Neural population dynamics underlying evidence accumulation in multiple rat brain regions

Brian DePasquale[1], Carlos D. Brody[1,2,3,*], Jonathan W. Pillow[1,4,*]

[1] Princeton Neuroscience Institute,

[2] Howard Hughes Medical Institute,

[3] Department of Molecular Biology,

[4] Department of Psychology,

Princeton University, Princeton NJ, USA

[*] These senior authors contributed equally to this work

Correspondence: B. DePasquale (depasquale@princeton.edu), C. Brody (brody@princeton.edu), and J. Pillow (pillow@princeton.edu)

Keywords: latent variable models, decision making, evidence accumulation

Abstract

Accumulating evidence in service of sensory decision making is a core cognitive function. However, previous work has focused either on the dynamics of neural activity during decision-making or on models of evidence accumulation governing behavior. We unify these two perspectives by introducing an evidence-accumulation framework that simultaneously describes multi-neuron population spiking activity and dynamic stimulus-driven behavior during sensory decision-making. We apply our method to behavioral choices and neural activity recorded from three brain regions — the posterior parietal cortex (PPC), the frontal orienting fields (FOF), and the anterior-dorsal striatum (ADS) — while rats performed a pulse-based accumulation task. The model accurately captures the relationship between stimuli and neural activity, the coordinated activity of neural populations, and the distribution of animal choices in response to the stimulus. Model fits show strikingly distinct accumulation models expressed within each brain region, and that all differ strongly from the accumulation strategy expressed at the level of choices. In particular, the FOF exhibited a suboptimal 'primacy' strategy, where early sensory evidence was favored. Including neural data in the model led to improved prediction of the moment-by-moment value of accumulated evidence and the intended—and ultimately made—choice of the animal. Our approach offers a window into the neural representation of accumulated evidence and provides a principled framework for incorporating neural responses into accumulation models.

Introduction

Accumulation of evidence is a critical mental process that facilitates decision making when information relevant to that decision unfolds in time. Behavioral experiments with precise control over the timing of evidence have offered an abundance of behavioral choice data from animals performing these tasks, and sophisticated behavioral models have been developed to use the choice data to infer the accumulation strategies that underlie these choices (Brunton, Botvinick, and Brody 2013). In parallel, neural correlates of accumulated evidence have been found in several brain regions (Shadlen and Newsome 2001; Joshua I. Gold and Shadlen 2007; Hanks et al. 2015; Yartsev et al. 2018; Huk and Shadlen 2005; Roitman and Shadlen 2002; Erlich, Bialek, and Brody 2011; Mante et al. 2013; Kim and Shadlen 1999; Long Ding and Gold 2010), and a variety of methods have been developed to describe the statistical relationship between neural activity and accumulated evidence (Park et al. 2014; Beck et al. 2008; Latimer et al. 2015; Aoi, Mante, and Pillow 2020; Zoltowski et al. 2019; Zoltowski, Pillow, and Linderman 2020; Latimer and Freedman 2021).

However, it remains an open problem to integrate these two perspectives to obtain a comprehensive account of how accumulated evidence underlies simultaneous neural activity and observer choice. For example, few analysis methods which use precise spike timing information take into account the timing of stimulus information (e.g. (Latimer et al. 2015)) and likewise few analysis methods which use the precise timing of stimulus information use neural responses to infer accumulated evidence (e.g. (Hanks et al. 2015)). To address this gap, we developed a framework for inferring a probabilistic evidence accumulation model jointly from choice data and neural activity recorded while animals perform a pulse-based evidence accumulation task. Our framework builds on prior work that learns latent accumulator models from choice data (Brunton, Botvinick, and Brody 2013) by incorporating spiking neural responses as additional observations. We developed a computationally-efficient method to fit the model and applied it to choices and neural responses from three brain regions known to be involved in evidence accumulation.

We found that a model with a shared latent variable representing accumulated evidence could accurately account for both neural and behavioral data. Application of the method identified distinct signatures of accumulation employed by each brain region, all of which differed from the accumulation model that governed behavior, suggesting that whole-organism accumulation strategies may result from the collective action of multiple accumulation processes. Incorporating neural activity into accumulation models greatly reduced the uncertainty in the moment-by-moment value of accumulated evidence when compared to similar models fit only to animal choices. This reduction in uncertainty led to a more refined picture of the moment-by-moment value of accumulated evidence, which in turn made the model more informative about what choice the animal intended to, and ultimately, made.

Broadly, our framework offers a unified, mechanistic and probabilistic description of the moment-by-moment accumulation process that underlies decision-making. Our flexible framework offers a computationally efficient method for identifying a key normative decision-making model using multiple different types of data, and can easily accommodate simultaneous recordings from many neurons, as well as recordings performed sequentially over many days. It provides a platform for quantitatively characterizing choice-related information

in neural responses, and can be used to understand how different brain regions differentially implement an accumulation strategy.

Results

We analyzed behavioral and neural data from rats trained to perform a previously developed decision-making task (Brunton, Botvinick, and Brody 2013). In this task, rats listen to two simultaneous series of randomly timed auditory clicks, one from a speaker to their left and one from a speaker to their right (**Fig 1A**). Upon cessation of the stimulus, the animal must orient in the direction of the greater number of clicks to receive a reward. **Figure 1A** illustrates the temporal structure of the task.

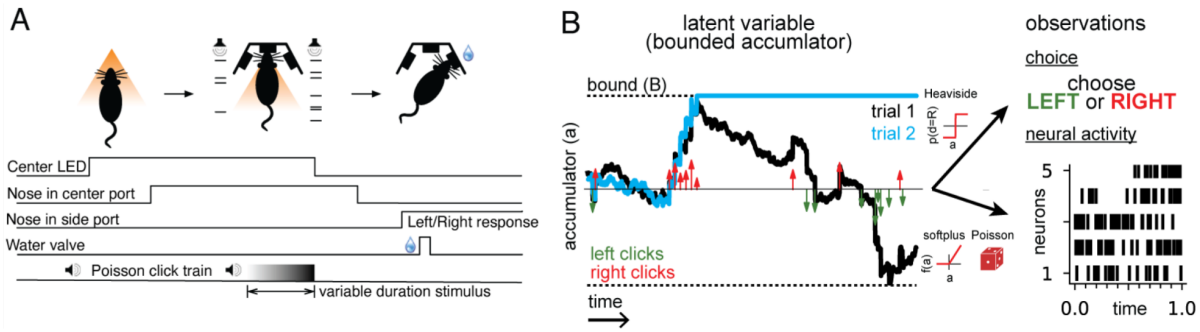


Fig 1: Accumulating evidence task and latent variable model. (A) Rats performed a pulsed-based evidence accumulation task. A central LED illuminates, indicating that the rat can begin a trial by poking its nose in a central port. After a delay of random duration, an auditory stimulus of variable duration is delivered—a series of brief auditory pulses played from a left and a right speaker. Upon cessation of the stimulus, the rat must orient to the direction of the greater number of pulses to receive a water reward. (B) The model relates the click-based sensory stimulus to two types of observations—the animal's choice and neural activity observed during the task. The latent variable model is a bounded accumulator. Left and right clicks (green and red arrows, respectively) push the variable to one side or the other; if the accumulator variable reaches the bound *B* (dotted line) accumulation ceases. Seven parameters govern the dynamics of *a(t)* (see main text). Two different hypothetical trajectories of *a(t)* are illustrated (black and blue) for the same click stimulus; the two trajectories differ due to the diffusive and stimulus noise in the model. *a(t)* relates to the animal's choice by a Heaviside step function and to neural activity by way of a softplus nonlinearity and a Poisson distribution. *a(t)* is common for all simultaneously recorded neurons and each neuron has its own parameters that determine its tuning curve.

