

# 1 Trial-history biases in evidence accumulation can give 2 rise to apparent lapses

3 Diksha Gupta<sup>1</sup>, Brian DePasquale<sup>1</sup>, Charles D. Kopec<sup>1</sup>, Carlos D. Brody<sup>1,2</sup>

4 <sup>1</sup>*Princeton Neuroscience Institute, Princeton University, Princeton, United States.*

5 <sup>2</sup>*Howard Hughes Medical Institute, Princeton University, Princeton, United States.*

6 Trial history biases and lapses are two of the most common suboptimalities observed  
7 during perceptual decision-making. These suboptimalities are routinely assumed to  
8 arise from distinct processes. However, several hints in the literature suggest that  
9 they covary in their prevalence and that their proposed neural substrates overlap –  
10 what could underlie these links? Here we demonstrate that history biases and appar-  
11 ent lapses can both arise from a common cognitive process that is normative under  
12 misbeliefs about non-stationarity in the world. This corresponds to an accumulation-  
13 to-bound model with history-dependent updates to the initial state of the accumulator.  
14 We test our model’s predictions about the relative prevalence of history biases and  
15 lapses, and show that they are robustly borne out in two distinct rat decision-making  
16 datasets, including data from a novel reaction time task. Our model improves the abil-  
17 ity to precisely predict decision-making dynamics within and across trials, by positing  
18 a process through which agents can generate quasi-stochastic choices.

## 19 **Introduction**

20 It has long been known that experienced perceptual decision makers deviate from the predictions of  
21 optimal decision-theory, displaying several suboptimalities in their decision-making. Among the  
22 most pervasive of these is the dependence of behavior on the recent history of observed stimuli,  
23 performed actions, or experienced outcomes, despite it being disadvantageous and leading to worse  
24 performance (Cho et al. 2002; Gold, Law, et al. 2008; Busse et al. 2011; Carandini and Churchland  
25 2013; Zhang et al. 2014; Fründ et al. 2014; Scott et al. 2015; Abrahamyan et al. 2016; Odoemene et  
26 al. 2018; Akrami et al. 2018; Pinto et al. 2018; Urai et al. 2019; Hermoso-Mendizabal et al. 2020;  
27 Mendonça et al. 2020; Lak et al. 2020; Mochol et al. 2021; Roy et al. 2021; The International Brain  
28 Laboratory et al. 2021; schematized in Fig 1A top). History biases may arise due to a strategy that  
29 is optimized for naturalistic settings, where continual learning of priors, action-values, or other  
30 decision variables helps agents adapt to changing environments, but is maladaptive in experimental  
31 settings where the statistics of the environment are stationary (Yu and Cohen, 2009; Molano-  
32 Mazon et al., 2021). To date, decision-theoretic models have accommodated history biases by

33 modeling them as a biasing factor on the perceptual evidence that drives choices (Nevin, 1969;  
34 Ratcliff and Rouder, 1998; Bogacz et al., 2006; Busse et al., 2011; Goldfarb et al., 2012; Gardner,  
35 2019; Urai et al., 2019; Hermoso-Mendizabal et al., 2020). In the predominant conceptualization  
36 of these models, history biases can be overcome with sufficient perceptual evidence.

37 A second widely-recognized but less studied suboptimality is the tendency to “lapse”, or  
38 make (asymptotic) errors that are immune to strong evidence (Wichmann and Hill 2001; Law  
39 and Gold 2009; Busse et al. 2011; Gold and Ding 2013; Brunton et al. 2013; Carandini and  
40 Churchland 2013; Wang et al. 2018; Pinto et al. 2018; Pisupati et al. 2021; Shushruth and Shadlen  
41 2021; schematized in Fig 1A bottom). Because lapses appear to be evidence-independent, they are  
42 assumed to arise from nuisance mechanisms that are separate from the perceptual decision-making  
43 process and are often imputed to ad-hoc noise sources such as inattention, motor errors etc.

44 However, several recent results suggest that these two suboptimalities may be linked in their  
45 origin. In primates, learning reduces dependence on recent trial history (Gold, Law, et al., 2008)  
46 as well as lapse probabilities (Law and Gold, 2009). Intriguingly, mice trained on a visual detec-  
47 tion task showed higher levels of history dependence on sessions with higher lapse probabilities  
48 (Busse et al., 2011). Moreover, lapses occur in runs (i.e. display Markov dependencies), rather  
49 than occurring with the traditionally assumed independent probabilities across trials (Ashwood et  
50 al., 2022). Furthermore, lapses have been proposed to reflect forms of exploration (Pisupati et al.,  
51 2021) that are sensitive to trial-by-trial updates of variables such as action value. Likewise, neural  
52 perturbations of secondary motor cortex and striatum in rodents have been shown to substantially  
53 impact both lapses (Erlich, Bialek, et al., 2011; Erlich, Brunton, et al., 2015; Yartsev et al., 2018;  
54 Guo et al., 2019; Pisupati et al., 2021; Sindreu et al., 2021) and trial-history influences on deci-  
55 sions (Siniscalchi et al., 2019; Sindreu et al., 2021). Together, these observations challenge the  
56 assumption that history biases and lapses have independent causes and raises the possibility that  
57 some of the variance ascribed to lapses emerges from history dependence.

58 We explore the idea that history biases reflect a misbelief about non-stationarity in the world,  
59 and demonstrate that normative decision-making under such beliefs gives rise to choices that are  
60 both history-dependent and appear to be evidence-independent (i.e. akin to lapses). This corre-  
61 sponds to an accumulation to bound process with a history dependent initial state. We fit this model  
62 to a large dataset of choices made by 152 rats trained on an auditory decision-making task. Despite  
63 heterogeneity in history biases and lapse rates in this population, we show that the a substantial

64 fraction of lapses can be explained by the presence of history dependence during evidence accu-  
65 mulation. Further, our model predicts the time it takes to make decisions. We test these predictions  
66 in a novel task in rats with reaction time reports, and show that it captures patterns of choices, reac-  
67 tion times, and their history dependence. This model significantly improves our ability to predict  
68 the temporal dynamics of decision variables within and across trials in perceptual decision making  
69 tasks, rendering choices that were previously thought to be stochastic, predictable.

## 70 **Results**

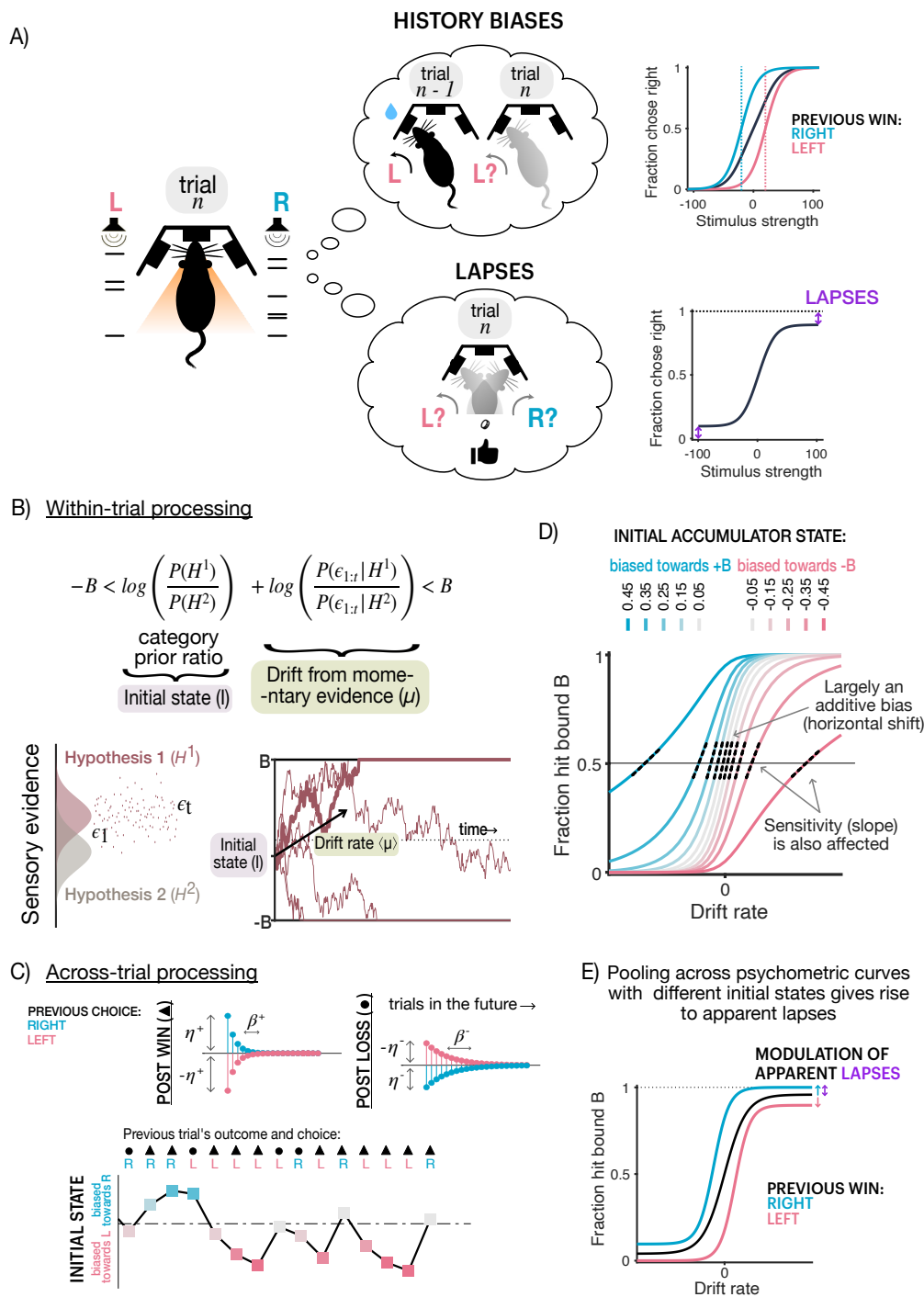
71 **A common mechanism produces history biases and apparent lapses** It is often assumed that  
72 well trained subjects in two-alternative forced choice (2AFC) tasks have faithfully learnt the like-  
73 lihood function and priors that determine the structure of the task (Bogacz et al., 2006; Gold and  
74 Shadlen, 2007). Under this assumption, the optimal decision making strategy entails combining  
75 any knowledge about prior prevalence of available options with the stream of incoming evidence  
76 until a desired threshold of confidence is reached in favor of one of the options <sup>1</sup> (Gold and Shadlen,  
77 2007; Dayan and Daw, 2008; Drugowitsch, Moreno-Bote, et al., 2012); Fig 1B top). This strategy  
78 converges to a drift diffusion model (DDM) when evidence is sampled continuously (Bogacz et al.,  
79 2006). In a DDM, one’s belief about the correct option maps onto a diffusing particle that drifts  
80 between two boundaries, where the first boundary the particle crosses determines the decision (Fig  
81 1B). Correspondingly, the initial state of this particle encodes the prior belief, and the drift rate is  
82 set by the likelihood of incoming evidence (Fig 1B). We refer to the evolving state of the particle  
83 in this model as ‘accumulated evidence’.

84 However, in general, subjects may not know that the task structure is stationary, and might  
85 incorrectly assume that it is constantly changing (Yu and Cohen, 2009). In this case, even ex-  
86 periented subjects would not converge to a static estimate of prior probabilities and likelihood  
87 functions, but would instead continually update them from trial to trial. Here we consider choice  
88 behavior that results from non-stationary beliefs about priors, which result in trial-to-trial updates  
89 to the initial accumulator states<sup>2</sup>.

---

<sup>1</sup>In tasks where the reliability of incoming evidence (controlled by stimulus strength) varies from one trial to the next, it has been shown that ideal observers should have time-varying bounds on the posterior (Drugowitsch, Moreno-Bote, et al. 2012. However under certain circumstances, stationary bounds over the summed stimulus have been shown to implement close-to-optimal collapsing bounds on the posterior, which is the regime we assume here for simplicity.

<sup>2</sup>For a treatment of non-stationary likelihood functions which yield variability in drift rate, see (Drugowitsch, Mendonça, et al., 2019; Mendonça et al., 2020)



**Figure 1: Trial history-dependent initial states give rise to apparent lapses (A)** Schematic of two commonly observed suboptimalities in decision-making: history biases (top) and lapses (bottom). (Left): Rat performing a perceptual decision making task, where it has to make one of two decisions (left, right) based on accumulated sensory evidence (auditory clicks on either side). Caption continued on next page.

Figure 1: (Previous page.) (Top left): History biases i.e. a tendency for the decision on the current trial ( $n$ ) to be inappropriately influenced by what happened on the previous trial ( $n-1$ ) in addition to the accumulated sensory evidence. In this example, a previously rewarding leftward decision is likely to be repeated. (Top right): Typically assumed effect of history bias on the psychometric curve, which is the proportion of rightward decisions as a function of the stimulus strength. History biases are thought to most strongly affect decisions when the sensory evidence is weak i.e. around the inflection point of the curve (threshold parameter), shifting it horizontally. (Bottom left): Lapses i.e. a tendency to make seemingly random choices on some trials, irrespective of the accumulated sensory evidence. (Bottom right): Typically assumed effect of lapses on the psychometric curve is vertically scaling the endpoints or asymptotes of the curve. **(B)** Standard normative model of within-trial processing during evidence accumulation. (Top) Decision rule that produces the most accurate decisions in the shortest amount of time, in which a decision is made when the summed log-ratios of category priors and likelihoods exceeds one of two decision bounds. This corresponds to a drift-diffusion process where the prior term sets the initial state ( $I$ ) and the rate of accumulating evidence sets the drift rate ( $\mu$ ). (Bottom left): Schematic of the generative model, where one of two hypotheses ( $H1$ ,  $H2$ ) produce noisy samples of evidence over time ( $\epsilon_t$ ). (Bottom right): Schematic of the aforementioned drift-diffusion process, showing a sample trajectory based on noisy evidence (bold line) that leads to a rightward decision when the positive bound is hit. Thin lines depict alternate trajectories based on different noisy instantiations of the same drift rate (black arrow). **(C)** Model of across trial processing that can accommodate several forms of prior updates. Past choices and outcomes can additively affect the initial state with different magnitudes ( $\eta$ ) and exponentially decaying timescales ( $\beta$ ) depending on whether they were wins (top left) or losses (top right). (Bottom): Example sequence of trials, labelled by whether they follow a previous win (triangles) or previous loss (circles) on right (R) or left (L) choices, showing the cumulative effect of trial history on initial state updates. Colors denote different initial state biases, same as (C). **(D)** Effect of initial state values on psychometric functions. Colors denote different initial state levels, towards the positive (blue) or negative (pink) decision bounds. Small deviations from 0 in the initial state (grey) lead to largely additive, horizontal biases in the psychometric curve whereas larger deviations (saturated colors) have more complex effects, additionally reducing its effective slope (dotted black lines) or “sensitivity” to the stimulus. **(E)** Pooling different initial state biases gives rise to apparent lapses. Psychometric function (black) pooled across trials with different initial state biases (due to history-based updating) has apparent lapses (purple arrow), moreover conditioning the psychometric curve on whether the previous trial was a rightward (blue) or leftward (pink) win reveals a modulation of these apparent lapses by trial history.

