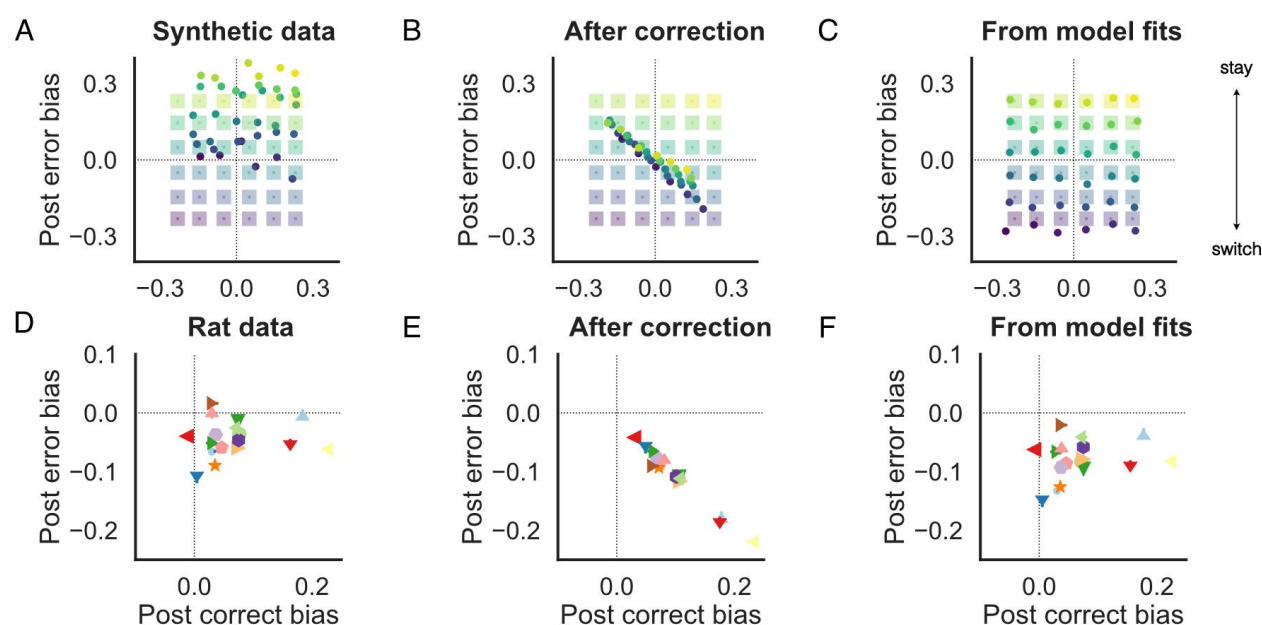
Figure 4:

A) Biases measured by the difference between conditioned psychometric curves following left and right wins (x-axis) or losses (y-axis) at the point of indifference (stimulus strength at which probability of choosing either option is equal). Squares denote the biases caused by systematic trial by trial updates, and dots denote the observed biases when slow drifts are present. Colors represent unique parameter settings.

B) Distribution of biases after applying MF correction

C) Distribution of inferred history biases from fits to a model that includes slow drift and trial-by-trial updating.

D) Distribution of observed history biases from 19 rats (Brunton 2013 dataset). symbols denote individual rats.

E) Distribution of biases after applying correction from Lak et al. '20. MF correction distorts the distribution, uniformly producing equal degrees of win-stay and lose-switch biases.

F) Distribution of inferred history biases from fits to a model that includes slow drift and trial-by-trial updating. The distribution is shifted relative to A), but individual rats retain their relative position and asymmetries between post-correct and post-error updating.

Discussion

Nonstationarities in decision-relevant variables, if overlooked, can bias one's estimates of psychophysical parameters and obscure strategies that underlie behavior (Fründ, Haenel, and Wichmann 2011). Previous work (Lak et al 2020) has shown that unaccounted for slow drifts in decision criterion could obscure trial-by-trial updates produced by active learning, and proposed a correction to remove its influence (Figure 1). Here we investigate the performance of this correction under a range of different generative models, and demonstrate that it fails to selectively remove the influence of slow drifts in the presence of systematic trial-by-trial updates of decision variables (Figure 2). Moreover, applying the correction corrupts the estimates of trial-history influences, biasing them towards a small subset of possible strategies. To mitigate these shortcomings, we propose an alternate approach of explicitly modeling slow fluctuations in the decision-making process (Figure 3). We demonstrate that this model-based approach can successfully disambiguate systematic updates from non-specific drifts, hence preserving the structure of individual variability in behavioral strategies (Figure 4).