We analyzed behavioral choice data and electrophysiological neural recordings from 11 rats. In total, we analyzed 37,179 behavioral choices and 141 neurons from three brain areas—the posterior parietal cortex (PPC), the frontal orienting fields (FOF) and the anterior-dorsal striatum (ADS). Prior electrophysiological and lesions studies have shown that these three brain regions play a key role in the accumulation process (Hanks et al. 2015; Yartsev et al. 2018; L. Ding and Gold 2013; Joshua I. Gold and Shadlen 2007; Shadlen and Newsome 2001; Kim and Shadlen 1999; Mante et al. 2013; Roitman and Shadlen 2002; Erlich et al. 2015; J. I. Gold and Shadlen 2000; Huk and Shadlen 2005; Erlich, Bialek, and Brody 2011; Long Ding and Gold 2010).

The analyzed neurons represented a subset of the entire population of recorded neurons; they were selected because of their significant tuning for choice during the stimulus period (two-sample t-test, $p < 0.01$) which we used as a surrogate measure of each neuron's task involvement. Data were collected after the animals were well-trained on this task and exhibiting a high-level of performance, with an average correct choice probability of 0.74. These data have been analyzed previously with other methods (Hanks et al. 2015; Yartsev et al. 2018). Information about the data is summarized in **Supplementary Table 1**. 68 neurons within the FOF and 17,382 behavioral choices were recorded from five rats over 46 behavioral sessions. 25 neurons within the PPC and 9,037 behavioral choices were recorded from three rats over 24 sessions. 48 neurons within the ADS and 10,760 behavioral choices were recorded from three rats over 27 behavioral sessions.

A latent variable model of behavioral choice and neural activity

One of the most commonly-used normative models used to account for the internal mental processes that underlie evidence accumulation is the drift-diffusion to bound model (DDM; **Fig 1B**) (Ratcliff and McKoon 2008; Brunton, Botvinick, and Brody 2013). Previous work fit this model using the choices subjects made (or reaction time information, if that data is available; Brunton, Botvinick, and Brody 2013). Here we seek to jointly model the relationship between accumulated evidence, choices and neural activity.

A DDM has a single latent variable, which we denote as *a(t)*, that evolves in time. The variable *a(t)* represents the current, inner mental representation of the evidence in support of a leftward or rightward choice at each moment in time. The temporal evolution of *a(t)* is governed by the following dynamical equation,

$$da = \lambda a dt + \Delta(t) dt + \sigma_a dW + \sigma_s \Sigma(t) \eta dt, \qquad \text{(Equation 1)}$$

where da is the amount *a(t)* changes in a time dt, λ is a leak or instability parameter, $\Delta(t)$ represents the click evidence at time t, $\sigma_a dW$ is a diffusive Gaussian noise process (or Weiner process) with scaling σ_a, and $\sigma_s \Sigma(t) \eta$ is noise induced by each click input, with standard deviation $\sigma_s \Sigma(t)$, where $\Sigma(t)$ is the sum of clicks at time t. Each parameter is discussed in greater detail below. The key distinction of our approach, described below, is that the accumulator variable *a(t)* drives both choices and neural activity.

The four terms of Equation 1 each account for specific ways in which *a(t)* might reflect accumulated evidence. The first term specifies how recent values of *a(t)* influence future values of *a(t)*, and is governed by the learnable parameter λ that determines the timescale of this effect. Values of λ greater than zero cause values of *a(t)* at the beginning of a trial to have greater impact on future states of *a(t)* than later values of *a(t)*. Values of λ less than zero cause recent values of *a(t)* to have greater impact on future states of *a(t)* than values of *a(t)* that occurred long in the past. When λ equals zero, the current value of *a(t)* is retained indefinitely, a signature of perfect memory of the recent past.

The second term specifies how the click stimulus is incorporated into *a(t)*. $\Delta(t)$ indicates the difference of the number of left and right clicks at time t, after their magnitude has been adapted based on recent stimulus history (see Methods). Because the goal of the task is to report whether there was a greater number of left or right clicks, only the click difference is required to correctly perform the task.

The first two terms are designed to account for the deterministic (non-random) dynamics exhibited by *a(t)*. The dynamics of *a(t)* however also evolve due to a non-deterministic component; for example, it is likely that *a(t)* would not follow the exact same dynamical trajectory if the animal experienced an identical sensory stimulus more than once. The final two terms account for these effects by introducing 'noise' into *a(t)*. The effect this noise has on *a(t)* is illustrated in the blue trajectory in **Figure 1B**— although the click inputs have the same effect on *a(t)* as they do for the black trajectory, noise leads to differences in the overall trajectory.

The third term is diffusive noise that corrupts *a(t)* by a random amount continuously in time; σ_a specifies the magnitude of this noise. The final term introduces noise into *a(t)* at moments when clicks are experienced, in proportion to the total number of clicks that occur at that moment. Defined similarly to $\Delta(t)$, $\Sigma(t)$ indicates the sum of the adapted click magnitude at time t. $\sigma_s \eta$ is a Gaussian random variable with a mean of zero and standard deviation σ_s.

The dynamics of *a(t)* also includes a condition such that if *a(t)* becomes greater in magnitude than a symmetric boundary with magnitude B (**Fig 1B**, dotted lines), then $da = 0$, and accumulation ceases for the duration of the trial. For example, the blue trajectory in **Figure 1B** crosses B and as a result does not change in time after that point.

To incorporate both types of data into the model, we model choices as arising from the accumulated evidence at the end of the trial. Specifically, we describe the probability of a rightward choice given the evidence *a(T),* where T is the length of the stimulus period, using a step function with lapses. With probability γ the animal lapses and picks one of the two sides at random, independent of the stimulus. With probability $(1 - \gamma)$ the animal makes a rightward choice if $a(T) > c$ and a leftward choice if $a(T) < c$, where c denotes the criterion. In equations, this model can be expressed:

$$P(d = R) = \gamma/2 + (1 - \gamma)H(a_T - c) \qquad \text{(Equation 2)}$$

where $H(\cdot)$ is the Heaviside step function.

To model neural spike train data, we describe the time-varying firing rate of each neuron as a soft-rectified linear function of the accumulated evidence:

$$f_{\theta_n}(a(t)) = \text{softplus}(\theta_n a(t) + \theta_n^0(t)), \qquad \text{(Equation 3)}$$

where the softplus function (**Fig 1B**) is given by softplus(x) = log(1+exp(x)) and θ_n denotes the slope of the linear relationship between *a(t)* and firing rate. θ_n is fit separately for each neuron. $\theta^0_{n,t}$ is a second, learnable parameter for each neuron that captures time-varying changes in

firing rate that do not depend on *a(t)* (Methods). The spikes of each neuron are modeled as a Poisson process with a time-dependent conditional intensity function $f_{\theta_n}(a(t))$(See Methods for details).