90 We assume that the initial state of the accumulator ( $I$ ) is set based on the exponentially  
 91 filtered history of choices and outcomes on past trials. Each unique choice-outcome pair (denoted  
 92 by  $h$ ; Fig 1C) is tracked by its own exponential filter ( $i^h$ ). On each trial  $n$ , each filter  $i^h$  decays by  
 93 a factor of  $\beta^h$  and is incremented by a factor of  $\eta^h$  depending on the choice-outcome pair on the  
 94 previous trial:

$$i^h(n) = \beta^h i^h(n-1) + \eta^h 1^h(o_{n-1}) \quad \text{where } h = \{Rw, Lw, Rl, Ll\} \quad (1)$$

95  $\{Rw, Lw, Rl, Ll\}$  represent the possible choice-outcome pairs: right-win, left-win, right-loss, and  
 96 left-loss respectively.  $o_{n-1}$  is the choice-outcome pair observed on trial  $(n-1)$  and  $1^h(o_{n-1})$  is an

97 indicator function that is 1 when  $o_{n-1} = h$  and is 0 otherwise. The initial state of accumulation,  $I$   
98 on trial  $n$  is given by the sum of these individual exponential filters:

$$I(n) = i^{Rw}(n) + i^{Lw}(n) + i^{Rl}(n) + i^{Ll}(n) \quad (2)$$

99 Such a filter can approximate optimal updating strategies under a variety of non-stationary beliefs.  
100 As an example, we show that this exponential filter can successfully approximate initial state up-  
101 dates during Bayesian learning of priors under the belief that the prior probabilities of the two  
102 hypotheses can undergo unsignaled jumps (Supp Fig. 1; Yu and Cohen 2009; Zhang et al. 2014).  
103 Nevertheless, we use this more flexible parameterization to allow for asymmetric learning from  
104 different choices and outcomes, which could be beneficial under generative models where one be-  
105 lieves that one category persists for longer than another (requiring different decay rates), or correct  
106 and incorrect outcomes are not equally informative (requiring different update magnitudes). For  
107 instance, in a prior-tracking experiment where previous correct choices had a cumulative effect,  
108 but errors had a resetting effect (Hermoso-Mendizabal et al., 2020), this could be captured in the  
109 exponential filter by faster decay rates for errors.

What are the consequences of such trial-by-trial updating of initial accumulator states for  
choice behavior? In a DDM, for a given initial state  $I$  and drift rate  $\mu$ , the probability of choosing  
the option corresponding to bound  $B+$  is given by:

$$P(B+) = \frac{1 - e^{-2\mu(B+I)/\sigma^2}}{1 - e^{-4\mu B/\sigma^2}} \quad (3)$$

110 where  $B$  is the magnitude of the bound and  $\sigma^2$  is the squared diffusion coefficient (derived from  
111 Palmer et al. 2005). The resultant psychometric curves for different values of initial accumulator  
112 states are plotted in Fig 1D. This expression reduces to a logistic function of  $\mu B/\sigma^2$  only when  
113  $I = 0$ . Small deviations in the initial state largely manifest as additive biases to the evidence,  
114 shifting psychometric curves horizontally towards the option favored by the initial state. This  
115 corresponds to a change in the psychometric “threshold”, i.e. the x-axis value at its inflection point  
116 (Fig 1D lighter colors). Interestingly, large deviations in the initial state produce qualitatively  
117 different effects on choices (Fig 1D darker colors). They not only bias the choices towards the  
118 option consistent with the initial state but additionally reduce the effective “sensitivity” to evidence.  
119 This can be seen as reduction in slope at the inflection point of the psychometric curve (Fig 1D  
120 dashed lines). Therefore, trial to trial deviations in the initial state produce history-biased choices  
121 which have differently diminished dependence on the evidence.

122 The average choice behavior obtained by pooling choices with different history-biased ini-  
123 tial states is a mixture of psychometric curves with varying thresholds and sensitivity to perceptual  
124 evidence. Such a psychometric curve is heavy-tailed (Shen and Ma, 2019; Nguyen, Josić, et al.,  
125 2019) and appears to have asymptotic errors or “lapse rates” (Fig 1E, black curve). These asymp-  
126 totic errors are not truly evidence-independent, random decisions or *true lapses*, rather they are  
127 “*apparent lapses*” arising from evidence accumulation with deterministic history-based updates to  
128 the initial accumulator state. In such a setting, the psychometric curves obtained by conditioning on  
129 past trials’ choice and outcome, or *history conditioned* psychometric curves, are both horizontally  
130 and vertically shifted, i.e. they show history-dependent modulations in both threshold and lapse  
131 rate parameters (Fig 1E, Supp Fig. 2B). In this formulation, trial-history modulated lapse rates  
132 are uniquely produced by history-biased initial accumulator states (and therefore reflect apparent  
133 lapses), in contrast to lapse rates observed in the unconditioned psychometric curve which might  
134 have additional extraneous causes (Wichmann and Hill, 2001; Ashwood et al., 2022; Pisupati et  
135 al., 2021), and therefore reflect both apparent and true lapses.

136 In this model, because history modulations of psychometric thresholds and lapse rates arise  
137 from one unified process, they are not allowed to vary independently of the decision-making pro-  
138 cess, or of each other. Rather their relative magnitudes are intimately coupled with and constrained  
139 by accumulation variables. For instance, increased magnitudes or timescales of initial state updat-  
140 ing produce large fluctuations in the initial accumulator state across trials. This in turn reduces the  
141 effective sensitivity of the accumulation process to evidence, giving rise to more apparent lapses  
142 and history biases (Supp Fig. 2A). Similarly, changes in within-trial parameters of accumulation  
143 can dramatically influence these history modulations (Supp Fig. 2C). Decisions made with smaller  
144 accumulator bounds are more sensitive to initial state modulations, and therefore give rise to more  
145 apparent lapses and higher modulations of lapse rates and thresholds. Higher levels of sensory  
146 noise have a similar effect, yielding more apparent lapses, consistent with recent reports of lapse  
147 rates being modulated by sensory uncertainty (Pisupati et al., 2021). Finally, impulsive integration  
148 strategies that overweigh early evidence rather than accumulating uniformly (Bogacz et al., 2006)  
149 exaggerate the influence of initial states, producing more apparent lapses and history biases.

150 Some definitions:

- 151 • *Lapse rate*: Lapse rates capture the difference between perfect performance and  
152 observed performance at the asymptotes, measured through sigmoidal fits to the

153 psychometric curves.

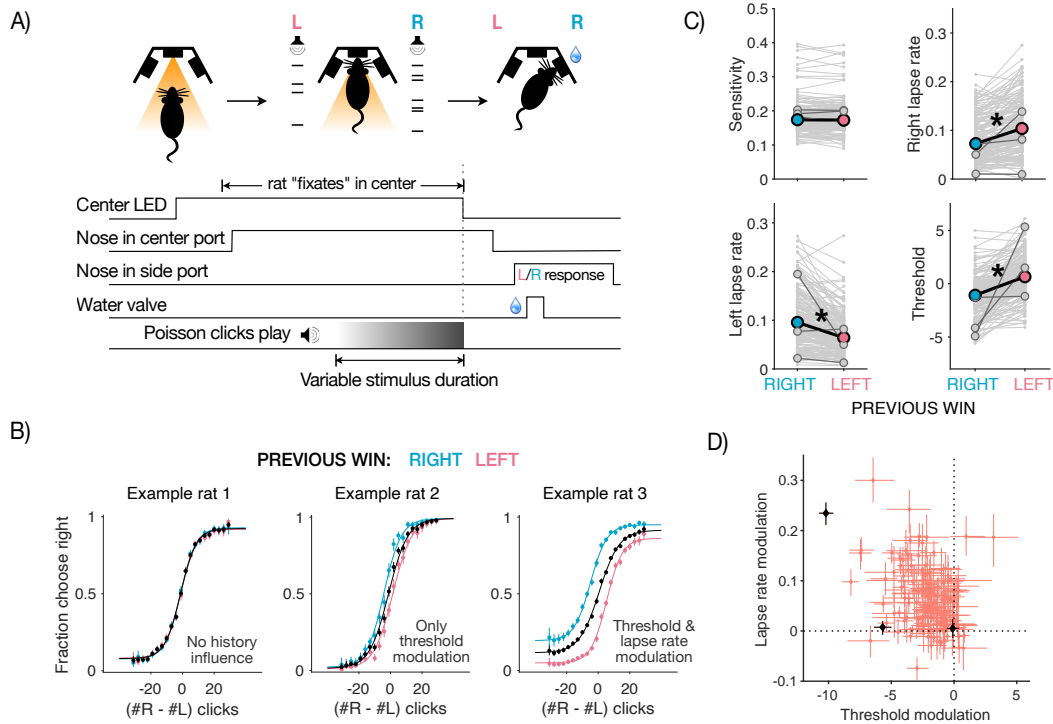
- 154 • *True lapse*: A true lapse is an independent cognitive process through which
- 155 agents generate stochastic evidence-independent choices.
- 156 • *Apparent lapses*: Apparent lapses are deterministic evidence-dependent choices,
- 157 that nonetheless contribute to lapse rates.

158 **Rats display varying degrees of history-dependent threshold and lapse rate modulation** We  
159 sought to test if the co-modulations posited by our model are present in rat decision-making  
160 datasets, in order to ascertain whether a unified explanation could underlie the links between his-  
161 tory biases and lapses.

162 We first examined whether and how rat decision-making strategies were affected by trial  
163 history. We analyzed choice data from 152 rats ( $37522 \pm 22090$  trials per rat, mean  $\pm$  SD; Supp  
164 Fig.3A) trained on a previously developed task that requires accumulation of pulsatile auditory  
165 evidence over time ('Poisson Clicks' task, Brunton et al. 2013). In this task, the subject is presented  
166 with two simultaneous streams of randomly-timed discrete pulses of evidence, one from a speaker  
167 to their left and the other to their right (Fig 2A). The subject must maintain fixation throughout  
168 the stimulus, and subsequently orient towards the side which played the greater number of clicks  
169 to receive a water reward. The trial difficulty, stimulus duration, and correct answer were set  
170 independently on each trial. Because this task delivers sensory evidence through randomly but  
171 precisely timed pulses, it provides high statistical power to characterize decision variables that  
172 give rise to the choice behavior.

173 Rats performed this task accurately ( $0.79 \pm 0.04$ , mean accuracy  $\pm$  SD, Supp Fig.3B). Per-  
174 formance was stable with little to no change in accuracy across trials (mean slope  $\pm$  SD across rats  
175 of linear fit to hit rate over trials:  $1.13 \times 10^{-7} \pm 8.90 \times 10^{-7}$ ; Supp Fig.3C) reflecting asymptotic  
176 behavior rather than task acquisition. Rats showed history dependence in their choices, largely  
177 tending towards a "win-stay, lose-switch" dependence (Supp Fig.3E). We found substantial indi-  
178 vidual variability in the dependence of rats' choices on history in the dataset. Some rats were  
179 weakly influenced by history (Fig 2B left) while others showed a history-dependent modulation of  
180 the psychometric threshold parameter (Fig 2B middle) or a history-dependent modulation of both  
181 threshold and lapse rate parameters (Fig 2B right). The population as a whole most closely resem-  
182 bles the Example rat 3, with both threshold and lapse rate parameters being significantly different  
183 following left and right wins while sensitivity is not affected ( $p = 0.8$  for sensitivity,  $3 \times 10^{-17}$



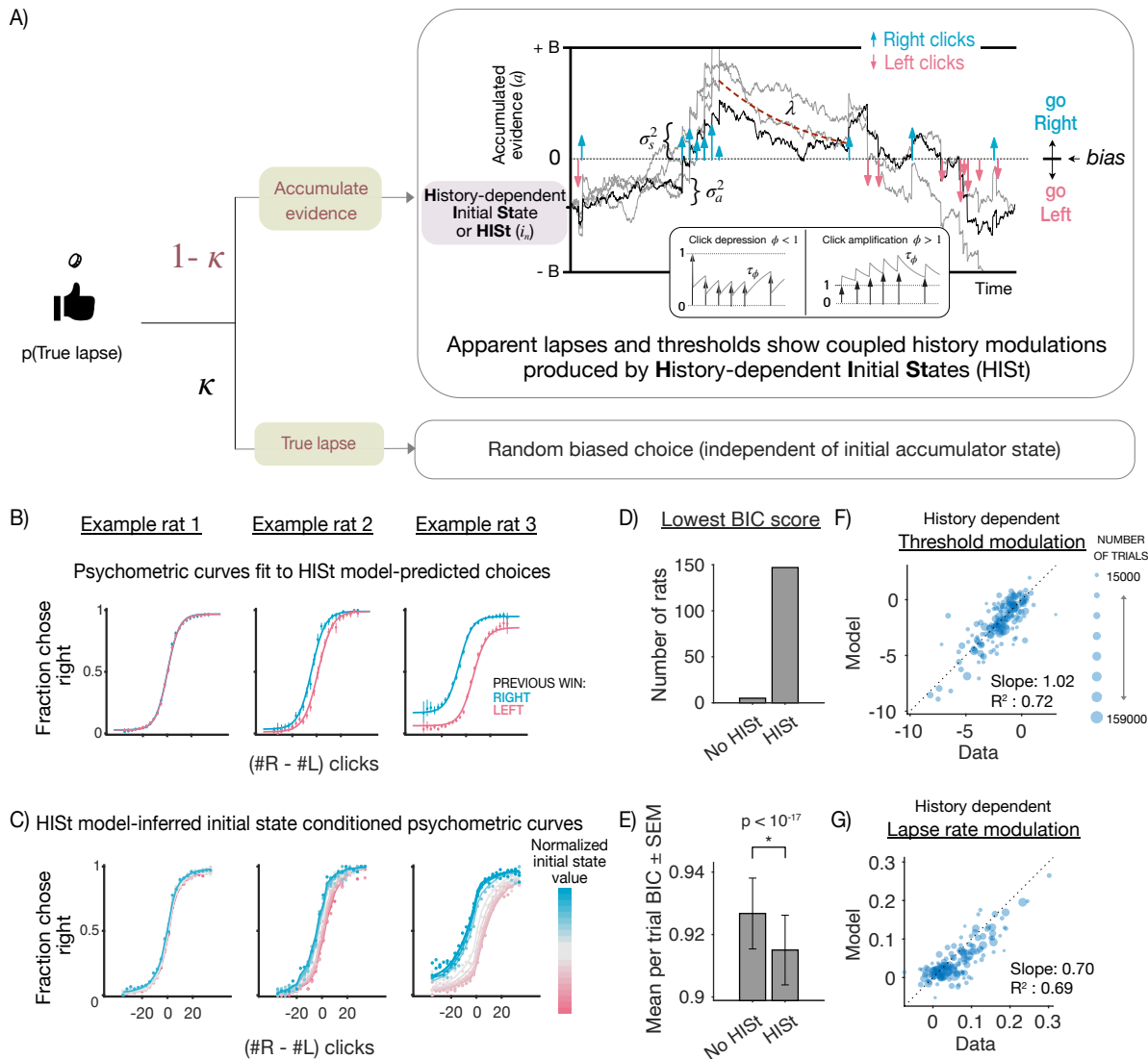


**Figure 2: History-dependent threshold and lapse rate modulations in a large-scale rat dataset (A)**

Schematic of evidence accumulation task in rats: (Top): Phases of the ‘Poisson clicks’ task, including trial initiation in center port (left), evidence accumulation based on two streams of Poisson-distributed auditory clicks (middle) and choice report in one of two side ports followed by water reward for correct choices (right). (Bottom): Time-course of trial events in a typical trial. (B) Individual differences in history-dependence: Psychometric functions of three example rats from a large-scale dataset, displaying different kinds of history modulation. Choices are plotted conditioned on previous left (blue), right (pink) or all wins (black). (Left): Example rat with no history-dependence in choices, resembling the ideal observer. (Middle): Example rat with modulations of the threshold parameter alone, resembling the dominant conceptualization of history bias. (Right): Example rat with history-dependent modulation of both threshold and lapse rate parameter, similar to the majority of the population. (C) Dataset displays significant modulations of both threshold and lapse rate parameters: Scatters showing parameters of psychometric functions following leftward wins (post left, blue) or rightward wins (post right, pink). Each pair of connected gray points represents an individual animal, solid colored dots represent average parameter values across animals. Trial history does not significantly affect the sensitivity parameter (top left) but significantly affects left, right lapse rate and threshold parameters (top right and bottom panels). (D) Scatter comparing threshold and lapse rate modulations in the entire population. Each dot is an individual animal, error bars are  $\pm 95\%$  bootstrap CIs. Black points represent example rats. The majority of the population lies in the top left quadrant, showing co-modulations of both threshold and lapse rate parameters by history.