Non-specific drifts in decision variables may arise from fluctuations in internal states such as attention, arousal and motivation, or from other internal sources of noise (Renart and Machens 2014). Hence, some authors have used neural measurements to gauge their dynamics (Ridderinkhof, Nieuwenhuis, and Bashore 2003; Eichele et al. 2008; Cowley et al. 2020). Of particular interest is the work of Cowley and colleagues who leverage slow drifts in stimulus encoding in visual and prefrontal cortex to explain fluctuations in behavior and pinpoint the affected decision-variable.

Even in the absence of such detailed measurements, it may be possible to account for unsystematic influences by modelling the noisy dynamics of unobserved variables, as we have done in this manuscript for

the decision criterion (see Smith et al. 2004 for formative work). This approach has also been used to identify systematic and random contributions to action value learning (Findling et al. 2019), allowing for the decomposition of errors into noise-driven and choice-driven components.

In general, it is likely that multiple decision variables drift over time, or that systematic updates occur over longer timescales, in which case it is important to incorporate those assumptions into the model. While the model we suggest can be readily extended to estimate additional forms of dynamics, other authors have put forward alternate approaches. Previous studies have proposed regression models to infer the trajectory of psychometric parameters over trials, regularizing their estimates with empirically determined priors (Bak et al. 2016; Kattner, Cochrane, and Green 2017; Bak and Pillow 2018; Roy et al. 2018). In another study authors consider a generative model in which the latent state which governs the setting of psychometric parameters undergoes discrete changes (Ashwood et al. 2020). In parallel, some generative models might allow for the use of filtering methods to recover signal from the observations. Indeed, previous studies have shown that low-pass filtering the sequence of choices using a moving average filter can adequately help estimate the slow drifts in bias (Cowley et al. 2020; Mochol, Kiani, and Moreno-Bote 2020) without inducing the kind of biases induced by the MF correction. The relative utility of these different approaches (state space models, regression, filtering, corrections) would depend on the desired level of explanatory power, number of data samples and efficiency of the inference algorithms among other factors. In any case, studies would benefit from explicit comparison between different hypothesized generative models.

Outside the realm of trial-based behavior many exciting advances have been made in inferring the dynamics underlying naturalistic behaviour (Wiltschko et al. 2015; Eyjolfsdottir et al. 2016; Sharma et al. 2018; Calhoun, Pillow, and Murthy 2019). In the future, it would be interesting to bring these advances into the study of learnt perceptual behaviour and develop inference strategies for more sophisticated generative models. Exploring the origins and logic of slow fluctuations inferred with such models might help advance our understanding of the principles that underlie behaviour and learning.

Methods

Simulation Details

We simulated a signal detection theory observer that compares a noisy sample of the stimulus (range 0 to 1, corrupted by logistic noise) to a decision criterion, making a rightward choice if the sample exceeds the criterion and leftward choice otherwise.

The decision criterion b_t of this observer varies from trial to trial relative to a fixed baseline β_0, due to unsystematic variations arising from random drift z_t and systematic variations based on the choice and

outcome of the previous trial c_{t-1}, o_{t-1}. The drift z in decision criterion b over trials t was simulated according to the following autoregressive process (discrete time Ornstein-Uhlenbeck process):

$$z_t = (1 - \lambda)z_{t-1} + \sigma_d \epsilon_t$$

where ϵ_t is an i.i.d sequence of standard Gaussian random variables. Throughout the study the decay rate was fixed to a small value ($\lambda = 5e - 4$) to prevent the criterion from drifting too far from the true boundary and the standard deviation of the Gaussian noise, σ_d was set at 0.05 when applicable. The systematic variations based on previous choice and outcome took the form of an additive bias to the decision criterion. The bias was $\beta_C/ - \beta_C$ following rightward/leftward choices and positive outcomes, and $\beta_E/ - \beta_E$ following rightward /leftward choices and negative outcomes.