We refer to the set of all parameters that govern *a(t)*, and its relationship to the neural activity and choice data as $\Theta = \{\sigma_i, B, \lambda, \sigma_a, \sigma_s, \phi, \tau_\phi, \theta_{1:N}, c, \gamma\}$, where σ_i is the variance of *a(t)* at the start of the trial, and ϕ and τ_ϕ determine how the magnitude of each click is adapted based on the timing of recent clicks (see Methods). These parameters can be divided into three groups, based on their role in the model: those governing the latent accumulation dynamics ($\theta_a = \{\sigma_i, B, \lambda, \sigma_a, \sigma_s, \phi, \tau_\phi\}$), those relating *a(t)* to choice ($\theta_d = \{c, \gamma\}$), and those relating *a(t)* to the neural responses ($\theta_y = \{\theta_{1:N}\}$).

We fit Θ separately for each brain region using maximum likelihood (see Methods). Briefly, maximizing the likelihood of the data requires computing the temporal evolution of the probability distribution of *a(t)* over the duration of a single trial, for all trials, and computing the probability of the observed spikes and choices under this distribution. The dynamics of this probability distribution can be expressed using the Fokker-Planck equation, and previous work has developed methods for numerically solving it (Brunton, Botvinick, and Brody 2013; see Methods). This method allows the likelihood to be computed with a single forward pass over the data. We refer to the value of Θ that maximizes the likelihood of the data as $\hat{\Theta}$. We verified that our method was able to recover the parameters that generated synthetic physiologically-relevant spiking and choices data (**Supp. Fig 1**).

A shared accumulator model captures both neural responses and choices for each brain region

We fit the model separately to data from each brain region (see **Supp. Table 1** for information about the data). To verify that the model fits were consistent with the data, we compared the peri-stimulus time histograms (PSTHs; (**Fig 2A** & **B**) and psychometric curves (**Fig 2C**) of the empirical data to synthetic data simulated from the fitted model for each brain region. The PSTH of most neurons showed a characteristic choice preference that increased over time, consistent with accumulation, and the model was able to capture this effect (**Fig 2A**). The model provided an accurate account of mean responses in all three brain areas (**Fig 2B**), with a median R^2 of 0.91, 0.68, and 0.87 for the FOF, PPC, and ADS respectively (**Fig 2B**, colored lines). **Fig 2C** shows a comparison between true psychometric curves and the psychometric curve of the fitted model, confirming that the model also accounted for psychophysical choice behavior. (R^2: 0.99-FOF; 0.99- PPC; ADS-0.97; see Methods for details).

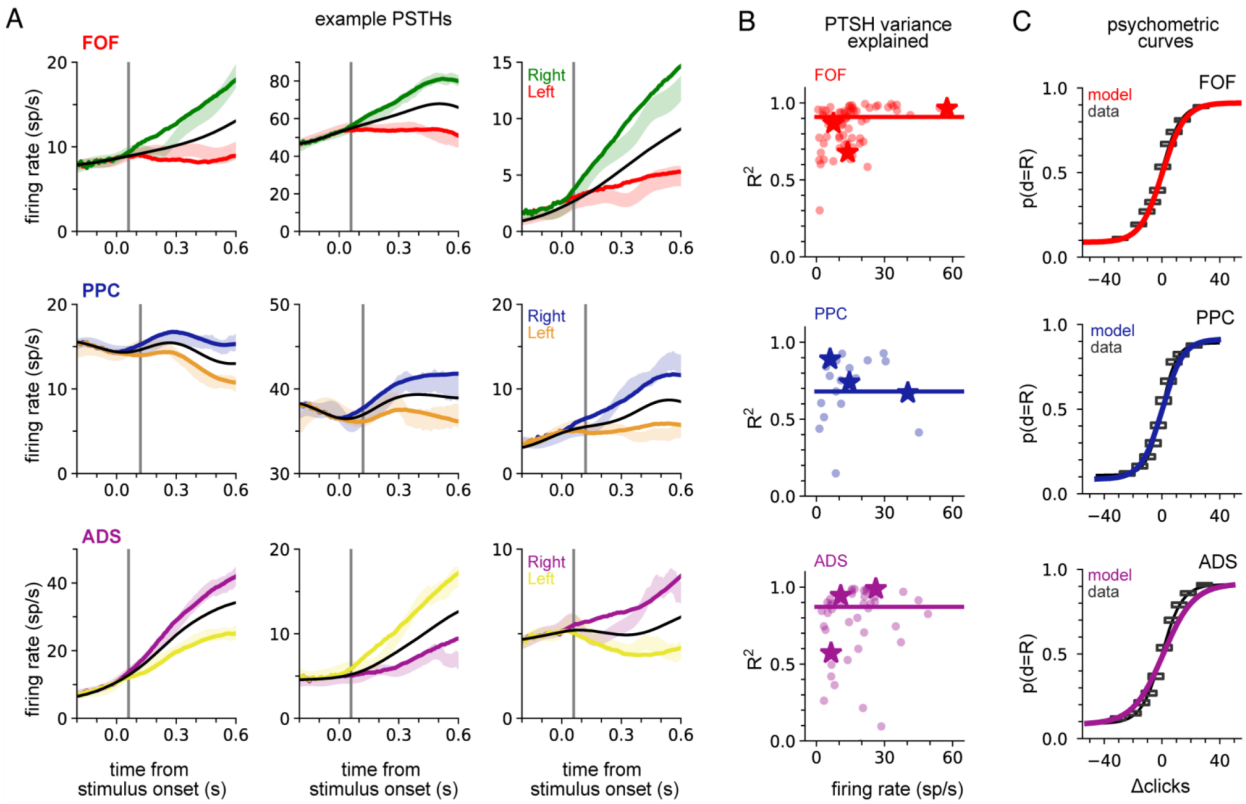


Fig 2: A shared accumulator model captures both neural response and choice for each brain region. (A) Peri-stimulus time histograms (PSTHs) of three example neurons for each brain region (each row; FOF: red/green, PPC: blue/orange, ADS: purple/yellow). Spike trains were binned, filtered with a Gaussian kernel (std = 50 *ms*), grouped based on the strength of evidence, and then averaged. Transparent shaded regions are +/- 1 standard error of the mean for the empirical data for each grouping. Colored curves are the mean of synthetic data simulated from the model with the parameters that maximze the likelihood of the data, grouped in a similar fashion. The black curve shows the trial-averaged firing rate, for all evidence strengths. Grey vertical lines indicate the average delay between the stimulus and the response for each brain region (see Methods). (B) Coefficient of determination (R^2) between PSTH generated with empirical data and PSTH generated with simulated data, for each neuron in each brain region. The data are plotted as a function of average firing rate; error decreases for higher firing rate neurons. The median across the population is shown as a line. Points indicated with a 'star' refer to the data plotted in (A). (C) Probability of making a rightward choice as a function of cumulative difference in the number of clicks (psychometric curves) for empirical data (black lines) and data simulated from the model with the best fitting parameters (colored curves; FOF: red, PPC: blue, ADS: purple). Each curve is the curve of best fit, as computed by logistic regression.