184 for bias,  $8 \times 10^{-8}$  for left lapse,  $6 \times 10^{-7}$  for right lapse, Mann-Whitney U-test, Fig 2C). As pre-  
 185 dicted by our model (Fig 1E), trial-history biased both threshold and lapse rate parameters in the

186 same direction (e.g. both biased toward rightward choices following right rewards). Moreover,  
187 the vast majority of rats show co-modulations of both parameters by history (Pearson's correlation  
188 coefficient:  $r = -0.35$ ,  $p = 7.28 \times 10^{-6}$ ; Fig 2D). Across rats, on average  $17 \pm 12\%$  of total lapse  
189 rates are modulated by trial history and therefore could potentially reflect apparent rather than true  
190 lapses (Supp Fig.3D). These findings support the conclusion that rat decision-making strategies,  
191 while idiosyncratic, largely show history-dependent effects consistent with our model. Next, we  
192 tested the model more directly using trial-by-trial model fitting.



**Figure 3: History-dependent initial states capture comodulations in thresholds and lapse rates in the data** (A) Schematic of the model (accumulator with HIST) used to fit rat data in the Poisson Clicks task. (Top): Schematic of accumulation-to-bound model whose initial states are modulated by trial history producing history-dependent apparent lapses and threshold modulations. The model consists of sensory noise ( $\sigma_s^2$ ) in click magnitudes, adaptation of successive click magnitudes based on an adaptation scale ( $\phi$ ) and timescale ( $\tau_\phi$ ), accumulator noise ( $\sigma_a^2$ ) added at each timestep, leak in the accumulator ( $\lambda$ ), and decision bounds  $\pm B$ . We refer to this accumulator model with History-dependent Initial States as ‘HIST’ (Bottom): On  $\kappa$  fraction of trials, the model occasionally chooses a random action irrespective of the initial state and stimulus, with some bias ( $\rho$ ) reflecting a motor errors or random exploration. These true lapses are not modulated by history, such that any history modulations arise from the initial states alone. (B) Model fits to individual rats: Psychometric data from 3 example rats conditioned on previous rightward (blue) or leftward (wins), overlaid on model-predicted psychometric curves (solid line) from the accumulation with HIST model. The three example rats were chosen to illustrate the diversity of history effects in the dataset, ranging from no history effects (left) - to history effects that largely created horizontal biases (center) and history effects that additionally affected lapse rates (right). Caption continued on next page.

Figure 3: (Previous page.) **(C)**: Psychometric curves (solid line) from the same example rats conditioned on model-inferred initial states (colors from pink to blue), showing a similar pattern to analytical predictions in Fig 1D. **(D)** Distribution of best fitting models for individual rats: Overall bar height for each model denotes the total number of rats for which that model scored the lowest BIC score. **(E)** Model comparison using BIC by pooling per trial BIC score across rats and computing mean. Lower scores indicate better fits. Mean of per trial BIC scores across rats were significantly lower for model with HIsT ( $p = 9.85 \times 10^{-18}$ , paired t-test). Error bars are SEM. **(F)** Individual variations in history modulations captured by the accumulator model with HIsT: History modulations of threshold parameters measured from psychometric fits to the raw data (x-axis) v.s. model predictions (y-axis). Individual points represent individual rats, point sizes indicate number of trials. The model captures a majority of the variability as evidenced by the points lying close to the unity line. **(G)** same as (F) but for history dependent lapse rate modulations. The model captures a majority of the variability in lapse rate modulations, implying that the magnitude of threshold and lapse rate modulations are coupled as predicted by our model, and that history-dependent initial accumulator states contribute to apparent lapses in this dataset.

193 **History-dependent initial states capture comodulations in thresholds and lapse rates in the**  
194 **data** To test whether the observed history modulations in thresholds and lapse rates arise from  
195 trial-by-trial updates to the initial accumulator state, we extended an accumulator model previ-  
196 ously adapted to this pulsatile task (Brunton et al., 2013) to incorporate History-dependent Initial  
197 States (abbreviated as HIsT, Fig 3A). As before, we model this history-dependence using an expo-  
198 nential filter over past trials’ choices and outcomes (Fig 1C). Hence, across trials the accumulator  
199 model with HIsT produces apparent lapses, as well as coupled history modulations in psychometric  
200 threshold and lapse rate parameters.

201 Within a trial, our accumulator model leverages knowledge of the timing of each evidence  
202 pulse to model the sensory adaptation process as well as to estimate the noise and drift of the  
203 accumulator variable (Fig 3A top bubble, Methods). The model includes a feedback parameter  
204 that controls whether integration is leaky, perfect, or impulsive. Following Brunton et al. 2013,  
205 this model also includes (biased) random choices independent of the accumulator value on a small  
206 fraction of trials ( $\kappa$ ) - we consider decisions arising from this process to be “true lapses” because  
207 they are evidence independent, unlike apparent lapses which still retain some evidence-dependence  
208 (Fig 3A bottom bubble).

209 We performed trial-by-trial fitting of the accumulator model with and without History- de-  
210 pendent Initial States (HIsT) to choices from each rat using maximum likelihood estimation (Meth-  
211 ods). We find that the accumulator model with HIsT captures both psychometric curve threshold  
212 and lapse rate modulations well across different regimes of rat behavior, as evident from fits to

213 example rats (Fig 3B). Moreover, conditioning rats' psychometric curves on model-inferred initial  
214 state values reveals that the initial state captures a large amount of variance in choice probabil-  
215 ities (Fig 3C), resembling theoretical predictions (Fig 1C). This shows that the initial state is a  
216 key explanatory variable underlying choice variability both across and within individuals, that  
217 jointly modulates multiple features of the empirical psychometric curves in a parametric fashion.  
218 We used Bayes Information Criterion (BIC) to determine whether adding HlSt to the accumula-  
219 tor model was warranted (Fig 3D-E). Individual BIC scores recommended that adding HlSt was  
220 warranted in 147/152 rats (Fig 3D). This model also best captured choices across the population  
221 as a whole, with significantly lower mean BIC scores across rats (Mean per trial BIC score for  
222 HlSt:  $0.91 \pm 0.01$  vs no HlSt:  $0.93 \pm 0.01$ ,  $p = 9.85 \times 10^{-18}$ , paired t-test; Fig 3E). Next, we  
223 compared the psychometric threshold and lapse rate modulations produced by this model to the  
224 modulations in the data, as determined by conditioning the psychometric functions on trial-history  
225 (Fig 3B). As predicted, the model successfully accounted for modulations in both these distinct  
226 psychometric features via the singular process of trial-by-trial history-dependent updates to the  
227 initial accumulator state. Across individuals, the model with HlSt captured a substantial amount  
228 of variance [ $R^2 = 0.72$  (threshold parameter),  $R^2 = 0.69$  (lapse rate parameter)] and showed good  
229 correspondence to the empirical modulations in data [slope= 1.02 (threshold parameter), slope  
230 = 0.70 (lapse rate parameter)].

231 In our model, apparent lapses show history modulations since they are produced by history-  
232 dependent initial accumulator states, while true lapses do not since they result from an occasional  
233 flip in the final choice and are independent of the accumulator value (following Brunton et al.,  
234 2013). Such kinds of true lapses could reflect errors in motor execution or random exploratory  
235 choices made despite successful accumulation (Supp Fig 4B). However true lapses could also occur  
236 due to inattention, i.e. an occasional failure to attend to the stimulus. In such cases, the optimal  
237 strategy devoid of sensory evidence is to deterministically choose the side favored by the initial  
238 accumulator state (Supp Fig 4C). Therefore, inattentive true lapses, while remaining evidence  
239 independent, may nevertheless be modulated by history due to their initial state dependence. In  
240 order to account for this possibility, we fit an additional "inattentive" variant of the accumulator  
241 model with HlSt (Supp Fig 4A,C), and found that it was closely matched on BIC scores with  
242 the previous model which we label as the "motor error" variant (Supp Fig 4E,F). Moreover, the  
243 inattentive variant, which additionally allows true lapses to depend on history, only captured  
244 slightly more variance in history modulations of lapse rates, at the expense of history modulations  
245 of thresholds (Supp Fig 4D). These minor improvements support the hypothesis that apparent

246 lapses produced by history-dependent initial states (rather than true lapses due to motor error or  
247 inattention) are the major driver of history-dependent co-modulations in psychometric thresholds  
248 and lapse rates in the dataset.

249 To summarize, our model predicted that the initial accumulator state should be the under-  
250 lying variable that jointly drives history-dependence in thresholds and lapse rates – implying that  
251 our accumulator model with HIsT should be able to simultaneously capture variability in both  
252 these parameters across rats. Our rat dataset strongly supports this prediction, lending evidence to  
253 the hypothesis that history-dependent initial states give rise to apparent lapses, and are the com-  
254 mon cognitive process that underlie links between these two suboptimalities that were previously  
255 thought to be distinct from each other.

256 **Reaction times support history-dependent initial state updating** In our model with history-  
257 dependent initial accumulator states, the time it takes for the accumulation variable to hit the bound  
258 determines the duration that the subject deliberates for, before committing to a choice. Therefore  
259 in addition to choices the model makes clear predictions about subjects' reaction times (RTs). We  
260 sought to test if these predictions are borne out in subject RTs.

261 To this end, we trained rats ( $n = 6$ ) on a new variant of the auditory evidence accumulation  
262 task (Fig 2A), with two key modifications that allowed us to collect reaction time reports (Fig 4A).  
263 First, in this new task the stimulus is played as long as the rat maintains their nose in the center  
264 port (or “fixates”) and stops immediately when this fixation is broken. Second, in this task the  
265 rat has to correctly report which speaker's auditory click train is sampled from a higher Poisson  
266 rate to receive a water reward (unlike the non-reaction time task where subject has to report the  
267 side which played the greater number of clicks). Rats perform this task with high accuracy (Fig  
268 4B left panel, average accuracy:  $0.75 \pm 0.02$ , number of trials  $37205 \pm 14247$ , mean  $\pm$  SD).  
269 Similar to the previously analyzed data, their choices are impacted by recent trial history (Fig 4B  
270 right panel). Moreover, trial-history dependent modulation of psychometric function parameters  
271 (Fig 4C) resembles that of the non-reaction time task (Fig 2C;  $p = 0.69$  for sensitivity, 0.004 for  
272 threshold, 0.02 for left lapse rate, 0.02 for right lapse rate, Mann-Whitney U-test). Once again,  
273 this history modulation of both psychometric threshold and lapse rate parameters in tandem is  
274 consistent with our singular accumulator model with history-dependent initial states.

275 Moreover, reaction times (RTs) of these rats display several signatures predicted by our

276 model (Fig 4D-F). First, trial-to-trial variability in the initial state of the accumulator is expected to  
277 give rise to shorter RTs on error trials compared to correct trials (Fig 4E, left; Ratcliff and Rouder  
278 1998). This is because trials in which the initial state is closer to the incorrect bound are more  
279 likely to be errors, but because of the closer bound they are also likely to hit it faster. This is unlike  
280 a standard DDM with no trial-to-trial variability in parameters, where RTs for correct and error  
281 trials are of similar magnitudes (Fig 4D, left). Indeed in the rat dataset, error RTs are consistently  
282 shorter than correct RTs across rats (Fig 4F, left). Second, initial state updates towards previously  
283 rewarded choices (such as in a win-stay agent) are expected to produce shorter RTs when the cur-  
284 rent stimulus favors the previously rewarded choice (Fig 4E, middle; Yu and Cohen 2009; Goldfarb  
285 et al. 2012). We find that this signature is also present in the dataset across rats (Fig 4F, middle).  
286 Finally, variability in the initial state is most influential early in the decision process, predicting  
287 that the majority of history dependence in choices occurs on trials with fast RTs (Fig 4E, right;  
288 Urai et al. 2019). Indeed, the data displays this pattern as well, with repetition bias being most  
289 prominent for short RTs, disappearing and turning into a weak alternation bias for long RTs (Fig  
290 4F, right). Taken together, these three signatures offer strong, complementary evidence from RTs  
291 for the prevalence of history-dependent initial states in rats performing this evidence accumulation  
292 task.

293 We directly test if our model can simultaneously capture reaction time patterns and history-  
294 modulation of psychometric threshold and lapse parameters by jointly fitting choices and RTs  
295 of individual subjects in a trial-by-trial fashion (see Methods). We find that the history-dependent  
296 initial state model jointly captures patterns of choices, reaction times, and their history modulations  
297 in the data (Fig 4G - fits from example rat, Supp Fig. 5 - fits from all rats). This model accounts  
298 for substantial variance in history-dependent threshold and lapse rate modulations (Fig 4H). We  
299 also fit a hybrid variant of the accumulator model with HIST that flexibly allows true lapses to be  
300 motor-error like and unaffected by history, or inattention-like and additionally be modulated by  
301 history (Supp Fig. 6A,B). While this model has a better BIC and leads to a slight improvement in  
302 correspondence to the history modulation of psychometric lapse rates, it does so at the cost of

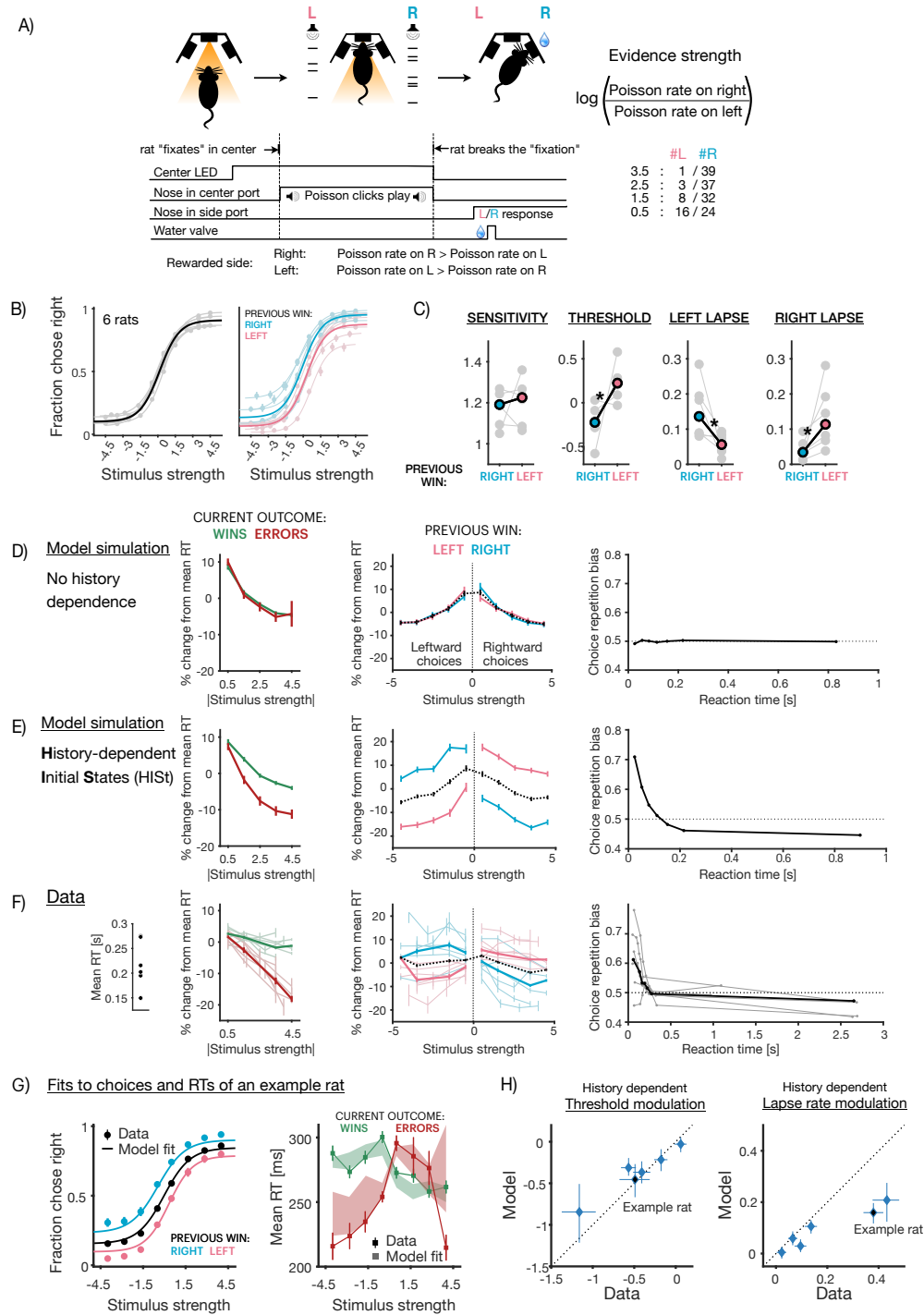


Figure 4: **Model predictions about reaction times are borne out in data** (A) Schematic of reaction time task in rats, with similar structure to (Fig 2A), with two modifications: rats are allowed to break “fixation” anytime during the trial and make a choice, and are rewarded for choosing the side with the higher Poisson rate, encouraging longer sampling for more accurate estimates. Caption continued on next page.