Therefore, the decision criterion on any trial was given by the following equation:

$$b_t(c_{t-1}, o_{t-1}, z_t) = \beta_0 + z_t$$
$$+ \beta_C \mathbb{I}(c_{t-1} = 1, o_{t-1} = 1) - \beta_C \mathbb{I}(c_{t-1} = 0, o_{t-1} = 1)$$
$$+ \beta_E \mathbb{I}(c_{t-1} = 1, o_{t-1} = 0) - \beta_E \mathbb{I}(c_{t-1} = 0, o_{t-1} = 0)$$

where $\mathbb{I}(c_{t-1}, o_{t-1})$ denotes an indicator variable for a particular combination of previous choice c_{t-1} and outcome o_{t-1}. The baseline value β_0 was set to produce an equal proportion of leftward and rightward choices when stimulus carried no information i.e. stimulus strength of 0.5.

Finally, choice c on trial t depended upon the stimulus strength on the current trial s_t, and the decision criterion b_t. The probability of making a rightward choice was given by the logistic function:

$$p(c_t = 1 | s_t, b_t) = \frac{1}{1 + e^{-(s_t - b_t)/\beta_S}}$$

We simulated 40000 trials at a time, and the plotted psychometric functions were fit using the logistic regression function from the scikit-learn module in Python (Pedregosa et al. 2011).

History Bias and Correction

History bias post-correct/error at each stimulus strength is computed as the difference between the probability of going right following rightward and leftward correct/error trials. Therefore, the history bias following correct trials is given by

$$bias_{t|t-1}^{PC} = \frac{1}{2}\left(p(c_t = 1 | s_t, o_{t-1} = 1, c_{t-1} = 1) - p(c_t = 1 | s_t, o_{t-1} = 1, c_{t-1} = 0)\right)$$. And similarly, following error trials history bias is computed as

$$\frac{1}{2}\left(p(c_t = 1 | s_t, o_{t-1} = 0, c_{t-1} = 1) - p(c_t = 1 | s_t, o_{t-1} = 0, c_{t-1} = 0)\right)$$.

To correct for the influence of slow drifts, we subtract the acausal bias due to the past trial on the choice preceding it from the bias due to the past trial on current choice $bias_{t-2|t-1}$ (Lak et al., 2020 and Mendonca et al., 2020) , i.e. following a rewarded trial the corrected bias for each stimulus strength was computed according to:

$$bias_{t|t-1}^{PC} - bias_{t-2|t-1}^{PC} = \frac{1}{2}\big(p(c_t = 1|s_t, o_{t-1} = 1, c_{t-1} = 1) - p(c_t = 1|s_t, o_{t-1} = 1, c_{t-1} = 0)\big) -$$
$$\frac{1}{2}\big(p(c_{t-2} = 1|s_{t=2}, o_{t-1} = 1, c_{t-1} = 1) - p(c_{t-2} = 1|s_{t-2}, o_{t-1} = 1, c_{t-1} = 0)\big)$$

The corrected bias following error trials was similarly computed, but by conditioning on previous error trials i.e. $o_{t-1} = 0$.

Model-based fitting

We denote the choice on trial $t \in 1...T$ as c_t such that $c_t = 1$ if the choice is towards right and 0 otherwise. The choice on any given trial is formed by comparing the value of a noisy perceptual sample (centered around the true stimulus s_t with logistic noise) to the category boundary or criterion b_t, which we assume is time-varying and is modulated by trial-history. If the value of the sample exceeds the criterion, a rightward choice is made. Therefore, conditioned on the drifting criterion and the stimulus the choice of an agent on any given trial is given by a Bernoulli distribution with mean

$$\mathbb{E}[c_t|b_t, s_t] = \frac{1}{1 + \exp(-(\beta_S s_t - b_t))}$$

where β_S determines the sensitivity of an agent to the perceptual stimulus s_t.

We assume that the drift z_t in criterion b_t evolves according to a random walk with step size σ_d, we assume that the mean and variance of the initial state are known $(z_0 = 0, \sigma_0^2 = 0.01)$:

$$z_1 \sim \mathcal{N}(z_0, \sigma_0^2)$$
$$z_{t+1}|z_t \sim \mathcal{N}(z_t, \sigma_d^2)$$

The decision criterion b_t is the summation of a fixed baseline b_0, the drift, and trial-history influences:

$$b_t = z_t + b_0 + \beta_C(\mathbb{I}_{t-1}^{RC} - \mathbb{I}_{t-1}^{LC}) + \beta_E(\mathbb{I}_{t-1}^{RE} - \mathbb{I}_{t-1}^{LE})$$

where $\mathbb{I}_{t-1}^{RC}$, $\mathbb{I}_{t-1}^{LC}$, $\mathbb{I}_{t-1}^{RE}$ and $\mathbb{I}_{t-1}^{LE}$ are indicator variables denoting successes or failures on the previous trial ($t - 1$) following rightward and leftward choices. β_C and β_E determine how much previous trial's choice and outcome modulate the criterion.