These analyses support the conclusion that a shared accumulator model for each brain region is sufficient to capture the primary features seen in the data, namely the animals' choice sensitivity to the stimulus and each strength of accumulated evidence reflected in each neuron's response.

Accumulation in different brain regions is governed by different dynamics that differ from the dynamics governing choice

The primary motivation of our study was to learn accumulator models that incorporate the precise stimulus timing information that our task affords and describe the animal's choice as well as the rich temporal structure in neural activity. Previous efforts have only modeled choice

using stimulus timing information (Brunton, Botvinick, and Brody 2013) or have modeled neural activity without choice for tasks without detailed stimulus timing information (Latimer et al. 2015; Zoltowski, Pillow, and Linderman 2020). Throughout, we refer to our model that describes both neural activity and choice data as the 'joint neural-behavioral model' or the 'joint model'.

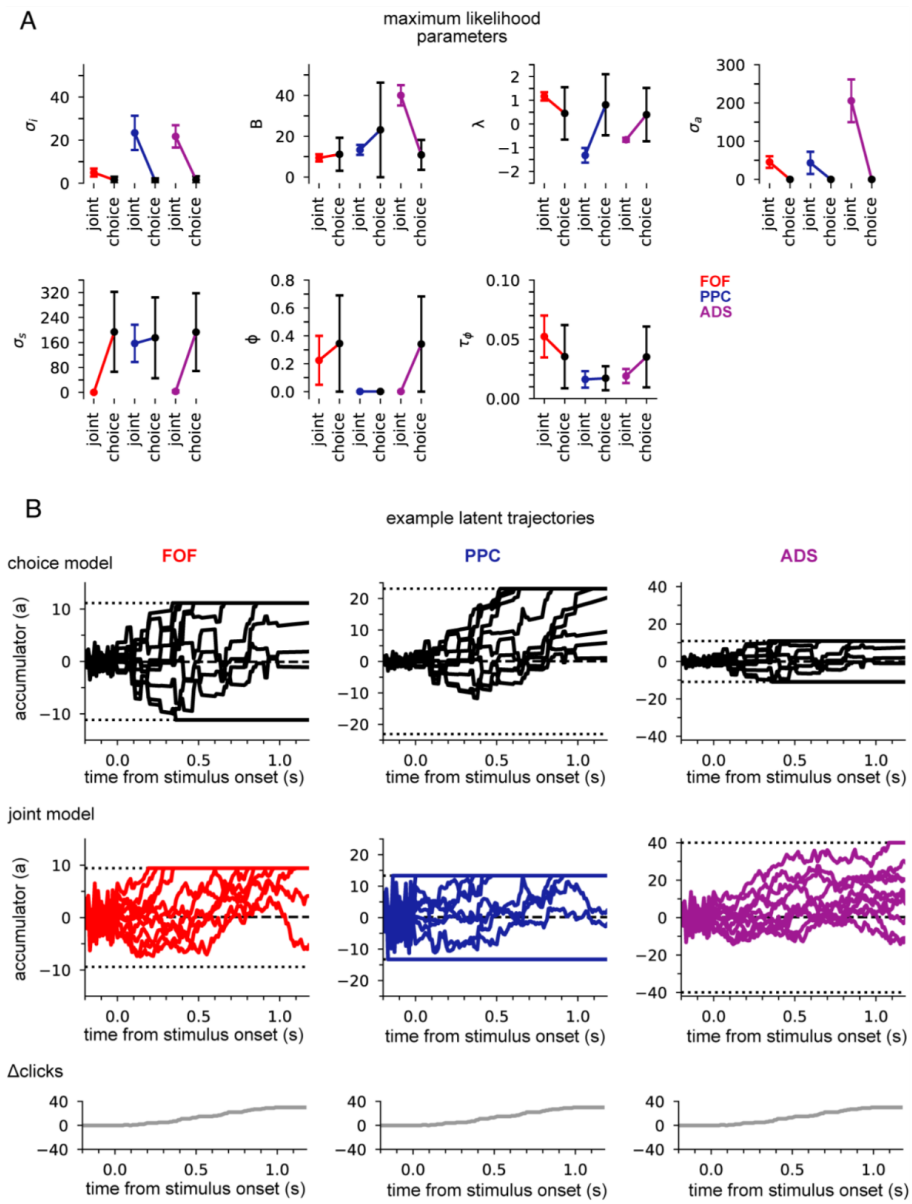


Fig 3: Accumulation in different brain regions is governed by different dynamics that differ from the dynamics governing choice. (A) Maximum likelihood parameters that govern *a(t)* for the joint neural-behavioral model and the behavior-only model. Error bars, computed by the Laplace approximation (Methods), are +/- 2 standard deviations. Parameters are σ_i: initial variance, B: accumulation bound, λ: drift, σ_a: accumulation noise variance, σ_s: click noise variance, ϕ: adaptation strength, τ_ϕ: adaptation timescale. (B) 10 example trajectories with different noise instantiations for one trial for the choice model (top) and the joint model (middle) model for each brain region, and cumulative sum of the click stimulus for each trial (bottom). The dotted black lines (top and middle) indicate the boundary value for each model.

We first compared the joint neural-behavioral model to an accumulator model where only the stimulus is used to model the animal's choice (i.e. neural activity is not used). To fit such a 'choice'-only accumulator model we fit the same latent variable model for *a(t)* using only choice data (see Methods).

Figure 3A shows the maximum likelihood parameter estimate for the joint and choice-only accumulator models for each brain region. In the case of the choice model, no neural data was used, so the brain region designates only the cohort of animals from which the choice data was taken. We found that both fitted models exhibited strong adaptation ($\phi << 1$) consistent with prior work using choice accumulator models (Brunton, Botvinick, and Brody 2013). This indicates that each stimulus pulse does not equally impact *a(t)*. Rather, clicks that occur soon after another click will have a smaller effect on *a(t)* than an isolated click.

However, we found that *a(t)* was impacted by very different forms of noise in the two models. The fitted choice-only models exhibited small amounts of diffusive noise ($\sigma_a \approx 0$) and large amounts of stimulus noise ($\sigma_s >> 1$), consistent with earlier findings. By contrast, the joint model fits exhibited large levels of diffusive noise ($\sigma_a > 0$) and large variability in the initial value of *a(t)* ($\sigma_i >> 0$). A likely cause for the large initial variability in *a(t)* for the joint model is the variability in neural responses prior to stimulus onset (Churchland et al. 2010). The effect of these different combinations of parameters can be seen in **Figure 3B**, which shows 10 noise realizations of one trial for both models in each brain region. Choice accumulator models have smaller variability in the initial state and more variability when clicks arrive, while joint accumulator models have larger initial variance and diffusive noise that is constant in time. This difference suggests that the accumulation dynamics underlying neural activity suffers from specific forms of noise that are resolved at the level of a behavioral accumulator model. It's notable that only the PPC joint model had a significant stimulus noise parameter, suggesting that the source of stimulus noise for behavioral accumulators resides in a small number of brain regions.