Figure 4: (Previous page.) **(B)** Average choice behavior on all trials (left) and following previous right or left wins (right) of 6 rats on the reaction time task (solid line), overlaid on individual rat behavior (translucent lines) **(C)** Average parameters (solid points) of history-conditioned psychometric curves, overlaid on individual parameters (translucent points) showing significant history modulations in threshold and lapse rate parameters ( $p < 0.05$ , Mann-Whitney U-test) **(D-F)** Reaction time signatures D) expected from accumulator models with no history dependence in initial states, E) expected from accumulator models with history dependent initial states and F) observed in data. First, error reaction times are expected to be shorter if initial states are history dependent, as seen in data (Left column, red curves are below green curves in E,F). Second, reaction times on trials following right wins are expected to be lower on rightward stimuli (positive half of x-axis), and similarly following left wins (Middle column, blue (pink) curves on the right (left) are lower than dotted lines in E,F). Finally, repetition biases in choices are expected to occur more frequently for short reaction times, when the effect of initial states is strong (Right column, curves are above dotted line for smaller RTs in E,F). **(G)** Joint fits of the accumulator model with history-dependent initial states to choices (left) and reaction times (right) of an example rat show good correspondence to data. Data represented by points (circles: choices, squares: reaction times) and model fits represented by lines (choices) or shaded bars (reaction times, thickness represents 95% bootstrap prediction intervals). Reaction times (right) are split by wins (green) or errors (red). **(H)** Scatter plot showing correspondence between history modulations in threshold (left) or lapse rate (right) parameters derived from data (x-axis) and model fits (y-axis). Individual points represent individual rats.

303 correspondence to modulations in psychometric thresholds (Supp Fig. 6C-E), once again largely  
304 implicating HIsT and its resultant apparent lapses (rather than true lapses) in the co-modulation of  
305 both parameters.

306 Overall, these results show that the history-dependent initial state updates that we invoked to  
307 explain apparent lapses in rodent data are corroborated by their reaction times, and accounting for  
308 them can help render a sizable fraction of decisions — that would have been otherwise attributed  
309 to noise — more predictable both within and across trials.

## 310 Discussion

311 History biases and lapses have both long been known to impact perceptual decision-making across  
312 species. However, they have largely been assumed to be distinct from each other, despite their  
313 frequent co-occurrence and co-modulation. Here, we propose that normative accumulation under  
314 misbeliefs of non-stationarity can produce both history biases and apparent lapses, offering an ex-  
315 planatory link between the two suboptimalities. This corresponds to history-dependent trial-to-trial  
316 updates to the initial state of an evidence accumulator. We show that such updates produce choices  
317 with varying biases in psychometric thresholds as well as varying sensitivities to evidence, yield-

318 ing apparent, history-modulated lapse rates when choices are averaged across trials (Fig 1). Our  
319 model postulates that the initial state of the accumulator is a key underlying variable that jointly  
320 modulates psychometric thresholds and lapse rate parameters, with the exact nature of this comod-  
321 ulation determined by the within and across trial parameters governing evidence accumulation. We  
322 tested this model in a large rat dataset consisting of choices from 152 rats (Fig 2) and confirmed its  
323 predictions using detailed model-fitting. We found that the singular process of history-dependent  
324 initial states successfully captured a substantial amount of variance in history modulations of both  
325 thresholds and lapse rates in the dataset (Fig 3). Finally, we tested the reaction time predictions  
326 of the model in a novel task in rats, and confirmed that the data showed signatures of initial state  
327 updating. The model could successfully capture choices, reaction times, and history modulations  
328 in psychometric thresholds and lapse rates (Fig 4). Altogether, our results suggest that history bi-  
329 ases and a substantial amount of variance attributed to lapses may reflect a common mechanistic  
330 process, whose evolution can be precisely tracked both within and across trials.

331 History biases in perceptual decision making tasks have been modeled using initial state up-  
332 dates to DDMs in humans and non-human primates (Gold, Law, et al., 2008; Goldfarb et al., 2012;  
333 Zhang et al., 2014). These studies tended to have relatively small magnitudes of history bias, and  
334 miniscule lapse rates, hence being well captured by small deviations in the initial state of a DDM,  
335 which largely yield horizontal shifts in the psychometric function. This regime of initial state  
336 updates is well approximated by a logistic function with additive biases, which is the dominant  
337 descriptive model used to characterize history-dependent psychometric curves (Busse et al., 2011;  
338 Carandini and Churchland, 2013; Fründ et al., 2014; Abrahamyan et al., 2016; Gardner, 2019;  
339 Pinto et al., 2018; Odoemene et al., 2018; Urai et al., 2019; Hermoso-Mendizabal et al., 2020;  
340 Roy et al., 2021; Ashwood et al., 2022; Bolkan et al., 2022). However, as we demonstrate, when  
341 deviations in the initial state are large, this logistic approximation breaks down. This fact has been  
342 overlooked in much of the literature. Consequently, even in datasets with large history biases and  
343 lapses, the logistic formulation continues to be favored (Odoemene et al., 2018; The International  
344 Brain Laboratory et al., 2021; Roy et al., 2021; Ashwood et al., 2022), albeit requiring additional  
345 components. Our demonstration predicts that the full range of initial state effects should resem-  
346 ble concurrent, trial-by-trial changes in both threshold and sensitivity parameters of the logistic  
347 function. Indeed, Ashwood et al. 2022 found that apparent lapses in several rodent datasets can be  
348 better captured by runs of trials with such concurrent modulations, yielding biased “disengaged”  
349 states.

350 In our treatment, we only considered history-dependent updates to the initial state of a DDM.  
351 Such a mechanism is normative under non-stationary beliefs about the prior <sup>3</sup>, which is our fa-  
352 vored interpretation as it aligns with other studies of history biases (Gold, Law, et al., 2008; Yu  
353 and Cohen, 2009; Summerfield and Koechlin, 2010; Goldfarb et al., 2012; Mulder et al., 2012;  
354 Abrahamyan et al., 2016; Molano-Mazon et al., 2021). Nevertheless, these updates may also re-  
355 flect other heuristic strategies (Gigerenzer and Gaissmaier, 2011) which we accommodate using  
356 our flexible parameterization of initial state updates. Animals may entertain non-stationary be-  
357 liefs about other elements of the decision process, such as the rewards or likelihoods (Dayan and  
358 Daw, 2008; Mendonça et al., 2020; Lak et al., 2020; Pisupati et al., 2021). Normative updating  
359 in such situations still reduces to initial state updates in simple settings (for e.g. non-stationary  
360 rewards for a single difficulty; Simen et al. 2009; Rorie et al. 2010), but in more complex ones  
361 it *additionally* affects drift rates (Palmer et al., 2005; Eckhoff et al., 2008; Hanks, Mazurek, et  
362 al., 2011; Drugowitsch, Mendonça, et al., 2019; Fan et al., 2018; Urai et al., 2019; Mendonça  
363 et al., 2020). Trial-to-trial variability in drift rates is known to give rise to longer error RTs than  
364 correct RTs (Ditterich, 2006a; Ditterich, 2006b; Drugowitsch, Moreno-Bote, et al., 2012; Nguyen  
365 and Reinagel, 2020), which is a signature often reported in monkeys and humans (Roitman and  
366 Shadlen, 2002; Shevinsky and Reinagel, 2019). Although we don't see this reaction time signature  
367 of drift rate variability in our dataset – instead we see signatures of initial state variability, with  
368 error RTs being shorter than correct RTs, rather than longer – drift rate updates may be another  
369 potential mechanism by which history-modulated apparent lapses could arise.

370 Lapse rates are often considered to be a mixed bag comprising several different noise pro-  
371 cesses, yet most studies so far have focused on one or more of these component processes in  
372 isolation (Pisupati et al., 2021; Ashwood et al., 2022). In this work, we have attempted a more  
373 expansive approach of considering multiple processes at once, in an attempt to partition lapse rate  
374 variance into mixtures of deterministic and stochastic components. We distinguished apparent  
375 lapses that interact with sensory evidence from two models of “true” lapses that are both evidence  
376 independent — motor error or exploration, which does not interact with the accumulator, and inat-  
377 tention, which may still depend on its initial state. While we find that the behavior of our rats is  
378 best described by a mixture of apparent lapses and the latter two true lapse variants, it is primarily  
379 the apparent lapses (rather than either true lapse variant) that captures the links between the sub-  
380 optimality i.e. the history-dependent comodulations in psychometric thresholds and lapse rates.

---

<sup>3</sup>Note that this is the case if the agent assumes that a shift in the prior over stimulus categories maps onto an overall shift in the prior over stimulus difficulties — see (Drugowitsch, Mendonça, et al., 2019) for a detailed treatment

381 A previous study proposed an evidence-dependent model of true lapses, uncertainty-guided explo-  
382 ration (Pisupati et al., 2021), in order to account for the scaling of lapse rates with sensory noise.  
383 Although we don't explicitly consider this model, our model of apparent lapses already displays  
384 this property, with higher levels of sensory noise leading to more frequent apparent lapses.

385 Our model predicts that an increased reliance on history (i.e., larger shifts of the initial states)  
386 should produce more apparent lapses. Indeed, this could provide an explanation that links disparate  
387 sets of observations from previous studies: while some studies have reported that perturbations  
388 of secondary motor cortex and striatum give rise to higher lapse rates (Erllich, Brunton, et al.,  
389 2015; Yartsev et al., 2018; Guo et al., 2019; Sindreu et al., 2021; Pisupati et al., 2021), others  
390 have shown that the effects of perturbing these regions seems to resemble an increased history-  
391 dependence (Sindreu et al., 2021; Luo et al., 2021). Interpreting these results through the lens  
392 of our model, we'd conclude that these regions play a crucial role in the interaction of history-  
393 dependent initial states with sensory evidence. Indeed, Luo et al. 2021 find that this increased  
394 history dependence upon perturbation is mediated by increased bias in initial value of the neurally  
395 derived accumulator variable. Our model could also help explain why Busse et al. 2011 found  
396 that mice with higher lapse probabilities showed higher history dependence, or results from The  
397 International Brain Laboratory et al. 2021 who observed a modulation in lapse rates in addition to  
398 horizontal biases upon explicit manipulation of category priors.

399 One interesting future line of investigation is to probe the precise nature of the model of non-  
400 stationarity over priors assumed by animals in such tasks. The range of parameter values inferred  
401 using our flexible formulation could offer a useful starting point for this line of investigation. For  
402 instance, Dynamic Belief Models (Yu and Cohen, 2009; Ryali et al., 2018), a popular class of  
403 generative models over priors, correspond to a narrowly constrained set of parameter values in our  
404 model. Such an understanding would not only afford more reliable control of behavior and more  
405 accurate interpretation of neural correlates in stationary tasks, but could also yield insight into the  
406 inductive biases that allow animals to learn quickly and efficiently in non-stationary, naturalistic  
407 settings.

## 408 **Methods**

409 **Subjects** Animal use procedures were approved by the Princeton University Institutional Animal  
410 Care and Use Committee (IACUC #1853). All subjects were adult male Long Evans rats, typically

411 housed in pairs. Rats that trained during the day were housed in a reverse light cycle room. Rats  
 412 had free access to food but in order to to motivate them to work for water reward, they were placed  
 413 on a controlled water schedule: 2-4 hours per day during task training, usually 7 days a week and  
 414 between 0 and 1 hour ad lib following training.

**Drift diffusion model of decision-making** We use a standard formulation of sequential decision-making based on (Bogacz et al., 2006; Drugowitsch, Moreno-Bote, et al., 2012), in which an agent is faced with a stream of noisy sensory evidence  $\epsilon_{1:t}$  coming from one of two hypotheses  $H_1$  and  $H_2$ . The agent has to decide between sampling for longer or choosing one of two actions  $L, R$  (reaction time regime) or has to choose one of two actions after a fixed amount of evidence (fixed duration regime). Such a problem can be formulated as one of finding an optimal policy  $\pi_t$  in a partially-observable markov decision process (Rao, 2010; Drugowitsch, Moreno-Bote, et al., 2012), whose solution can be written as a pair of thresholds on the log-posterior ratio  $\log(\frac{g(t)}{1-g(t)})$ , where  $g(t) = p(H_1|\epsilon_{1:t})$ :

$$\pi_t = \begin{cases} \text{choose L,} & -B \geq \log(\frac{g(t)}{1-g(t)}) \\ \text{sample,} & -B < \log(\frac{g(t)}{1-g(t)}) < B \\ \text{choose R,} & \log(\frac{g(t)}{1-g(t)}) \geq B \end{cases}$$

The log posterior ratio can be further broken down into a sum of log prior ratios and log likelihood ratios, using Bayes rule:

$$\log \frac{p(H_1|\epsilon_{1:t})}{p(H_2|\epsilon_{1:t})} = \log \frac{p(H_1)}{p(H_2)} + \log \frac{p(\epsilon_{1:t}|H_1)}{p(\epsilon_{1:t}|H_2)}$$

The optimal policy can equivalently be expressed in terms of the prior and sum of momentary sensory evidence  $x(t) = \sum_t \epsilon_t$ , which are sufficient statistics of the posterior (Drugowitsch, Moreno-Bote, et al., 2012; Piet et al., 2018). In the continuous time limit, when the average rate of evidence increments or drift rate is  $\mu$ , and the standard deviation of sensory noise is  $\sigma$ , this corresponds to a drift diffusion model that terminates when it reaches one of two bounds (Bogacz et al., 2006) and whose initial state  $I$  is proportional to the log prior ratio:

$$dx = \mu dt + \sigma dW, \quad x(0) = I = k \cdot \log \frac{p(H_1)}{p(H_2)}$$

In this case, the probability of choosing rightward actions, i.e. hitting the upper bound can be written analytically as follows (derived from Palmer et al. 2005):

$$P(B+) = \frac{1 - e^{-2\mu(B+I)/\sigma^2}}{1 - e^{-4\mu B/\sigma^2}}$$

415 In cases where trial difficulties (and hence drift rates) vary from trial to trial the optimal  
416 policy includes time-dependent, collapsing bounds on the posterior. However, under certain cir-  
417 cumstances, constant bounds on  $X_t = \sum_t \epsilon_t$  implement close-to-optimal collapsing bounds on the  
418 posterior (Denève, 2012; Drugowitsch, Moreno-Bote, et al., 2012), which is the regime we assume  
419 for our analysis.

**Models of initial state updating** We model initial state updating as a sum of exponential filters over past choice-outcome pairs (*Rw*: right-wins, *Lw*: left-wins, *Rl*: right-loss, *Ll*: left-loss). So the initial state  $I$  at trial  $n + 1$  is given by:

$$I(n + 1) = i^{Rw}(n + 1) + i^{Lw}(n + 1) + i^{Rl}(n + 1) + i^{Ll}(n + 1)$$

where each filter  $i^h$  decays by a factor of  $\beta^h$ , and is incremented by a factor of  $\eta^h$  following the observation of that particular choice-outcome pair, i.e

$$i^h(n + 1) = \eta^h 1^h(o_n) + \beta^h i^h(n) \quad \text{where } h = \{Rw, Lw, Rl, Ll\}$$

420  $o_n$  is the choice-outcome pair observed on trial  $n$  and  $1^h(o_n)$  is an indicator function that is 1 when  
421  $o_n = h$  and is 0 otherwise.