This model for choice behavior is essentially a linear dynamical system with Bernoulli emissions. Therefore, to infer the parameters $\Theta = \{\sigma_d, b_0, \beta_S, \beta_C, \beta_E\}$ of this latent variable model we use an

16

expectation-maximization algorithm, similar to those described before (Smith and Brown 2003, Macke et al. 2011, Macke et al. 2015). For the E-step, we require the log of posterior distribution $P(\mathbf{z}|\mathbf{c},\Theta)$ over the latent drift given the observed pattern of choices and our current estimate of the parameters Θ:

$$\log P(\mathbf{z}|\mathbf{c},\Theta) = \sum_{t=1}^{T}\left(c_t h_t + \log\left(1 - \frac{1}{1+e^{-h_t}}\right)\right) - \frac{(z_1 - z_0)^2}{2\sigma_0^2} - \frac{1}{2\sigma_d^2}\sum_{t=1}^{T-1}(z_{t+1} - z_t)^2 + \mathrm{const}$$

where $h_t = \beta_S s_t - b_0 - z_t - \beta_C \mathbb{I}_t^C - \beta_E \mathbb{I}_t^E$. Owing to Bernoulli observations this distribution is not available in closed-form but is concave, so we approximate it with a multivariate Gaussian distribution (Laplace approximation) $P(\mathbf{z}|\mathbf{c},\Theta) \sim \mathbb{N}(\mu^*,\Sigma^*)$ where $\mu*$ is a vector of size T and is set to maximize the posterior $\mu^* = \arg\max_{\mathbf{z}} P(\mathbf{z}|\mathbf{c},\Theta)$ and Σ^* is of size $T \times T$ and is set to the inverse Hessian of the log posterior evaluated at μ^* i.e. $\Sigma^* = -[\nabla_{\mathbf{z}}^2 \log P(\mathbf{z}|\mathbf{c},\Theta)|_{\mathbf{z}=\mu^*}]^{-1}$

The M-step updates parameters by optimizing the expected total data log-likelihood with respect to the parameters

$$\Theta^* = \arg\max_{\theta} \int [\log(\mathbf{c}|\mathbf{z},\Theta) + \log P(\mathbf{z}|\Theta)]\mathbb{N}(\mathbf{z}|\mu^*,\Sigma^*)d\mathbf{z}$$

We update the parameters using L-BFGS except σ_d which we update analytically. We use the *ssm* package (Linderman et al. 2020) in Python for this inference.

Dataset studied

We analyze the trial-history effects of nineteen rats (50223 ± 21915 trials per rat) performing an evidence accumulation task, originally published in Brunton et al. 2013. In this task, the subjects were presented with two simultaneous streams of randomly-timed discrete pulses of evidence (clicks), one from a speaker to their left and the other to their right, for a predetermined duration. The subjects were required to maintain fixation throughout the entire stimulus, failure to do so led to a violation trial. At the end of the stimulus, the subjects had to orient towards the side which played the greater number of clicks to obtain a water reward. The discrimination difficulty was controlled on each trial by varying the ratio of the generative Poisson rates of the two click streams. Successive trial's difficulty and rewarded side were independently sampled. Since the animals neither made a choice nor received an outcome on violation trials, we ignore them while computing trial-history effects.

Data and Code availability

The code and data associated with this manuscript is available here: https://doi.org/10.24433/CO.8821874.v1 as a standalone executable.

Acknowledgements

References

Abrahamyan, Arman, Laura Luz Silva, Steven C. Dakin, Matteo Carandini, and Justin L. Gardner. 2016. "Adaptable History Biases in Human Perceptual Decisions." *Proceedings of the National Academy of Sciences* 113 (25): E3548–57. https://doi.org/10.1073/pnas.1518786113.