Because each brain region was fit separately, we could compare the best-fit parameters across the three brain regions (**Fig. 3A**). We focus on one of the most salient differences in parameters across the three regions — the leak or instability parameter λ. Although there was no significant difference in the value of λ across the cohorts of animals in the choice-only model, we found substantial differences across brain regions in the joint model fits (**Fig 3A**). Fits to the PPC and ADS data revealed leaky accumulator models ($\lambda < 0$). Surprisingly, however, the model fit to FOF data exhibited unstable accumulation dynamics ($\lambda > 0$) meaning that the accumulator (and thus firing rates) are more strongly affected by early stimulus clicks. The stronger weighting of earlier clicks was compounded further by the low accumulation bound exhibited by the FOF. Such a low bound, in conjunction with unstable accumulation, causes *a(t)* to stop evolving early in the trial (see **Fig 3B**). This results in a phenomenon known as 'primacy encoding', in which early clicks more strongly impact the animal's choice while later clicks are ignored. We confirmed this finding in the FOF using a generalized linear model (GLM) regression analysis (see Methods & **Supp Fig 2**). This result is consistent with previous work suggesting that the FOF has a categorical representation of *a(t)* (Hanks et al. 2015). We expand on this finding in light of other studies of the FOF in the Discussion.

ADS is better described by independent accumulators

Next, we sought to examine whether the model could account for the correlations simultaneously recorded neurons exhibited. The model we introduced describes the activity of any group of simultaneously-recorded neurons as being jointly generated from a shared, one-dimensional latent accumulation variable. Can this model provide an accurate description of the joint firing statistics of these populations? For these analyses, we focused on the FOF and ADS datasets, which were the only regions with enough simultaneously-recorded neurons to make meaningful comparisons (see **Supp Table 1**).

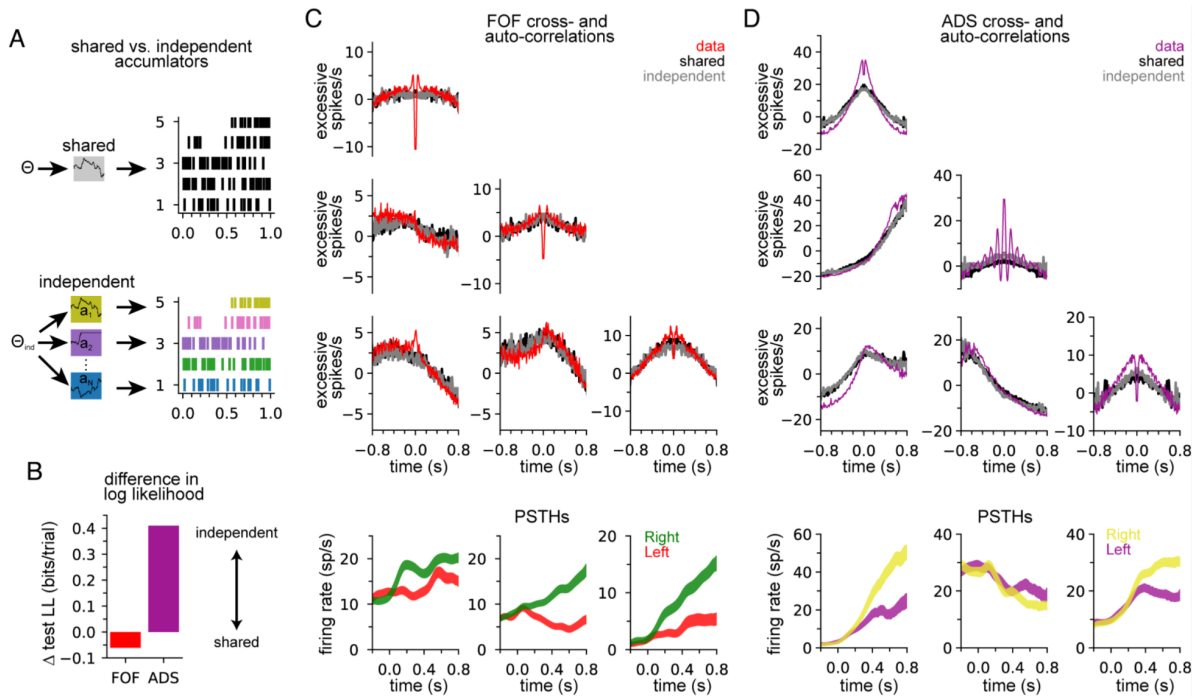


Fig 4: ADS is better described by independent accumulators. (A) Schematic of the independent accumulator model. For the shared accumulator model (top), a set of parameters defines the dynamics of a single accumulator, which drives the spiking activity of the entire population. In the independent accumulator model, a set of parameters defines the dynamics of an ensemble of independent accumulator models, which each individually determine the spiking of a single neuron. (B) Difference in test log-likelihood (bits/trial) for the shared versus independent accumulator models. Values greater than zero indicate that the independent accumulator model more accurately captures the data. (C) Empirical (red) and synthetic (shared: black; independent; grey) auto- and cross-correlations for three simultaneously recorded neurons from the FOF. Auto-correlations are shown along the diagonal. Corresponding PSTHs are shown below for reference. (D) Empirical (purple) and synthetic auto- and cross-correlations (shared: black; independent; grey) for three (of five) simultaneously recorded neurons from the ADS. Corresponding PSTHs are shown below for reference.

To understand how well a one-dimensional accumulator model captures shared neural variability, we compared the basic model described above to an 'independent accumulator' in which each neuron is driven by an independent accumulation variable with its own independent noise (**Fig 4A**). If trial-to-trial variability in neural spiking is largely driven by

temporal covariation in the accumulator, a model of independent accumulators (which do not share this covariation) should not account for the data as well.

We refer to the set of all parameters of the independent accumulator model as Θ_{ind} and identified Θ_{ind} using the same optimization method as described above, albeit with modified objective function (see Methods). For both brain regions, the maximum likelihood parameters for the shared and independent accumulator models were qualitatively similar, with one minor difference in the magnitude of the initial variance of the accumulator in the ADS (**Supp. Fig 3**).

To characterize how well each model captures shared neural response variability, we computed cross-correlation functions (Methods; **Fig 4C** & **D**) for all pairs or groups of simultaneously recorded neurons. We first computed these metrics for the empirical data (**Fig 4C** & **D**; *colored traces*) and also for synthetic data generated using the maximum likelihood parameters of each model (**Fig 4C** & **D**; *black trace*-shared model; *grey trace*- independent model). Visually, the synthetic data of both models accurately captured the major trends in the auto- and cross-correlation functions at a slower time scale consistent with covariations in firing rate, but failed to capture fluctuations on short time scales, both in auto- and cross-correlations. Prior work has illustrated how more elaborate spike-history terms that our models lack can account for the latter effect (Park et al. 2014; Pillow et al. 2008).

When we considered the model's ability to match the empirical auto- and cross-correlations across all pairs of simultaneously recorded neurons across the entire dataset (70 neurons in total) we found a high level of agreement. We found that the shared and independent accumulation models provided approximately equally accurate fits to the auto- and cross-correlations between simultaneously recorded neurons (median R^2 of 0.86 and 0.80 for shared model vs. 0.86 and 0.80 for independent model, for FOF and ADS respectively). This shows that a single, shared accumulator is not necessary to generate the correlations observed in simultaneously recorded spike trains.