422 When  $\beta^h$  and  $\eta^h$  are the same  $\forall h$ , this rule reduces to an approximation of the Bayesian  
423 update for the Dynamic Belief Model (Yu and Cohen, 2009), which tracks a prior that undergoes  
424 discrete unswitched switches at a fixed rate. We compared this unconstrained model to models  
425 with various constraints on the decay and magnitude parameters (same parameters for corrects  
426 v.s. errors, left v.s. right etc). While model comparison revealed that not every rat required all  
427 parameters to be different, the unconstrained model is the most general form that best captures  
428 behavior across rats.

**Psychometric curves** Psychometric curves model the probability of a subject choosing one of the options (e.g. right) as a function of stimulus strength. We parametrize the psychometric curve as a 4-parameter logistic function:

$$P(\text{choose Right}) = \kappa_0 + \frac{\kappa_1}{1 + e^{-b(x-x_0)}}$$

429 where  $x_0$  is the threshold parameter that additively biases the stimulus  $x$ ,  $b$  measures sensitivity to  
430 the stimulus,  $\kappa_0$  is the left asymptote or left lapse rate and  $\kappa_1$  scales the logistic function. Therefore,  
431 the right asymptote is given by  $\kappa_0 + \kappa_1$  and the right lapse rate itself is given by  $1 - (\kappa_0 + \kappa_1)$ .  
432 We fit all four of these parameters  $\{\kappa_0, \kappa_1, x_0, b\}$  to choices generated by either the DDM (Fig 1),  
433 rats (Fig 2, 3, 4), or accumulator models adapted to the tasks (Fig 3, 4) using a gradient-descent  
434 algorithm (interior-point) to maximize the (Binomial) log likelihood of choices using MATLAB's  
435 constrained optimization function *fmincon*.  $\kappa_0$  and  $\kappa_1$  were both constrained to lie within the  
436 interval  $[0, 1]$ . 95% confidence intervals on these parameters were generated using bootstrapping.

437 **History modulation of psychometric parameters:** To summarize the effects of trial history  
438 on psychometric parameters we fit independent psychometric curves to choices conditioned on  
439 1-trial back choice-outcome history i.e. following rightward wins (Rw) and leftward wins (Lw).  
440 Modulation of the threshold parameter by history was then computed as  $x_0^{Rw} - x_0^{Lw}$ . To quantify the  
441 modulation of lapse rate parameter by history we first computed the difference in the left and right  
442 asymptotes following rightward and leftward wins:  $\kappa_0^{Rw} - \kappa_0^{Lw}$  and  $(\kappa_0^{Rw} + \kappa_1^{Rw}) - (\kappa_0^{Lw} + \kappa_1^{Lw})$   
443 respectively. The net modulation of lapse rates with trial history is given by the sum of these  
444 differences:  $2(\kappa_0^{Rw} - \kappa_0^{Lw}) + (\kappa_1^{Rw} - \kappa_1^{Lw})$ .

## 445 Behavioral tasks

446 **Auditory evidence accumulation task:** Rats were trained with previously established protocol  
447 (Brunton et al., 2013; Hanks, Kopec, et al., 2015; Erlich, Brunton, et al., 2015; Yartsev et al.,  
448 2018) using the BControl system. Briefly, rats were put in an operant chamber with three nose  
449 ports. They were trained to begin a trial by poking their nose into the middle port. This initiated  
450 two simultaneous streams of randomly-timed discrete auditory clicks for a predetermined duration  
451 after a variable delay (0.5-1.3s), one from a speaker to their left and the other to their right. Rats  
452 were required to maintain “fixation” throughout the entire stimulus (1.5s), failure to do so led to

453 a violation trial. At the end of the stimulus, rats had to poke towards the side which played the  
454 greater number of clicks to obtain a water reward. Stimulus difficulty was varied from trial-to-trial  
455 by changing the ratio of the generative Poisson rates of the two click streams. Trial difficulty and  
456 rewarded side were independently sampled on each trial.

457 We analyzed rats which performed greater than 30,000 trials, at 70% or more accuracy.  
458 Sessions with less than 300 trials or less than 60% accuracy for either of the choices were excluded.  
459 Since rats typically perform this task for many months after having passed the final training stage,  
460 to minimize nonstationarities in the data (due to break in training because of holiday closures etc.)  
461 and ensure that we are analyzing asymptotic performance, we identified temporally contiguous  
462 sessions with stable accuracy by performing change-point detection on smoothed trial hit rate  
463 using MATLAB's *findchangepts* function. The partition with most number of trials was included  
464 in the analysis. Since the animals neither made a choice nor received an outcome on violation  
465 trials, we ignore them while computing trial-history effects.

466 **Auditory evidence accumulation task with reaction time reports** To measure rats' reaction  
467 times in addition to choices we modified the auditory evidence accumulation task in two ways.  
468 First, we relaxed the "fixation" requirement and instead allowed rats to sample the stimulus for  
469 as long as they want. As soon as rats broke fixation by removing their nose from the center port,  
470 the stimulus stopped and the rats were required to report their decision by poking into one of the  
471 side ports. For any given trial, the time that the rat spent sampling the stimulus was its reaction  
472 time. Second, we rewarded rats if they correctly reported the side which had greater underlying  
473 Poisson rate rather than the side which played the greater number of clicks. This helped eliminate  
474 the trivial strategy of culminating a decision after the first click and having perfect accuracy by  
475 simply reporting the side of that click without any need for evidence accumulation.

476 In practice, we followed the same training protocol as the interrogation task (Brunton et al.,  
477 2013) but with the modified reward rule. Once the rats were fully trained on the interrogation  
478 protocol we gradually reduced the duration of delay between stimulus onset and trial initiation as  
479 well as the fixation period. Most rats maintained high accuracy (>70%) upon this manipulation,  
480 if rats performance did not meet this criterion even after a week of training, they were excluded.  
481 Rats tended to have worse accuracy early in the session, so we omitted the first 50 trials from our  
482 analysis. After the first 50 trials, we confirmed that the accuracy in the first and second halves of



483 the session was comparable.

#### 484 **Data modeling methods**

**Accumulator model** To model subjects choices and RTs, we used accumulation to bound model modified to take into account the discrete nature of evidence in our behavioral tasks (Brunton et al., 2013). In the model, the evolution of accumulated evidence  $a(t)$  in response to the left ( $\delta_L$ ) and right ( $\delta_R$ ) click trains on trial  $n$  is given by:

$$da = \begin{cases} 0, & \text{if } |a| \geq B \\ \lambda a dt + (\delta_{R,t} C_R(t) \eta_R - \delta_{L,t} C_L(t) \eta_L) dt + \sigma_a dW & \text{otherwise} \end{cases}$$

where  $\frac{dC}{dt} = \frac{1-C}{\tau_\phi} + (\phi-1)C(\delta_{R,t} + \delta_{L,t})$  and  $a(t=0) = I(n)$

485 where  $\lambda$  is the inverse time constant of the consistent drift in memory of  $a(t)$ .  $C_R(t)$  and  $C_L(t)$  are  
 486 the magnitudes of each right and left click respectively after undergoing sensory adaptation (with  
 487 adaptation strength  $\phi$  and adaptation time constant  $\tau_\phi$ ). The sensory noise that accompanies each  
 488 click is represented by  $\eta_R, \eta_L$  which are Gaussian random variables with mean 1 and variance  $\sigma_s^2$ .  
 489 The accumulation variable  $a$  also undergoes Brownian diffusion through the addition of a Wiener  
 490 process ( $W$ ) with variance  $\sigma_a^2$ .  $B$  represents the absorbing decision bound that prevents  $a(t)$  from  
 491 evolving further, if crossed. The initial value of the accumulator variable  $a$  varies from trial-to-trial  
 492 and is set based on exponentially filtered history of previous choices and outcomes (see Methods  
 493 section on Models of initial state updating). A choice is made by comparing the final value of the  
 494 accumulator  $a(T)$  to a side bias. A rightward choice is made if  $a(T) > \text{bias}$ .

Since the model quantifies noise sources on each trial, it requires estimating the evolution of a noise-induced probability distribution  $P(a(t))$ . We compute  $P(a(t))$  by solving the Fokker-Planck equations that correspond to model dynamics (see Brunton et al. 2013; DePasquale et al. 2021 for numerical methods). The probability of making a rightward choice at the end time-point  $T$  of a trial, given accumulation model parameters  $\theta^{acc}$  is:

$$P(\text{choose R} | \delta_R, \delta_L, \theta^{acc}) = \int_{a=\text{bias}}^{\infty} da P(a(T) | \delta_R, \delta_L, \theta^{acc})$$

495 **Models of true lapses** We assume that some fraction of choices  $\kappa$  arise from processes extra-  
 496 neous to evidence accumulation such as motor error/exploration or inattention. We parameterize  
 497 these processes with  $\theta^{lapse}$  and refer to them as “true lapses”:

- 498 • In the motor error/exploration variant, the probability of making a choice towards the right -  
 499 when lapsing - is given by  $\rho$ .

$$P(\text{choose R}|\theta^{lapse}) = \rho$$

- In the inattention variant (Supp Fig 4C), the subject lapses towards the side favored by the initial state relative to a bias  $\rho$ . So the probability of a rightward choice due to inattention on trial  $n$  is:

$$P(\text{choose R}|\theta^{lapse}) = \begin{cases} 1 & \text{if } i(n) - \rho > 0 \\ 0.5 & \text{if } i(n) - \rho = 0 \\ 0 & \text{if } i(n) - \rho < 0 \end{cases}$$

- 500 • In the hybrid variant (with motor error and inattention; Supp Fig 6), the probability of lapsing  
 501 towards right depends on the initial state through a sigmoidal function whose slope  $m$  (or  
 502 matching constant) as well as bias  $\rho$  is a free parameter:

$$P(\text{choose R}|\theta^{lapse}) = \frac{1}{1 + e^{-m(i(n)-\rho)}}$$

503 Hence the total probability of making a rightward choice due to accumulation and true lapses is:

$$P(\text{choose R}|\Theta) = (1 - \kappa)P(\text{choose R}|\delta_R, \delta_L, \theta^{acc}) + \kappa P(\text{choose R}|\theta^{lapse})$$

504 where  $\Theta = \{\theta^{acc}, \theta^{lapse}, \kappa\}$ .

**Model fitting** The model parameters were fit to individual rats by maximizing the log likelihood of the observed choices of the rat  $\mathbf{c}_{\text{obs}}$ , i.e. by maximizing

$$\ln \mathcal{L}(\mathbf{c}_{\text{obs}}|\delta_{\mathbf{R}}, \delta_{\mathbf{L}}, \Theta) = \sum_n \ln P(c_{\text{obs},n}|\delta_{R,n}, \delta_{L,n}, \Theta)$$

505 where  $n$  indexes trials. Constrained optimization was performed in Julia using Optim package. We  
506 computed gradients for parameter optimization using a forward-mode automatic differentiation  
507 package. The reported maximum likelihood parameters and likelihood values (used for model  
508 comparison) are from model fits to the entire dataset. We fit a random subset of 10 rats using 5-fold  
509 cross-validation (85% training dataset, 15% test dataset) but this yielded very similar maximum  
510 likelihood parameters and virtually identical test and training log-likelihoods. Hence, to save on  
511 computing time we fit the different model variants to each rat's entire dataset. This agreement  
512 between test and training likelihoods is likely due to the large number of trials in our dataset and  
513 the modest number of parameters in our model.

**Simultaneous modeling of choices and RTs** In decision-making tasks, observed reaction times (RTs) are often thought of as comprised of stimulus sampling or decision times (DTs, the time it takes for the subject's accumulated evidence to hit the bound) and non-decision related processing times (NDTs). In our datasets we observed that reaction times tended to be slower following incorrect trials and that they grew longer over the course of a session. These effects could be isolated just to RTs and were not observed in choice behavior. To model these trends we conceptualize non-decision times as arising from a separate drift diffusion process whose drift  $\nu$  is additionally modulated by current trial number  $n$  and previous trial's outcome. These non-decision time drift-diffusion processes terminate when the bound  $\omega$  is hit. We assume that the non-decision times for each choice  $k \in \{L, R\}$  have independent bounds ( $\omega_k$ ) and drifts ( $\nu_k$ ). So the non-decision times for a trial  $n$  are samples from the following Wald or Inverse Gaussian (*IG*) distribution:

$$\tau_n^{NDT} \sim IG\left(\frac{\omega_k}{\nu_k - \alpha n + \gamma_o 1_{(n-1)}^-}, \omega_k^2\right)$$

514 where  $k \in \{L, R\}$  and  $1_{(n-1)}^-$  is an indicator function which is 1 if the previous trial was incorrect  
515 and is 0 otherwise.  $\alpha$  parameterizes the impact of trial number on NDTs and  $\gamma_o$  parameterizes the  
516 impact of previous trial's outcome on current trial's NDT.

We fit the model by maximizing the joint log likelihood of the observed choices and RTs. For any given trial, we can compute the likelihood of observing a particular reaction time  $RT_{obs}$  and choice  $c_{obs}$  due to accumulation by marginalizing over possible decision or bound hitting times

$\tau_{c_{obs}}$  for the observed choice:

$$P(c_{obs}, RT_{obs} | \delta_R, \delta_L, \theta^{acc}, \theta^{NDT}) = \int_0^{RT_{obs}} P(\tau_{c_{obs}} | \delta_R, \delta_L, \theta^{acc}) P(c_{obs}, RT_{obs} | \theta^{NDT}, \tau_{c_{obs}}) d\tau_{c_{obs}}$$

On true lapse trials, RTs were assumed to arise from NDTs alone and therefore the joint likelihood due to accumulation and true lapses is given by:

$$\mathcal{L}(c_{obs}, RT_{obs} | \delta_R, \delta_L, \Theta) = (1 - \kappa) P(c_{obs}, RT_{obs} | \delta_R, \delta_L, \theta^{acc}, \theta^{NDT}) + \kappa P(c_{obs}, RT_{obs} | \theta^{lapse}, \theta^{NDT})$$

517 where  $\Theta = \{\theta^{acc}, \theta^{NDT}, \theta^{lapse}, \kappa\}$ .

## 518 Acknowledgements

519 We thank members of the Brody lab for experimental support and helpful feedback throughout the  
520 project especially Adrian Bondy, Thomas Luo, Emily Dennis, Tyler Boyd-Meredith, and Ahmed  
521 El-Hady. We also thank Jovanna Teran and Brody lab technicians for assistance with rat training.  
522 We are grateful to Sashank Pisupati, Jonathan Cohen, Sebastian Musslick, Jonathan Pillow, and  
523 Ilana Witten for helpful discussions at various points during the project. This work was supported  
524 by NIH grant R01MH108358 to CDB.

## 525 References

- 526 [1] R. Y. Cho et al., “Mechanisms underlying dependencies of performance on stimulus his-  
527 tory in a two-alternative forced-choice task,” en, *Cognitive, Affective, & Behavioral Neuro-*  
528 *science*, vol. 2, no. 4, pp. 283–299, 2002. DOI: 10.3758/CABN.2.4.283.
- 529 [2] A. Akrami, C. D. Kopec, M. E. Diamond, and C. D. Brody, “Posterior parietal cortex repre-  
530 sents sensory history and mediates its effects on behaviour,” en, *Nature*, vol. 554, no. 7692,  
531 pp. 368–372, 2018, Number: 7692 Publisher: Nature Publishing Group. DOI: 10.1038/  
532 nature25510.
- 533 [3] L. Pinto et al., “An Accumulation-of-Evidence Task Using Visual Pulses for Mice Nav-  
534 igating in Virtual Reality,” English, *Frontiers in Behavioral Neuroscience*, vol. 12, 2018,  
535 Publisher: Frontiers. DOI: 10.3389/fnbeh.2018.00036.