Ashwood, Zoe C., Nicholas A. Roy, Iris R. Stone, The International Brain Laboratory, Anne K. Churchland, Alexandre Pouget, and Jonathan W. Pillow. 2020. "Mice Alternate between Discrete Strategies during Perceptual Decision-Making." *BioRxiv,* October, 2020.10.19.346353. https://doi.org/10.1101/2020.10.19.346353.

Bak, Ji Hyun, Jung Yoon Choi, Athena Akrami, Ilana Witten, and Jonathan W. Pillow. 2016. "Adaptive Optimal Training of Animal Behavior." *Advances in Neural Information Processing Systems* 29: 1947–55.

Bak, Ji Hyun, and Jonathan W. Pillow. 2018. "Adaptive Stimulus Selection for Multi-Alternative Psychometric Functions with Lapses." *Journal of Vision* 18 (12): 4–4. https://doi.org/10.1167/18.12.4.

Brunton, Bingni W, Matthew M Botvinick, and Carlos D Brody. 2013. "Rats and Humans Can Optimally Accumulate Evidence for Decision-Making." *Science* 340 (6128): 95–98.

Calhoun, Adam J., Jonathan W. Pillow, and Mala Murthy. 2019. "Unsupervised Identification of the Internal States That Shape Natural Behavior." *Nature Neuroscience* 22 (12): 2040–49. https://doi.org/10.1038/s41593-019-0533-x.

Cowley, Benjamin R., Adam C. Snyder, Katerina Acar, Ryan C. Williamson, Byron M. Yu, and Matthew A. Smith. 2020. "Slow Drift of Neural Activity as a Signature of Impulsivity in Macaque Visual and Prefrontal Cortex." *Neuron* 108 (3): 551-567.e8. https://doi.org/10.1016/j.neuron.2020.07.021.

Dayan, P., and N. D. Daw. 2008. "Decision Theory, Reinforcement Learning, and the Brain." *Cognitive, Affective, & Behavioral Neuroscience* 8 (4): 429–53. https://doi.org/10.3758/CABN.8.4.429.

Drugowitsch, Jan, André G. Mendonça, Zachary F. Mainen, and Alexandre Pouget. 2019. "Learning Optimal Decisions with Confidence." *Proceedings of the National Academy of Sciences* 116 (49): 24872–80. https://doi.org/10.1073/pnas.1906787116.

Dutilh, Gilles, Don van Ravenzwaaij, Sander Nieuwenhuis, Han L. J. van der Maas, Birte U. Forstmann, and Eric-Jan Wagenmakers. 2012. "How to Measure Post-Error Slowing: A Confound and a Simple Solution." *Journal of Mathematical Psychology* 56 (3): 208–16. https://doi.org/10.1016/j.jmp.2012.04.001.

Eichele, Tom, Stefan Debener, Vince D. Calhoun, Karsten Specht, Andreas K. Engel, Kenneth Hugdahl, D. Yves von Cramon, and Markus Ullsperger. 2008. "Prediction of Human Errors by Maladaptive Changes in Event-Related Brain Networks." *Proceedings of the National Academy of Sciences* 105 (16): 6173–78. https://doi.org/10.1073/pnas.0708965105.

Eyjolfsdottir, Eyrun, Kristin Branson, Yisong Yue, and Pietro Perona. 2016. "Learning Recurrent Representations for Hierarchical Behavior Modeling." *ArXiv:1611.00094 [Cs]*, November. http://arxiv.org/abs/1611.00094.

Findling, Charles, Vasilisa Skvortsova, Rémi Dromnelle, Stefano Palminteri, and Valentin Wyart. 2019. "Computational Noise in Reward-Guided Learning Drives Behavioral Variability in Volatile Environments." *Nature Neuroscience* 22 (12): 2066–77. https://doi.org/10.1038/s41593-019-0518-9.

Fründ, Ingo, N. Valentin Haenel, and Felix A. Wichmann. 2011. "Inference for Psychometric Functions in the Presence of Nonstationary Behavior." *Journal of Vision* 11 (6): 16–16. https://doi.org/10.1167/11.6.16.

Gardner, Justin L. 2019. "Optimality and Heuristics in Perceptual Neuroscience." *Nature Neuroscience* 22 (4): 514–23. https://doi.org/10.1038/s41593-019-0340-4.