We also used 5-fold cross-validation to compare the accuracy of the two models, and found that the independent accumulator was better able to capture the ADS neural activity and choices than a shared model, while the FOF data weakly favor a shared model (**Fig 4B**). These results were suprisinging given the equal ability of both models to capture the auto- and cross-correlation functions of the ADS, but suggest that the neural activity of the ADS is not best described by a one-dimensional accumulator but rather a multi-dimensional latent process. Although ADS datasets with 4 or more neurons provided the primary contribution to the model-comparison results favoring the independent accumulator over the shared accumulator model (**Supp Fig 4A**), when the number of neurons in ADS datasets were subsampled to match the maximum number of neurons in FOF sessions (3 neurons) the ADS recordings still favored an independent accumulator model (**Supp Fig 4B**).

To gain additional insight into this finding, we computed a measure of response dimensionality known as the participation ratio (Litwin-Kumar et al. 2017) computed using the Eigenvalues of the covariance matrix of the firing rates (Methods). If the firing rates are all independent the Eigenvalues will all be equal and the participation ratio will equal the number of neurons. If the firing rates are correlated such that some Eigenvalues are small (or perhaps even zero) the participation ratio will reflect this and the dimensionality of the data will be less than the

number of neurons. Consistent with our modeling results, we found that responses in ADS had higher dimensionality than in FOF and that sessions with greater dimensionality were those that favored independent accumulator models (**Supp Fig 4C**).

Together, these analyses support the conclusion that within the ADS, individual neurons largely reflect independent accumulation processes, while neurons in the FOF appear to reflect a single latent accumulator.

Neural data provides more information about accumulated evidence on single trials than choice alone

Next we examined how neural data affected our ability to make inferences about the accumulated evidence $a(t)$. To do this, we computed the posterior distribution over $a(t)$ for the joint model, given choice data only or given neural data and choice data. The posterior distribution combines information from multiple sources — stimulus, choice and neural activity — to offer a concise window into the animal's internal state of evidence accumulation.

Figure 5A shows the posterior distribution over $a(t)$ under the model for three example trials (one for each brain region) — first, when only choice data was included, and second, when both choice and neural data were included. When we considered only choice data, the posterior over $a(t)$ was broad; a large set of $a(t)$ trajectories were all consistent with the animal's choice. However, when we considered both choice and neural spiking activity, we obtained a substantially narrower distribution over $a(t)$, meaning that the joint model including neural data offers greater confidence in the precise value of accumulated evidence at each moment within a trial.

To quantify this difference, we computed the standard deviation of the two posteriors (**Fig 5B**). For all brain regions, the median posterior standard deviation given neural data and choice was smaller than when conditioning only on choice (**Fig 5B**; median difference FOF: 0.46; PPC: 0.72; ADS: 2.23). This reduction in the standard deviation increased when more neurons were used to compute the posterior (**Fig 5C**). The increased confidence in $a(t)$ provided by neural activity makes intuitive sense: temporally specific spiking activity, e.g. in the middle of a trial, allows one to infer that $a(t)$ has increased in favor of a choice, whereas choice information can only offer certainty about the range of $a(t)$ at the end of the trial.

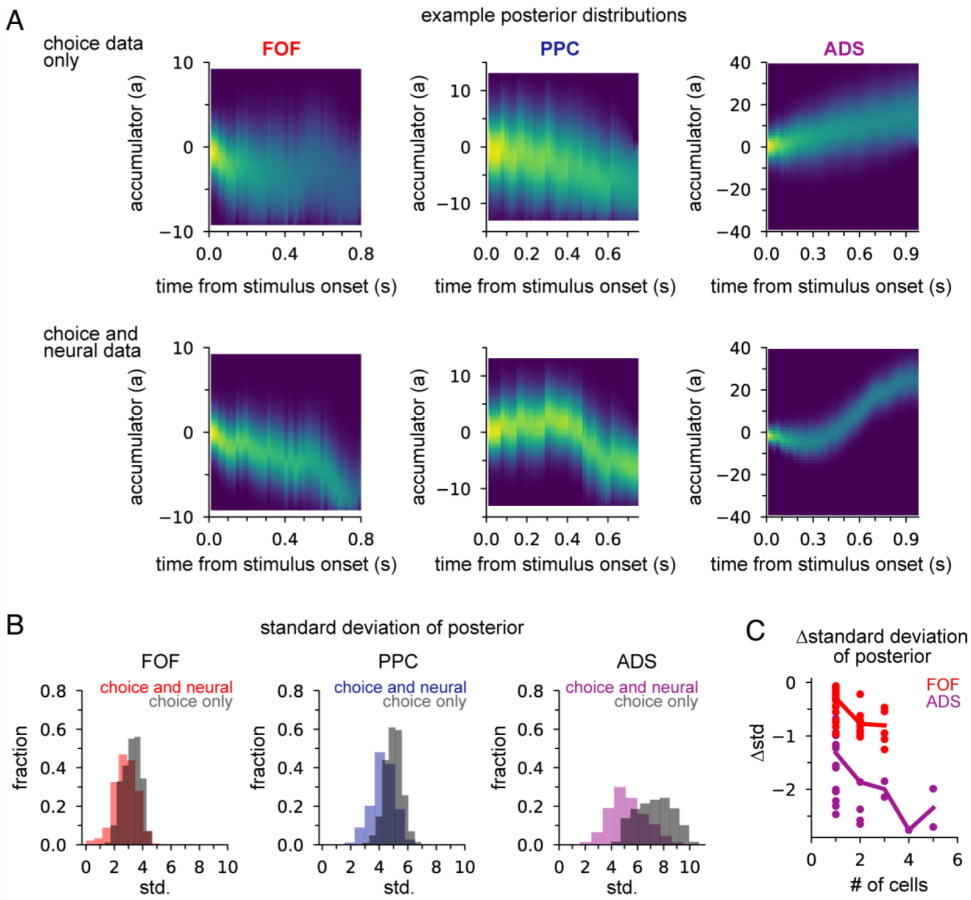


Fig 5: Neural data provides more information about accumulated evidence on single trials than choice alone. (A) Posterior distribution of *a(t)* (excluding captured mass at the boundary) given only the choice (top row) and given spike times and choice (bottom row), for a single example trial. Columns show example trials for different brain regions. (B) Histogram of joint model posterior standard deviations given choice data (*black*) or both neural and choice data (*colors*) for all three brain regions. (C) Difference in choice-conditioned joint posterior standard deviation and neural- and choice-conditioned joint posterior standard deviation as a function of the number of simultaneously recorded neurons. Each point is the difference in the average posterior standard deviation for a session. Negative values indicate that the neural- and choice-conditioned posterior had smaller average standard deviation than the choice-conditioned posterior.

Neural activity improves decoding of upcoming choice

To further test the ability of the joint model to make inferences about the moment-by-moment decision-making process using neural activity, we used the model to decode choices on single trials. We compared the choice decoding accuracy of the joint model to the choice-only model. We found that choices could be predicted more accurately under the joint model, which took into account both the stimulus and neural activity, than under the choice model, which used stimulus information alone. We quantified this improvement in both test log-likelihood and percent correct (**Fig 6A**). The joint model had higher test log-likelihood for choice data and choice prediction accuracy for all three brain regions, with the FOF joint model showing an almost 50% improvement in test log likelihood and a 6% increase in prediction accuracy.