- 536 [4] A. E. Urai, J. W. de Gee, K. Tsetsos, and T. H. Donner, “Choice history biases subsequent  
537 evidence accumulation,” en, *eLife*, vol. 8, e46331, 2019. DOI: 10.7554/eLife.46331.
- 538 [5] A. Hermoso-Mendizabal, A. Hyafil, P. E. Rueda-Orozco, S. Jaramillo, D. Robbe, and J. de  
539 la Rocha, “Response outcomes gate the impact of expectations on perceptual decisions,”  
540 en, *Nature Communications*, vol. 11, no. 1, p. 1057, 2020, Number: 1 Publisher: Nature  
541 Publishing Group. DOI: 10.1038/s41467-020-14824-w.
- 542 [6] A. G. Mendonça, J. Drugowitsch, M. I. Vicente, E. E. J. DeWitt, A. Pouget, and Z. F.  
543 Mainen, “The impact of learning on perceptual decisions and its implication for speed-  
544 accuracy tradeoffs,” en, *Nature Communications*, vol. 11, no. 1, p. 2757, 2020, Number: 1  
545 Publisher: Nature Publishing Group. DOI: 10.1038/s41467-020-16196-7.
- 546 [7] A. Lak et al., “Reinforcement biases subsequent perceptual decisions when confidence is  
547 low, a widespread behavioral phenomenon,” *eLife*, vol. 9, E. Salinas, M. J. Frank, E. Salinas,  
548 C. D. Brody, and L. Ding, Eds., e49834, 2020, Publisher: eLife Sciences Publications, Ltd.  
549 DOI: 10.7554/eLife.49834.
- 550 [8] G. Mochol, R. Kiani, and R. Moreno-Bote, “Prefrontal cortex represents heuristics that  
551 shape choice bias and its integration into future behavior,” en, *Current Biology*, vol. 31,  
552 no. 6, 1234–1244.e6, 2021. DOI: 10.1016/j.cub.2021.01.068.
- 553 [9] N. A. Roy, J. H. Bak, A. Akrami, C. D. Brody, and J. W. Pillow, “Extracting the dynamics of  
554 behavior in sensory decision-making experiments,” en, *Neuron*, vol. 109, no. 4, 597–610.e6,  
555 2021. DOI: 10.1016/j.neuron.2020.12.004.
- 556 [10] The International Brain Laboratory et al., “Standardized and reproducible measurement of  
557 decision-making in mice,” *eLife*, vol. 10, N. Uchida and M. J. Frank, Eds., e63711, 2021,  
558 Publisher: eLife Sciences Publications, Ltd. DOI: 10.7554/eLife.63711.
- 559 [11] J. I. Gold, C.-T. Law, P. Connolly, and S. Bennur, “The Relative Influences of Priors and Sen-  
560 sory Evidence on an Oculomotor Decision Variable During Perceptual Learning,” *Journal of*  
561 *Neurophysiology*, vol. 100, no. 5, pp. 2653–2668, 2008, Publisher: American Physiological  
562 Society. DOI: 10.1152/jn.90629.2008.
- 563 [12] L. Busse et al., “The Detection of Visual Contrast in the Behaving Mouse,” en, *Journal of*  
564 *Neuroscience*, vol. 31, no. 31, pp. 11 351–11 361, 2011, Publisher: Society for Neuroscience  
565 Section: Articles. DOI: 10.1523/JNEUROSCI.6689-10.2011.

- 566 [13] M. Carandini and A. K. Churchland, “Probing perceptual decisions in rodents,” en, *Nature*  
567 *Neuroscience*, vol. 16, no. 7, pp. 824–831, 2013, Number: 7 Publisher: Nature Publishing  
568 Group. DOI: 10.1038/nn.3410.
- 569 [14] S. Zhang, H. C. Huang, and A. J. Yu, “Sequential effects: A Bayesian analysis of prior  
570 bias on reaction time and behavioral choice,” en, *Proceedings of the Annual Meeting of the*  
571 *Cognitive Science Society*, vol. 36, no. 36, 2014.
- 572 [15] I. Fründ, F. A. Wichmann, and J. H. Macke, “Quantifying the effect of intertrial dependence  
573 on perceptual decisions,” *J Vis*, vol. 14, no. 7, pp. 9–9, 2014.
- 574 [16] B. B. Scott, C. M. Constantinople, J. C. Erlich, D. W. Tank, and C. D. Brody, “Sources of  
575 noise during accumulation of evidence in unrestrained and voluntarily head-restrained rats,”  
576 *eLife*, vol. 4, N. Uchida, Ed., e11308, 2015, Publisher: eLife Sciences Publications, Ltd.  
577 DOI: 10.7554/eLife.11308.
- 578 [17] A. Abrahamyan, L. L. Silva, S. C. Dakin, M. Carandini, and J. L. Gardner, “Adaptable  
579 history biases in human perceptual decisions,” en, *Proceedings of the National Academy of*  
580 *Sciences*, vol. 113, no. 25, E3548–E3557, 2016, Publisher: National Academy of Sciences  
581 Section: PNAS Plus. DOI: 10.1073/pnas.1518786113.
- 582 [18] O. Odoemene, S. Pisupati, H. Nguyen, and A. K. Churchland, “Visual Evidence Accumula-  
583 tion Guides Decision-Making in Unrestrained Mice,” en, *Journal of Neuroscience*, vol. 38,  
584 no. 47, pp. 10 143–10 155, 2018, Publisher: Society for Neuroscience Section: Research  
585 Articles. DOI: 10.1523/JNEUROSCI.3478-17.2018.
- 586 [19] A. J. Yu and J. D. Cohen, “Sequential effects: Superstition or rational behavior?” In *Ad-*  
587 *vances in Neural Information Processing Systems 21*, D. Koller, D. Schuurmans, Y. Bengio,  
588 and L. Bottou, Eds., Curran Associates, Inc., 2009, pp. 1873–1880.
- 589 [20] M. Molano-Mazon, D. Duque, G. R. Yang, and J. d. l. Rocha, “Pre-training RNNs on eco-  
590 logically relevant tasks explains sub-optimal behavioral reset,” en, Tech. Rep., 2021, Com-  
591 pany: Cold Spring Harbor Laboratory Distributor: Cold Spring Harbor Laboratory Label:  
592 Cold Spring Harbor Laboratory Section: New Results Type: article, p. 2021.05.15.444287.  
593 DOI: 10.1101/2021.05.15.444287.
- 594 [21] J. A. Nevin, “Signal detection theory and operant behavior,” *Journal of the Experimental*  
595 *Analysis of Behavior*, vol. 12, no. 3, pp. 475–480, 1969. DOI: 10.1901/jeab.1969.  
596 12-475.

- 597 [22] R. Ratcliff and J. N. Rouder, “Modeling Response Times for Two-Choice Decisions,” *Psy-*  
598 *chological Science*, vol. 9, no. 5, pp. 347–356, 1998, Publisher: SAGE Publications Inc.  
599 DOI: 10.1111/1467-9280.00067.
- 600 [23] R. Bogacz, E. Brown, J. Moehlis, P. Holmes, and J. D. Cohen, “The physics of optimal  
601 decision making: A formal analysis of models of performance in two-alternative forced-  
602 choice tasks,” *Psychological Review*, vol. 113, no. 4, pp. 700–765, 2006, Place: US Pub-  
603 lisher: American Psychological Association. DOI: 10.1037/0033-295X.113.4.700.
- 604 [24] S. Goldfarb, K. Wong-Lin, M. Schwemmer, N. E. Leonard, and P. Holmes, “Can Post-Error  
605 Dynamics Explain Sequential Reaction Time Patterns?” *Frontiers in Psychology*, vol. 3,  
606 2012. DOI: 10.3389/fpsyg.2012.00213.
- 607 [25] J. L. Gardner, “Optimality and heuristics in perceptual neuroscience,” en, *Nature Neuro-*  
608 *science*, vol. 22, no. 4, pp. 514–523, 2019, Number: 4 Publisher: Nature Publishing Group.  
609 DOI: 10.1038/s41593-019-0340-4.
- 610 [26] F. A. Wichmann and N. J. Hill, “The psychometric function: I. Fitting, sampling, and good-  
611 ness of fit,” en, *Perception & Psychophysics*, vol. 63, no. 8, pp. 1293–1313, 2001. DOI:  
612 10.3758/BF03194544.
- 613 [27] S. Shushruth and M. N. Shadlen, “Sequential sampling from memory underlies action selec-  
614 tion during abstract decision making,” en, Tech. Rep., 2021, Company: Cold Spring Harbor  
615 Laboratory Distributor: Cold Spring Harbor Laboratory Label: Cold Spring Harbor Labo-  
616 ratory Section: New Results Type: article, p. 2021.04.30.442176. DOI: 10.1101/2021.  
617 04.30.442176.
- 618 [28] C.-T. Law and J. I. Gold, “Reinforcement learning can account for associative and perceptual  
619 learning on a visual-decision task,” en, *Nature Neuroscience*, vol. 12, no. 5, pp. 655–663,  
620 2009, Number: 5 Publisher: Nature Publishing Group. DOI: 10.1038/nn.2304.
- 621 [29] J. I. Gold and L. Ding, “How mechanisms of perceptual decision-making affect the psy-  
622 chometric function,” en, *Progress in Neurobiology*, Conversion of Sensory Signals into  
623 Perceptions, Memories and Decisions, vol. 103, pp. 98–114, 2013. DOI: 10.1016/j.  
624 pneurobio.2012.05.008.
- 625 [30] B. W. Brunton, M. M. Botvinick, and C. D. Brody, “Rats and Humans Can Optimally Ac-  
626 cumulate Evidence for Decision-Making,” en, *Science*, vol. 340, no. 6128, pp. 95–98, 2013,  
627 Publisher: American Association for the Advancement of Science Section: Report. DOI:  
628 10.1126/science.1233912.

- 629 [31] H.-T. Wang et al., “Finding the needle in high-dimensional haystack: A tutorial on canonical  
630 correlation analysis,” *arXiv:1812.02598 [cs, stat]*, 2018, arXiv: 1812.02598.
- 631 [32] S. Pisupati, L. Chartarifsky-Lynn, A. Khanal, and A. K. Churchland, “Lapses in perceptual  
632 decisions reflect exploration,” *eLife*, vol. 10, D. Lee, Ed., e55490, 2021, Publisher: eLife  
633 Sciences Publications, Ltd. DOI: 10.7554/eLife.55490.
- 634 [33] Z. C. Ashwood et al., “Mice alternate between discrete strategies during perceptual decision-  
635 making,” en, *Nature Neuroscience*, vol. 25, no. 2, pp. 201–212, 2022, Number: 2 Publisher:  
636 Nature Publishing Group. DOI: 10.1038/s41593-021-01007-z.
- 637 [34] J. C. Erlich, M. Bialek, and C. D. Brody, “A Cortical Substrate for Memory-Guided Orient-  
638 ing in the Rat,” *Neuron*, vol. 72, no. 2, pp. 330–343, 2011.
- 639 [35] J. C. Erlich, B. W. Brunton, C. A. Duan, T. D. Hanks, and C. D. Brody, “Distinct effects of  
640 prefrontal and parietal cortex inactivations on an accumulation of evidence task in the rat,”  
641 *eLife*, vol. 4, p. 8166, 2015.
- 642 [36] M. M. Yartsev, T. D. Hanks, A. M. Yoon, and C. D. Brody, “Causal contribution and dynam-  
643 ical encoding in the striatum during evidence accumulation,” *eLife*, vol. 7, J. I. Gold, T. E.  
644 Behrens, J. I. Gold, and M. D. Humphries, Eds., e34929, 2018, Publisher: eLife Sciences  
645 Publications, Ltd. DOI: 10.7554/eLife.34929.
- 646 [37] L. Guo, J. T. Weems, W. I. Walker, A. Levichev, and S. Jaramillo, “Choice-Selective Neu-  
647 rons in the Auditory Cortex and in Its Striatal Target Encode Reward Expectation,” en,  
648 *Journal of Neuroscience*, vol. 39, no. 19, pp. 3687–3697, 2019, Publisher: Society for Neu-  
649 roscience Section: Research Articles. DOI: 10.1523/JNEUROSCI.2585-18.2019.
- 650 [38] C. Sindreu, R. Marín, Y. Fuentealba, D. Duque, L. Jimenez, and J. de la Rocha, “The causal  
651 role of the striatum in the encoding of task-adaptive expectation-based choice biases.,” in  
652 *Computational and Systems Neuroscience (Cosyne) 2021*, 2021, p. 117.
- 653 [39] M. J. Siniscalchi, H. Wang, and A. C. Kwan, “Enhanced Population Coding for Rewarded  
654 Choices in the Medial Frontal Cortex of the Mouse,” *Cerebral Cortex*, vol. 29, no. 10,  
655 pp. 4090–4106, 2019. DOI: 10.1093/cercor/bhy292.
- 656 [40] J. I. Gold and M. N. Shadlen, “The Neural Basis of Decision Making,” *Annual Review of*  
657 *Neuroscience*, vol. 30, no. 1, pp. 535–574, 2007, eprint: <https://doi.org/10.1146/annurev.neuro.29.051605.113038>.  
658 DOI: 10.1146/annurev.neuro.29.051605.113038.



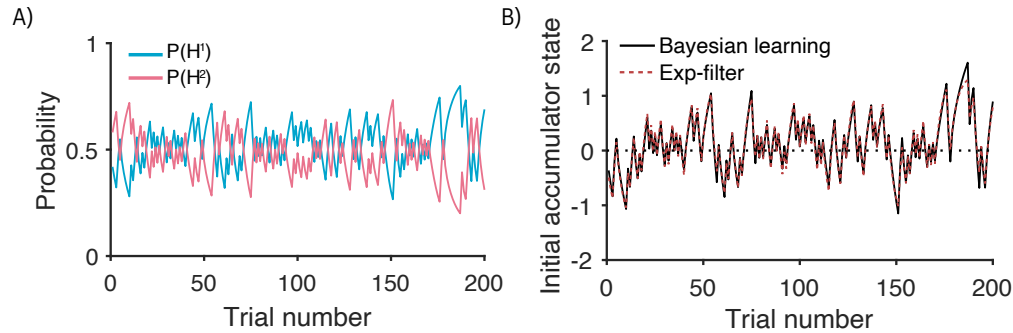
- 659 [41] J. Drugowitsch, R. Moreno-Bote, A. K. Churchland, M. N. Shadlen, and A. Pouget, “The  
660 Cost of Accumulating Evidence in Perceptual Decision Making,” en, *Journal of Neuro-*  
661 *science*, vol. 32, no. 11, pp. 3612–3628, 2012. DOI: 10.1523/JNEUROSCI.4010-  
662 11.2012.
- 663 [42] P. Dayan and N. D. Daw, “Decision theory, reinforcement learning, and the brain,” en,  
664 *Cognitive, Affective, & Behavioral Neuroscience*, vol. 8, no. 4, pp. 429–453, 2008. DOI:  
665 10.3758/CABN.8.4.429.
- 666 [43] J. Drugowitsch, A. G. Mendonça, Z. F. Mainen, and A. Pouget, “Learning optimal decisions  
667 with confidence,” en, *Proceedings of the National Academy of Sciences*, vol. 116, no. 49,  
668 pp. 24 872–24 880, 2019, Publisher: National Academy of Sciences Section: PNAS Plus.  
669 DOI: 10.1073/pnas.1906787116.
- 670 [44] J. Palmer, A. C. Huk, and M. N. Shadlen, “The effect of stimulus strength on the speed  
671 and accuracy of a perceptual decision,” en, *Journal of Vision*, vol. 5, no. 5, pp. 1–1, 2005,  
672 Publisher: The Association for Research in Vision and Ophthalmology. DOI: 10.1167/  
673 5.5.1.
- 674 [45] S. Shen and W. J. Ma, “Variable precision in visual perception.,” en, *Psychological Review*,  
675 vol. 126, no. 1, pp. 89–132, 2019. DOI: 10.1037/rev0000128.
- 676 [46] K. P. Nguyen, K. Josić, and Z. P. Kilpatrick, “Optimizing sequential decisions in the drift–diffusion  
677 model,” en, *Journal of Mathematical Psychology*, vol. 88, pp. 32–47, 2019. DOI: 10.  
678 1016/j.jmp.2018.11.001.
- 679 [47] S. S. Bolkan et al., “Opponent control of behavior by dorsomedial striatal pathways depends  
680 on task demands and internal state,” en, *Nature Neuroscience*, vol. 25, no. 3, pp. 345–357,  
681 2022, Number: 3 Publisher: Nature Publishing Group. DOI: 10.1038/s41593-022-  
682 01021-9.
- 683 [48] C. Summerfield and E. Koechlin, “Economic Value Biases Uncertain Perceptual Choices  
684 in the Parietal and Prefrontal Cortices,” English, *Frontiers in Human Neuroscience*, vol. 4,  
685 2010, Publisher: Frontiers. DOI: 10.3389/fnhum.2010.00208.
- 686 [49] M. J. Mulder, E.-J. Wagenmakers, R. Ratcliff, W. Boekel, and B. U. Forstmann, “Bias in the  
687 Brain: A Diffusion Model Analysis of Prior Probability and Potential Payoff,” en, *Journal*  
688 *of Neuroscience*, vol. 32, no. 7, pp. 2335–2343, 2012, Publisher: Society for Neuroscience  
689 Section: Articles. DOI: 10.1523/JNEUROSCI.4156-11.2012.