Gigerenzer, Gerd, and Wolfgang Gaissmaier. 2011. "Heuristic Decision Making." *Annual Review of*

Psychology 62 (1): 451–82. https://doi.org/10.1146/annurev-psych-120709-145346.

Hennig, Jay A., Emily R. Oby, Matthew D. Golub, Lindsay A. Bahureksa, Patrick T. Sadtler, Kristin M. Quick, Stephen I. Ryu, et al. 2021. "Learning Is Shaped by Abrupt Changes in Neural Engagement." *Nature Neuroscience* 24 (5): 727–36. https://doi.org/10.1038/s41593-021-00822-8.

Katahira, Kentaro. 2015. "The Relation between Reinforcement Learning Parameters and the Influence of Reinforcement History on Choice Behavior." *Journal of Mathematical Psychology* 66 (June): 59–69. https://doi.org/10.1016/j.jmp.2015.03.006.

Kattner, Florian, Aaron Cochrane, and C. Shawn Green. 2017. "Trial-Dependent Psychometric Functions Accounting for Perceptual Learning in 2-AFC Discrimination Tasks." *Journal of Vision* 17 (11): 3–3. https://doi.org/10.1167/17.11.3.

Lak, Armin, Emily Hueske, Junya Hirokawa, Paul Masset, Torben Ott, Anne E Urai, Tobias H Donner, et al. 2020. "Reinforcement Biases Subsequent Perceptual Decisions When Confidence Is Low, a Widespread Behavioral Phenomenon." Edited by Emilio Salinas, Michael J Frank, Emilio Salinas, Carlos D Brody, and Long Ding. *ELife* 9 (April): e49834. https://doi.org/10.7554/eLife.49834.

Lak, Armin, Michael Okun, Morgane M. Moss, Harsha Gurnani, Karolina Farrell, Miles J. Wells, Charu Bai Reddy, Adam Kepecs, Kenneth D. Harris, and Matteo Carandini. 2020. "Dopaminergic and Prefrontal Basis of Learning from Sensory Confidence and Reward Value." *Neuron* 105 (4): 700-711.e6. https://doi.org/10.1016/j.neuron.2019.11.018.

Linderman, Scott, Benjamin Antin, David Zoltowski, and Joshua Glaser. (2018) 2020. *SSM: Bayesian Learning and Inference for State Space Models* (version 0.0.1). Jupyter Notebook. https://github.com/lindermanlab/ssm.

Lyamzin, Dmitry R., Ryo Aoki, Mohammad Abdolrahmani, and Andrea Benucci. 2021. "Probabilistic Discrimination of Relative Stimulus Features in Mice." *BioRxiv*, February, 2020.12.20.423700. https://doi.org/10.1101/2020.12.20.423700.

MacKe, J. H., L. Buesing, and M. Sahani. 2015. "Estimating State and Parameters in State Space Models of Spike Trains." In *Advanced State Space Methods for Neural and Clinical Data*, edited by Zhe Chen, 137–59. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139941433.007.

Macke, Jakob H. 2011. "Empirical Models of Spiking in Neural Populations," 9.

Macke, Jakob H., Lars Buesing, John P. Cunningham, Byron M. Yu, Krishna V. Shenoy, and Maneesh Sahani. 2011. "Empirical Models of Spiking in Neural Populations." *Advances in Neural Information Processing Systems* 24: 1350–58.

Mendonça, André G., Jan Drugowitsch, M. Inês Vicente, Eric E. J. DeWitt, Alexandre Pouget, and Zachary F. Mainen. 2020. "The Impact of Learning on Perceptual Decisions and Its Implication for Speed-Accuracy Tradeoffs." *Nature Communications* 11 (1): 2757. https://doi.org/10.1038/s41467-020-16196-7.

Miller, Kevin J., Matthew M. Botvinick, and Carlos D. Brody. 2019. "From Predictive Models to Cognitive Models: An Analysis of Rat Behavior in the Two-Armed Bandit Task." *BioRxiv*, February, 461129. https://doi.org/10.1101/461129.

Mochol, Gabriela, Roozbeh Kiani, and Rubén Moreno-Bote. 2020. "Prefrontal Cortex Represents Heuristics That Shape Choice Bias and Its Integration into Future Behavior." Preprint. Animal Behavior and Cognition. https://doi.org/10.1101/2020.03.20.000224.