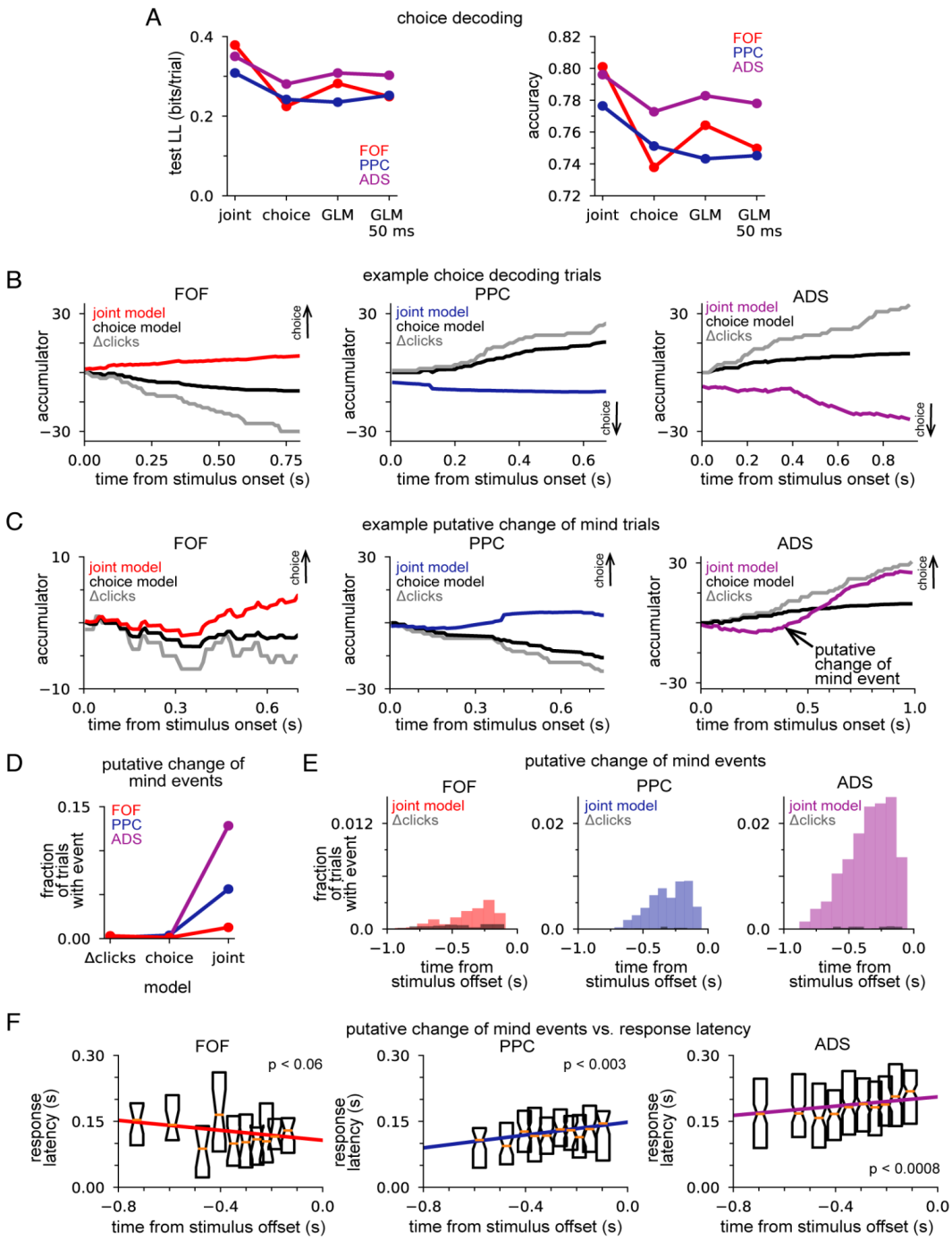


Fig 6: Neural activity improves decoding of choice on a moment-by-moment basis. (A) Choice-prediction accuracy, quantified with log-likelihood (left) and percent correct (right) on test choice data for 4 models: joint neural-behavioral model, choice-only model, and two logistic regression models (see Methods). Values greater than zero indicate that the model can predict choices better than a model that guesses randomly with a probability equal to the empirical distribution of choices. (B) Posterior mean of *a(t)* conditioned on the neural activity for the joint model (*colors*), the distribution of *a(t)* for the choice only model (*black*) and the cumulative click difference (*grey*) for three example trials (one for each brain

region). (C) Putative change-of-mind events, where the posterior mean of the joint model crossed the decision threshold. The corresponding distribution of *a(t)* for the choice only model (*black*) and the cumulative click difference (*grey*) for the same trial are shown for comparison. (D) Fraction of trials that contain at least one putative change-of-mind event for the cumulative click difference, the choice model, and the joint model, for each brain region. (E) Fraction of trials for which a putative change-of-mind event occurs at the specified time relative to the end of the stimulus for the joint model (*color*) and the cumulative click difference (*black*) for each brain region. (F) Response latency plotted as a function of timing of putative change-of-mind events relative to stimulus offset for each brain region. Bar plots show the 25-75 percentiles of the response latency for putative change-of-mind events occurring at similar times. The colored lines indicate the line of best fit for each brain region computed by linear regression.

The posterior mean of *a(t)* of the joint model and the posterior mean of *a(t)* of the choice model is shown in **Figure 6B f**or three example trials. In all examples, the joint model correctly predicted the choice the animal made (indicated by the arrow), whereas the choice-only model failed because its prediction was based on the stimulus. This increased performance derives from the choice-informative spiking information contained in the posterior that the choice model lacks.

We also compared the joint model to a model that predicted choice using temporally summed neural activity. We fit a logistic regression model (i.e. Bernoulli generalized linear model) to the choice using the final accumulated click difference and the trial-summed spike count for each neuron as regressors (Methods). Decoding under the joint model outperformed the logistic regression (**Fig 6A**, *GLM*). The performance of the GLM did not depend strongly on the time window considered. We performed decoding of choice using spikes from the last 50 ms (**Fig 6A**, *GLM* 50 ms), 100 ms, 150 ms, 200 ms and 250 ms before a decision, and found that all had virtually identical performance (**Supp. Fig 5**). This shows that the fine timescale dynamics of the joint model captures features of the spike trains that are useful for predicting the animal's choice, above and beyond the information carried by spike counts in particular time windows before the choice.

Neural activity identified putative changes of mind

Next we took a more in-depth look at the temporal dynamics of the posterior of *a(t)* for the joint model. We identified moments when the posterior mean, conditioned on neural activity, crossed from one side of the decision threshold to the other. We refer to these moments as 'putative change-of-mind' events (Peixoto et al. 2021), reflecting the fact that a model-based decoder of neural activity switches its instantaneous prediction of choice from one option to the other. To be considered an event, the crossing must persist for a certain duration of time, both before and after the crossing, and achieve a sufficiently large magnitude (see Methods for details).