- 690 [50] G. Gigerenzer and W. Gaissmaier, “Heuristic Decision Making,” *Annual Review of Psy-*  
691 *chology*, vol. 62, no. 1, pp. 451–482, 2011, eprint: [https://doi.org/10.1146/annurev-psych-](https://doi.org/10.1146/annurev-psych-120709-145346)  
692 [120709-145346](https://doi.org/10.1146/annurev-psych-120709-145346). DOI: [10.1146/annurev-psych-120709-145346](https://doi.org/10.1146/annurev-psych-120709-145346).
- 693 [51] P. Simen, D. Contreras, C. Buck, P. Hu, P. Holmes, and J. D. Cohen, “Reward rate optimiza-  
694 tion in two-alternative decision making: Empirical tests of theoretical predictions.,” *Journal*  
695 *of Experimental Psychology-Human Perception and Performance*, vol. 35, no. 6, pp. 1865–  
696 1897, 2009.
- 697 [52] A. E. Rorie, J. Gao, J. L. McClelland, and W. T. Newsome, “Integration of Sensory and Re-  
698 ward Information during Perceptual Decision-Making in Lateral Intraparietal Cortex (LIP)  
699 of the Macaque Monkey,” en, *PLOS ONE*, vol. 5, no. 2, e9308, 2010, Publisher: Public  
700 Library of Science. DOI: [10.1371/journal.pone.0009308](https://doi.org/10.1371/journal.pone.0009308).
- 701 [53] P. Eckhoff, P. Holmes, C. Law, P. M. Connolly, and J. I. Gold, “On diffusion processes  
702 with variable drift rates as models for decision making during learning,” en, *New Journal*  
703 *of Physics*, vol. 10, no. 1, p. 015 006, 2008, Publisher: IOP Publishing. DOI: [10.1088 /](https://doi.org/10.1088/1367-2630/10/1/015006)  
704 [1367-2630/10/1/015006](https://doi.org/10.1088/1367-2630/10/1/015006).
- 705 [54] T. D. Hanks, M. E. Mazurek, R. Kiani, E. Hopp, and M. N. Shadlen, “Elapsed Decision  
706 Time Affects the Weighting of Prior Probability in a Perceptual Decision Task,” en, *Journal*  
707 *of Neuroscience*, vol. 31, no. 17, pp. 6339–6352, 2011. DOI: [10.1523 / JNEUROSCI .](https://doi.org/10.1523/JNEUROSCI.5613-10.2011)  
708 [5613-10.2011](https://doi.org/10.1523/JNEUROSCI.5613-10.2011).
- 709 [55] Y. Fan, J. I. Gold, and L. Ding, “Ongoing, rational calibration of reward-driven perceptual  
710 biases,” *eLife*, vol. 7, P. Latham and R. B. Ivry, Eds., e36018, 2018, Publisher: eLife Sciences  
711 Publications, Ltd. DOI: [10.7554/eLife.36018](https://doi.org/10.7554/eLife.36018).
- 712 [56] J. Ditterich, “Evidence for time-variant decision making,” en, *European Journal of Neuro-*  
713 *science*, vol. 24, no. 12, pp. 3628–3641, 2006, eprint: [https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1460-](https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1460-9568.2006.05221.x)  
714 [9568.2006.05221.x](https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1460-9568.2006.05221.x). DOI: [10.1111/j.1460-9568.2006.05221.x](https://doi.org/10.1111/j.1460-9568.2006.05221.x).
- 715 [57] J. Ditterich, “Stochastic models of decisions about motion direction: Behavior and physiol-  
716 ogy,” en, *Neural Networks, Neurobiology of Decision Making*, vol. 19, no. 8, pp. 981–1012,  
717 2006. DOI: [10.1016/j.neunet.2006.05.042](https://doi.org/10.1016/j.neunet.2006.05.042).
- 718 [58] Q. N. Nguyen and P. Reinagel, “A qualitative difference in decision-making of rats vs.  
719 humans explained by quantitative differences in behavioral variability,” en, *Neuroscience*,  
720 preprint, 2020. DOI: [10.1101/2020.01.05.895268](https://doi.org/10.1101/2020.01.05.895268).

- 721 [59] J. D. Roitman and M. N. Shadlen, “Response of Neurons in the Lateral Intraparietal Area  
722 during a Combined Visual Discrimination Reaction Time Task,” en, *Journal of Neuro-*  
723 *science*, vol. 22, no. 21, pp. 9475–9489, 2002, Publisher: Society for Neuroscience Section:  
724 ARTICLE. DOI: 10.1523/JNEUROSCI.22-21-09475.2002.
- 725 [60] C. A. Shevinsky and P. Reinagel, “The Interaction Between Elapsed Time and Decision  
726 Accuracy Differs Between Humans and Rats,” *Frontiers in Neuroscience*, vol. 13, p. 1211,  
727 2019. DOI: 10.3389/fnins.2019.01211.
- 728 [61] T. Z. Luo, T. D. Hanks, D. Gupta, A. G. Bondy, and C. D. Brody, “Dorsomedial frontal  
729 cortex participates in both evidence accumulation and history-based updating,” 2021.
- 730 [62] C. Ryali, G. Reddy, and A. J. Yu, “Demystifying excessively volatile human learning: A  
731 Bayesian persistent prior and a neural approximation,” in *Advances in Neural Informa-*  
732 *tion Processing Systems 31*, S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-  
733 Bianchi, and R. Garnett, Eds., Curran Associates, Inc., 2018, pp. 2781–2790.
- 734 [63] R. P. N. Rao, “Decision Making Under Uncertainty: A Neural Model Based on Partially  
735 Observable Markov Decision Processes,” *Frontiers in Computational Neuroscience*, vol. 4,  
736 p. 146, 2010. DOI: 10.3389/fncom.2010.00146.
- 737 [64] A. T. Piet, A. El Hady, and C. D. Brody, “Rats adopt the optimal timescale for evidence  
738 integration in a dynamic environment,” en, *Nature Communications*, vol. 9, no. 1, p. 4265,  
739 2018, Number: 1 Publisher: Nature Publishing Group. DOI: 10.1038/s41467-018-  
740 06561-y.
- 741 [65] S. Denève, “Making Decisions with Unknown Sensory Reliability,” *Frontiers in Neuro-*  
742 *science*, vol. 6, 2012.
- 743 [66] T. D. Hanks, C. D. Kopec, B. W. Brunton, C. A. Duan, J. C. Erlich, and C. D. Brody,  
744 “Distinct relationships of parietal and prefrontal cortices to evidence accumulation,” *Nature*,  
745 vol. 520, no. 7546, pp. 220–223, 2015.
- 746 [67] B. DePasquale, J. W. Pillow, and C. Brody, “Neural population dynamics underlying ev-  
747 idence accumulation in multiple rat brain regions,” en, Tech. Rep., 2021, Company: Cold  
748 Spring Harbor Laboratory Distributor: Cold Spring Harbor Laboratory Label: Cold Spring  
749 Harbor Laboratory Section: New Results Type: article, p. 2021.10.28.465122. DOI: 10.  
750 1101/2021.10.28.465122.

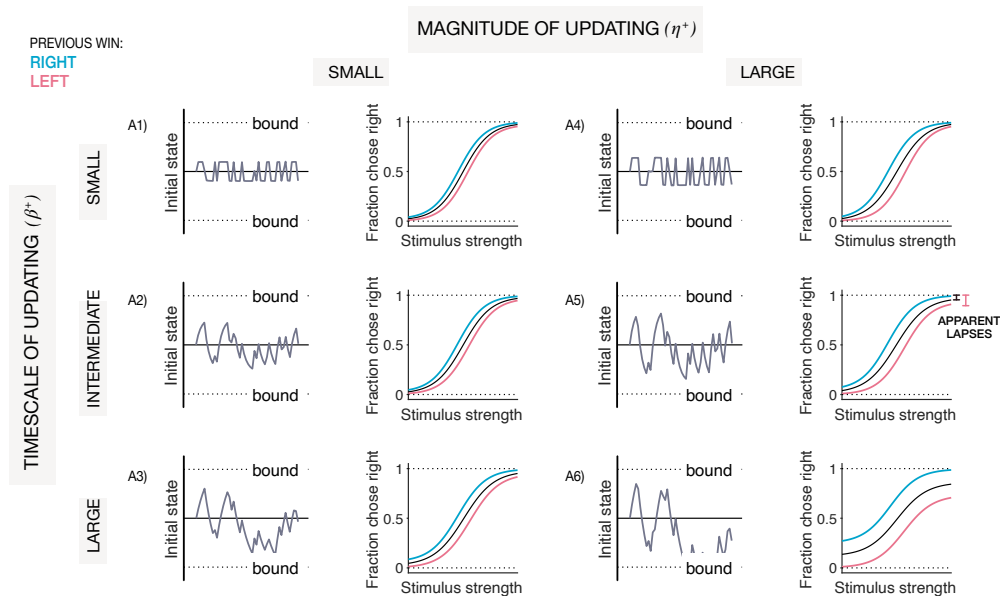
- 751 [68] P. Mamassian, M. Landy, and L. T. Maloney, “Bayesian modelling of visual perception,” in  
752 *Probabilistic models of the brain: Perception and neural function*, ser. Neural information  
753 processing series, Cambridge, MA, US: The MIT Press, 2002, pp. 13–36. DOI: 10.7551/  
754 mitpress/5583.001.0001.

755 **Supplementary materials**

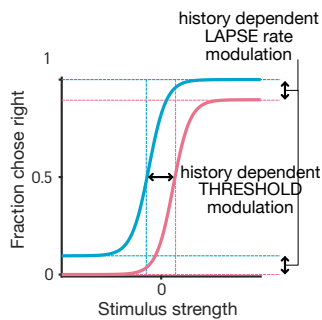


Supplementary Figure 1: **Exponential filtering for initial state setting approximates Bayesian prior updates under assumptions of non-stationarity** **A**: Example of a mis-belief in a non-stationary prior. Traces represent belief about prior probability of two hypotheses  $H^1$  and  $H^2$  inferred from a random sequence of trials drawn from a stationary symmetric prior, under the misbelief that the prior is occasionally undergoing unsignalled jumps. Such an assumed generative model is often referred to as the Dynamic Belief Model (DBM; Yu and Cohen 2009). **B**: Initial state updates corresponding exactly to the fluctuating prior beliefs in (A) that emerge from Bayesian learning (black line), plotted against approximate initial states derived from exponential filtering (dotted red line) of past choices and outcomes. The exponential filter provides a good approximation of exact Bayesian updates, while being more expressive and flexible to capture the possibility of other generative models and corresponding update rules.