Pedregosa, Fabian, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, et al. 2011. "Scikit-Learn: Machine Learning in Python." *Journal of Machine Learning Research* 12 (85): 2825–30.

Pisupati, Sashank, Lital Chartarifsky-Lynn, Anup Khanal, and Anne K Churchland. 2021. "Lapses in Perceptual Decisions Reflect Exploration." Edited by Daeyeol Lee. *ELife* 10 (January): e55490. https://doi.org/10.7554/eLife.55490.

Renart, Alfonso, and Christian K Machens. 2014. "Variability in Neural Activity and Behavior." *Current Opinion in Neurobiology*, Theoretical and computational neuroscience, 25 (April): 211–20. https://doi.org/10.1016/j.conb.2014.02.013.

Ridderinkhof, K. Richard, Sander Nieuwenhuis, and Theodore R. Bashore. 2003. "Errors Are Foreshadowed in Brain Potentials Associated with Action Monitoring in Cingulate Cortex in Humans." *Neuroscience Letters* 348 (1): 1–4. https://doi.org/10.1016/s0304-3940(03)00566-4.

Roy, Nicholas A., Ji Hyun Bak, Athena Akrami, Carlos D. Brody, and Jonathan W. Pillow. 2021. "Extracting

the Dynamics of Behavior in Sensory Decision-Making Experiments." *Neuron* 109 (4): 597-610.e6. https://doi.org/10.1016/j.neuron.2020.12.004.

Roy, Nicholas A., Ji Hyun Bak, Athena Akrami, Carlos Brody, and Jonathan W Pillow. 2018. "Efficient Inference for Time-Varying Behavior during Learning." In *Advances in Neural Information Processing Systems 31*, edited by S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, 5695–5705. Curran Associates, Inc. http://papers.nips.cc/paper/7812-efficient-inference-for-time-varying-behavior-during-learning.pdf.

Sharma, Anuj, Robert Johnson, Florian Engert, and Scott Linderman. 2018. "Point Process Latent Variable Models of Larval Zebrafish Behavior." *Advances in Neural Information Processing Systems* 31: 10919–30.

Smith, Anne C., and Emery N. Brown. 2003. "Estimating a State-Space Model from Point Process Observations." *Neural Computation* 15 (5): 965–91. https://doi.org/10.1162/089976603765202622.

Smith, Anne C., Loren M. Frank, Sylvia Wirth, Marianna Yanike, Dan Hu, Yasuo Kubota, Ann M. Graybiel, Wendy A. Suzuki, and Emery N. Brown. 2004. "Dynamic Analysis of Learning in Behavioral Experiments." *Journal of Neuroscience* 24 (2): 447–61. https://doi.org/10.1523/JNEUROSCI.2908-03.2004.

Sugawara, Michiyo, and Kentaro Katahira. 2021. "Dissociation between Asymmetric Value Updating and Perseverance in Human Reinforcement Learning." *Scientific Reports* 11 (1): 3574. https://doi.org/10.1038/s41598-020-80593-7.

The International Brain Laboratory, Valeria Aguillon-Rodriguez, Dora Angelaki, Hannah Bayer, Niccolo Bonacchi, Matteo Carandini, Fanny Cazettes, et al. 2021. "Standardized and Reproducible Measurement of Decision-Making in Mice." Edited by Naoshige Uchida and Michael J Frank. *ELife* 10 (May): e63711. https://doi.org/10.7554/eLife.63711.

Wiltschko, Alexander B., Matthew J. Johnson, Giuliano Iurilli, Ralph E. Peterson, Jesse M. Katon, Stan L. Pashkovski, Victoria E. Abraira, Ryan P. Adams, and Sandeep Robert Datta. 2015. "Mapping Sub-Second Structure in Mouse Behavior." *Neuron* 88 (6): 1121–35. https://doi.org/10.1016/j.neuron.2015.11.031.

Yu, Angela J, and Jonathan D Cohen. 2008. "Sequential Effects: Superstition or Rational Behavior?" *Advances in Neural Information Processing Systems* 21: 1873–80.

Zhang, Shunan, He Crane Huang, and Angela J. Yu. 2014. "Sequential Effects: A Bayesian Analysis of Prior Bias on Reaction Time and Behavioral Choice." *Proceedings of the Annual Meeting of the Cognitive Science Society* 36 (36). https://escholarship.org/uc/item/9jp3j5c7.