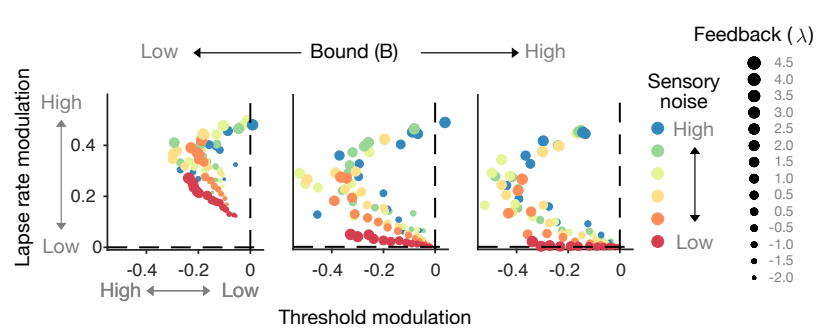
A) Influence of across-trial parameters



(B) Quantifying modulation

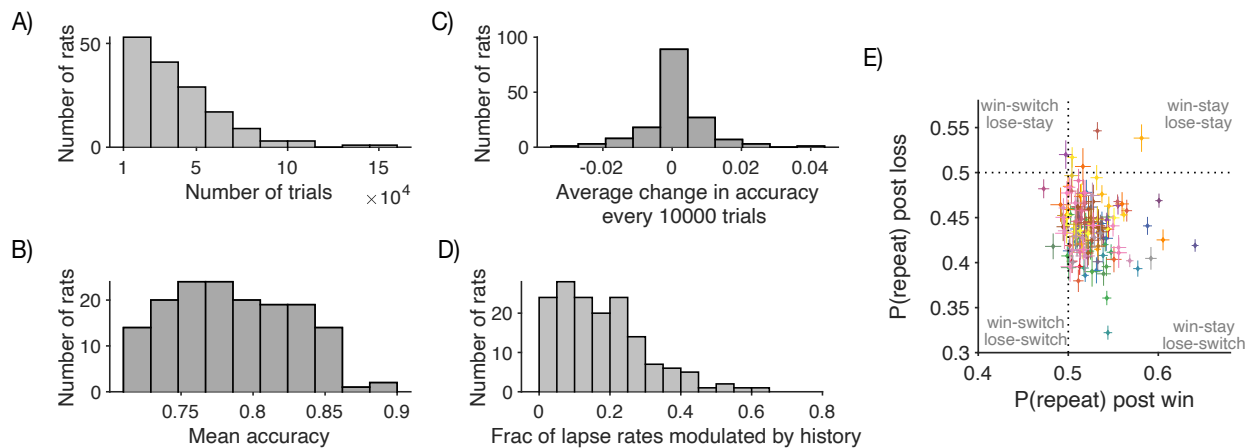


(C) Influence of within-trial parameters

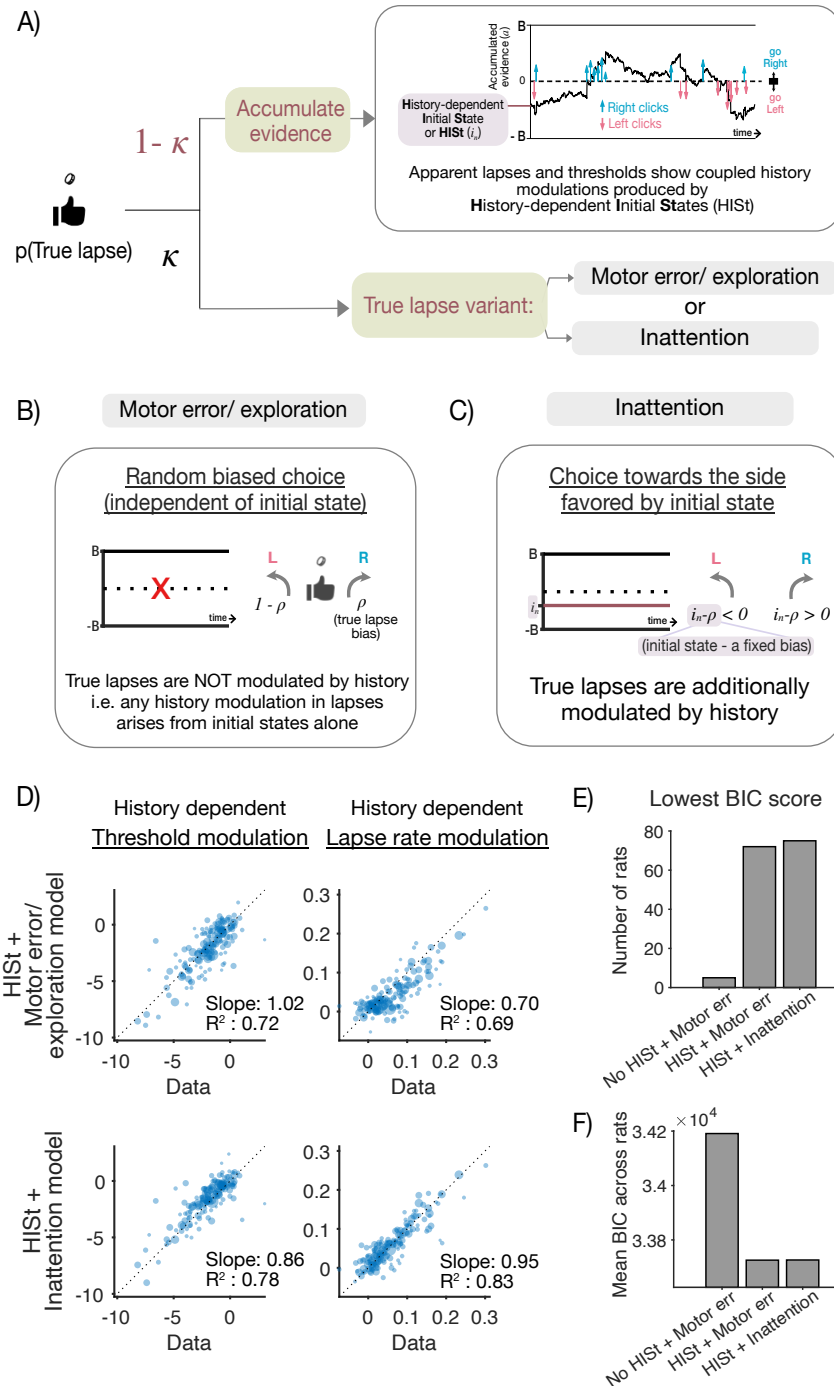


Supplementary Figure 2: **Influence of within- and across-trial parameters on history-dependent modulation of biases and lapses in the psychometric function.** (A) Influence of across-trial parameters on history-dependent modulation: Effects of varying magnitude of trial-by-trial updating ( $\eta$ ) - columns) and timescale of updating ( $\beta$ ) - rows) on initial state trajectories (gray lines) and psychometric curves (black - conditioned on all previous wins, blue - previous right wins, pink - previous left wins). (Top row) Small timescales of updating lead to fast fluctuations in initial states, and mostly horizontal shifts in psychometric curves with trial history, for both small and large magnitudes of updating (A1 and A4). (Middle row) Intermediate timescales of updating lead to slower fluctuations in initial state that have a cumulative effect across trials. For large magnitudes of updating (A5) these can give rise to apparent lapses (black intervals) as well as history-dependent modulation of these lapses (pink intervals). (Bottom row) Long timescales of updating lead to stronger cumulative initial state biases across trials, yielding apparent lapses and lapse rate modulations even for small magnitudes (A3). When combined with large magnitudes of updating (A6) lead to initial states that sometimes exceed the bounds, leading to a combination of apparent lapses (initial states within bounds) and deterministic, stimulus-independent decisions (initial states outside bounds). Caption continued on next page.

Supplementary Figure 2: (Previous page.) **(B)** Quantifying modulation of psychometric properties: The difference between psychometric curves conditioned on previous wins (blue) or losses (pink) can be quantified using two metrics - the horizontal distance between the midpoints of psychometric curves (“threshold modulation”) and the vertical distance between its asymptotes (“lapse rate modulation”) **(C)** Effects of varying the parameters of the within-trial drift diffusion model (DDM) on history-dependent threshold (x-axis) and lapse rate modulations (y-axis). Colors denote levels of sensory noise, size of dots denote values of the feedback parameter of the DDM. The feedback parameter determines if the accumulation is leaky ( $\lambda < 0$ , ignores early evidence), perfect ( $\lambda = 0$ , uses all evidence) or impulsive ( $\lambda > 0$ , ignores later evidence). Plots from left to right are for low, intermediate and high values of bound respectively. High bounds predominantly give rise to threshold modulations, however high positive values of feedback and higher levels of sensory noise additionally produce lapse rate modulations. Lapse rate modulations are dramatically increased by lower bounds for many different values of feedback and by higher values of sensory noise.



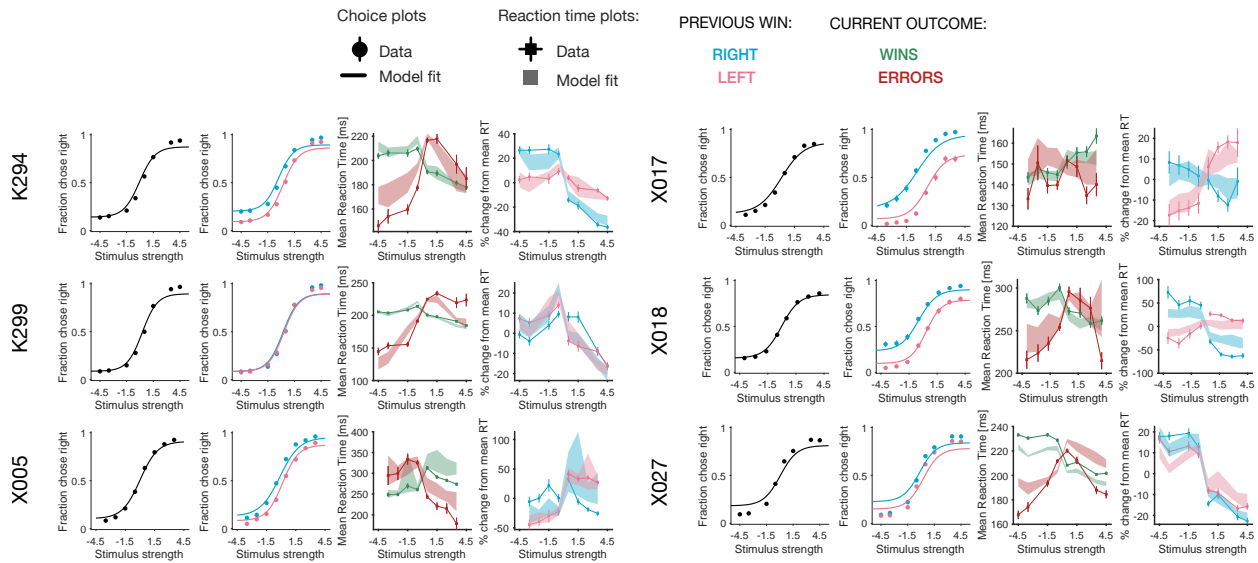
Supplementary Figure 3: **Performance measures across the rat dataset** **(A)** Histogram of trial counts for all rats in the population. Most rats completed on the order of  $1e^4$  trials. **(B)** Histogram of mean accuracy showing that rats showed good performance on the task (mean accuracy  $\pm$  SD:  $0.79 \pm 0.04$ ). **(C)** Average change in mean accuracy every 10000 trials, showing that rats’ performance was stable over time, reflecting asymptotic behavior rather than task acquisition. **(D)** Histogram of history-modulated lapse rates as a fraction of total lapse rates. A sizeable portion of the population had non-zero fractions, suggesting that history-dependence could potentially account for substantial lapse rate variance. **(E)** Scatter comparing repetition bias following wins and losses. Each point is a rat, error bars are Wilson binomial CIs. Most of the population occupied the bottom right quadrant, showing a “win-stay, lose-switch” bias i.e. repetitions following wins and alternations following losses.



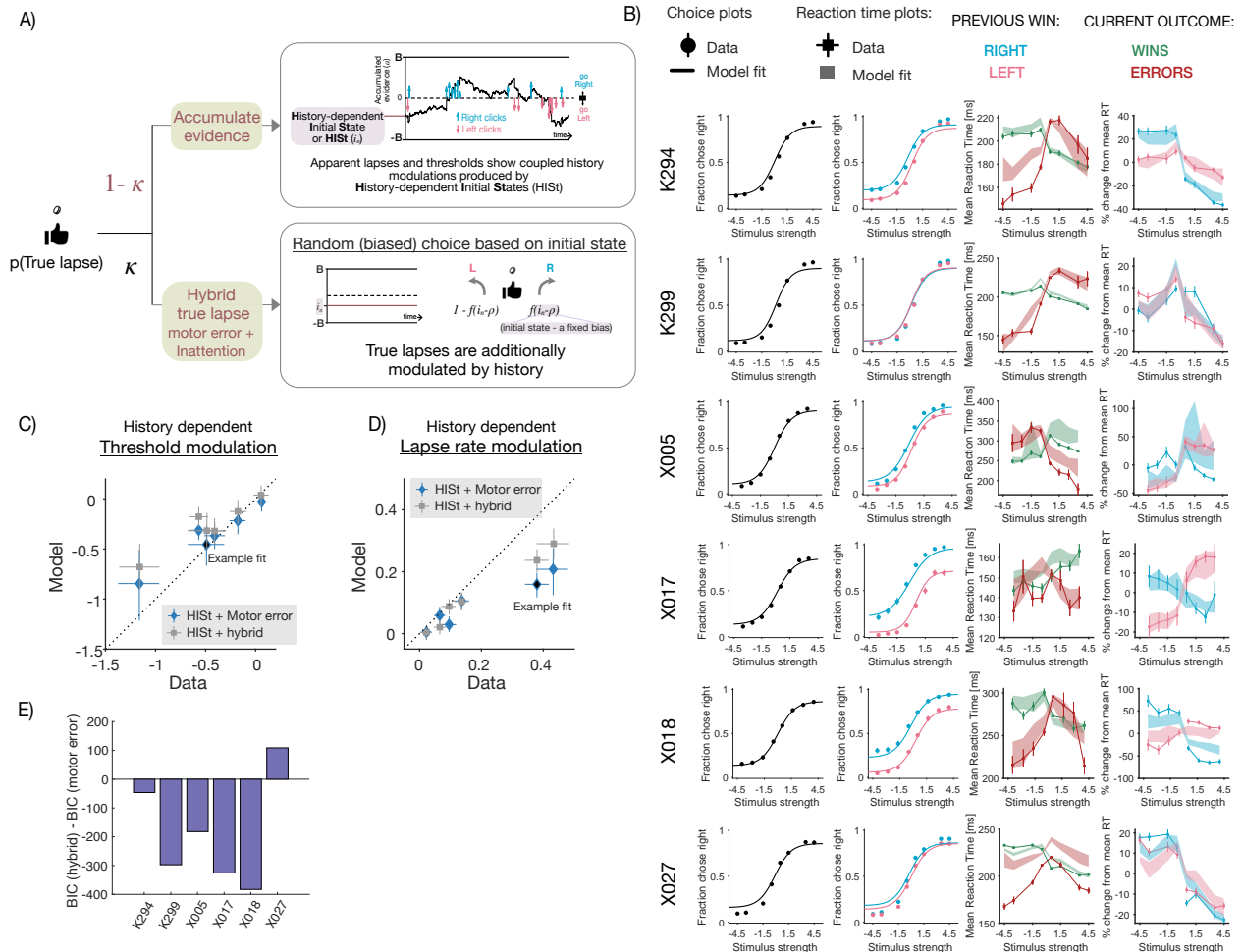
Supplementary Figure 4: **Two variants of the accumulator with HIST model with different kinds of true lapses perform equally well** (A) Schematic of accumulator with HIST (top), which produces apparent lapses and thresholds that are both modulated by history, with two variants of true lapses (bottom) - those due to motor errors/exploration, and those due to inattention. (B) Motor error/exploration variant, that occasionally chooses a random action with some bias ( $\rho$ ) irrespective of the initial state, reflecting an error in motor execution or random exploration. This model produces true lapses that are not modulated by history, such that any history modulations arise from HIST alone. Caption continued on next page.



Supplementary Figure 4: (Previous page.) **(C)** Inattentive variant, that occasionally fails to attend to the stimulus, and relies on the initial state to make an informed, deterministic decision based on the difference between the initial state and a bias ( $\rho$ ). In this model, true lapses are also modulated by history in addition to apparent lapses and thresholds. **(D)** Individual differences in history effects captured by different models: History modulations of threshold (left) and lapse rate (right) parameters measured from psychometric fits to the raw data (x-axis) v.s. model predictions (y-axis). (Top): Motor error/exploration model has no history dependence in true lapses, yet captures a majority of the variance in both threshold and lapse rate modulations [ $R^2 = 0.72$  (threshold parameter),  $R^2 = 0.69$  (lapse rate parameter)], and shows good correspondence with both parameters, as evidenced by the points lying close to the unity line [slope = 1.02 (threshold parameter), slope = 0.70 (lapse rate parameter)]. This suggests that these modulations can be captured by the comodulations in apparent lapses and thresholds produced by H1St. (Bottom): same as Top but for Inattention model. The inattention model allows true lapses to additionally depend on history, and captures slightly more variance in history modulations [ $R^2 = 0.78$  (threshold parameter),  $R^2 = 0.83$  (lapse rate parameter)]. However, it does so at the expense of correspondence with thresholds [slope = 0.86 (threshold parameter), slope = 0.95 (lapse rate parameter)]. This marginal improvement suggests that comodulations in thresholds and lapse rates largely reflect apparent lapses arising from H1St, rather than true lapses of either kind. **(E)** Distribution of best fitting model variants for individual rats: Overall bar height for each variant denotes the total number of rats for which that variant scored the lowest BIC score. Inattention variant won in marginally more rats than motor error (inattention: 75/152 rats, motor error: 72/152 rats). **(F)** Population model comparison using mean BIC score across rats. Lower scores indicate better fits. Scores are comparable across variants, marginally favoring motor-error over inattention (Mean BIC score for motor error/exploration: 33725.64, inattention: 33726.25 ).



**Supplementary Figure 5: Fits of the accumulator model with history-modulated initial states (and additional true lapses arising from motor error) to choices and reaction times of individual rats.** Each horizontal set of 4 panels shows fits to an individual rat, and each of the 4 columns depicts a different behavioral measure summarizing choices (first column, psychometric curve), reaction times (third column, win RTs in green and error RTs in red), and history modulations in choices/reaction time (second/fourth column, psychometric curves/RTs conditioned on previous right wins (blue) or left wins (pink)). Data represented by points (circles: choices, squares: reaction times) and model fits represented by lines (choices) or shaded bars (reaction times, thickness represents 95% bootstrap prediction intervals).



Supplementary Figure 6: **Accumulator model with HIST and true lapse variants fit to the RT dataset** (A) Flexible true lapse variant of the accumulator model with HIST, capable of producing both motor errors and inattentional true lapses: in this model, the subject chooses stochastically on true lapse trials, with a probability given by a sigmoidal function of the initial state (such a strategy is often called probability matching and although suboptimal, has found empirical support in many perceptual tasks e.g. Mamassian et al. 2002). The slope parameter of the sigmoid which is fit to the data – when this slope parameter goes to infinity, this model picks deterministically based on the initial state, similar to the previous inattention model. On the other hand when the slope goes to zero, choices on true lapse trials are no longer dependent on initial states, reducing to the motor error/exploration model, with intermediate parameters interpolating between these two extremes. This “hybrid true lapse” variant of the model can flexibly include many kinds of true lapses. (B) Fits of the hybrid model to individual rats in the RT dataset, showing choice, RT and history measures similar to Supp Fig. 5. (C-D) History modulations in psychometric thresholds (C) and lapse rates (D) for motor error (blue) and hybrid (grey) models. Once again, allowing for the possibility of history-modulated true lapses slightly improves correspondence to lapse rate modulations, at the cost of threshold modulations. (E) Difference in BIC scores between the hybrid and motor error models, for individual rats in the reaction time dataset. Negative values indicate that the more flexible hybrid model won, as is the case for most rats.