To illustrate the phenomenon, **Fig 6C** shows three example trials with putative change-of-mind events. We also plot the posterior mean for the choice model (*black*) as well as the cumulative sum of the click stimulus (*grey*) for comparison. In all three examples, the joint model crossed the decision threshold, ending on the side corresponding to the animal's actual choice. Sign changes in the accumulated difference in clicks were rare, as were putative change-of-mind events under the choice-only model, both of which could only be caused by the stimulus (**Fig 6D**). In contrast, putative change-of-mind events were observed frequently under the joint

model for all three brain regions (**Fig 6D**). This shows that the putative change-of-mind events reflect information about the accumulator carried in neural activity.

The animal's performance improved on putative change-of-mind event trials (fraction correct: FOF: 0.88 vs. 0.74; PPC: 0.87 vs. 0.74; ADS: 0.85 vs. 0.76; **Supp Fig 6A**). Furthermore, the prediction of the joint model was more accurate at decoding the animal's choice on these trials (fraction correct: FOF: 0.92 vs. 0.80; PPC: 0.88 vs. 0.77; ADS: 0.88 vs. 0.78; **Supp Fig 6**B).

Putative change-of-mind events were more likely to occur at later moments in the trial, usually not long before the stimulus ended (**Fig 6E**). To more firmly connect putative change-of-mind events to the animal's behavior, we performed linear regression to compare the time of the event relative to the end of the stimulus to the response latency (**Fig 6F**). We found a statistically significant effect for the PPC and the ADS (PPC: $p < 0.003$; ADS: $p < 0.0008$; two-sided t-test), while the FOF had a strong trend (FOF: $p < 0.06$). Both PPC and ADS showed a slower response time when a change-of-mind event occurred closer to the end of the stimulus, while the FOF showed the opposite trend. The difference exhibited by the FOF may reflect the unique role that it and other frontal areas play in the decision-making process and is perhaps related to the primacy accumulation strategy it displays as identified by the joint model (**Fig 3**).

Discussion

We developed a probabilistic latent process model to describe the joint distribution of neural activity and choice in a sensory decision-making task, which we fit using data from three different brain regions. Using the model, we found that the dynamics of evidence accumulation differed across brain regions, and from the accumulation model that governed the animal's choice behavior. Finally, we showed that integrating neural activity into these models provides rich, moment-by-moment information about the animal's choice and that this inferred accumulation model can be meaningfully related to other measures of behavior such as reaction time.

One of our most surprising discoveries was that accumulation within the FOF was consistent with a 'primacy' strategy in which early stimulus clicks have an out-size impact on neural activity and choice compared to later clicks (**Fig 3**). Coupled with the low accumulation bound found in FOF, our analysis suggests a model of FOF accumulation where a subject prematurely commits to a decision based on early sensory evidence. Previous analyses of these data identified an alternative interpretation of FOF activity: that neural activity in the FOF exhibited a strongly nonlinear encoding of accumulated evidence consistent with the FOF encoding a categorical representation of intended choice (Hanks et al. 2015). At a strategic level, these interpretations are consistent: because stimulus-induced change-of-mind events are rare in this task (**Fig 6D**) a strongly nonlinear encoding of an unbounded accumulator that does not switch signs will appear very much like an bounded accumulator. Additionally, because these prior analyses did not learn the accumulator variable from the neural activity, only from choices (Hanks et al. 2015), the instability of the FOF accumulator was not identified. Additional experiments (Piet, El Hady, and Brody 2018) and modeling is required to differentiate these two potential models of accumulation at a mechanistic level.

The primacy encoding model of the FOF was also supported by our putative change-of-mind analysis (**Fig 6C-F**) and was both supported by and offers context to prior inactivation studies of the FOF (Erlich et al. 2015). Putative change-of-mind events identified from neural activity occurred less frequently in the FOF than other brain regions, consistent with an early-commitment accumulation strategy in the FOF. Behavioral modeling of choices in conjunction with bilateral pharmacological inactivation found that FOF inactivation led to leakier accumulation when producing choices (Erlich et al. 2015). Leakier accumulation at the level of choice also implies that later stimulus information disproportionately impacts choice, precisely the impact predicted if an early stimulus favoring brain region, such as the FOF, was silenced.

Although a single accumulator model captured some aspects of the ADS data, multiple independent accumulators described the data better (**Fig 4**). Prior work has identified the ADS as a key node in the accumulation network (Yartsev et al. 2018). The independence with which individual ADS neurons encode accumulated evidence may be critical for encoding accumulated evidence with higher fidelity than if their activity strongly covaried.

Prior studies that fit this model to choices developed specific interpretations of the accumulation strategy used by animals (Brunton, Botvinick, and Brody 2013). One difference between choice accumulator models and joint neural-behavior models is the differential impact of accumulator noise versus stimulus noise (**Fig 3**). Choice models have typically indicated that stimulus noise is the primary cause of systematic behavioral uncertainty (Brunton, Botvinick, and Brody 2013), whereas our joint models suggest that this impact is weaker than diffusion noise. One interpretation of this difference is that at the level of a single neural population, diffusive noise plays a stronger role in producing uncertainty in $a(t)$ than stimulus noise, whereas at the level of the entire brain's encoding of accumulated evidence, this diffusive noise 'averages' out and residual stimulus noise remains. Understanding how multiple brain regions work together to produce a model of accumulated evidence at the level of behavior is an important future direction of this work.

Several direct extensions of our framework are readily apparent, and we highlight some of the most interesting directions here. Perhaps most importantly, we illustrated how increasing the number of neural recordings leads to an improved estimate of $a(t)$. As the density of neural recordings from animals performing accumulation tasks increases (Luo et al. 2020), the explanatory power of our model will increase. Additionally, although we have extended the evidence accumulation model to include neural responses and choice, we could easily extend it further to describe additional physiological or behavioral variables (e.g. from annotated video data, or pupil-dilation measurements). Including these additional measures of behavior is likely to further inform the inferred accumulator model, providing a clearer window into the internal factors governing choice behavior across multiple timescales.

Likewise, although we considered a very specific form of an evidence accumulation model due to its normative interpretation, our framework can readily accept modifications and extensions to the dynamical equations that govern evidence accumulation (e.g. (Genkin, Hughes, and Engel 2020)). More extreme discounting of recent evidence, or a more refined specification as to how noise impacts accumulation are two such examples. Additionally, our framework can be modified to consider more elaborate and/or appropriate (given the type of data) relationships

between accumulated evidence and neural responses, such as would be necessary to use calcium imaging data instead of electrophysiological data.

Additionally, although our framework was developed with the specific application to a pulsed-based accumulation task in mind, it is certainly not confined to working only in this context. Our framework can be adapted to any task where noisy temporal accumulation of evidence is thought to play a role, and for which neural recordings and behavioral choice are thought to reflect this process (The International Brain Laboratory et al. 2020).

One additional related extension of the model would be to include additional latent processes unrelated to accumulated evidence (Zoltowski, Pillow, and Linderman 2020), that could be learnable with more flexible methods. While a major motivation of our approach was to develop a framework for identifying a specific normative and mechanistic accumulation model in data, the rigidity of the model makes it difficult to capture varying features present in the data. Including more flexible model components alongside a rigid accumulation model would enable the model to simultaneously account for currently unexplained variance in the data, while preserving the model's ability to account for variance with an accumulation model. Doing so may offer a clearer picture of the evidence accumulation process by sweeping away unrelated variance with a more flexible, but less interpretable, latent process model.

Finally, we have only scratched the surface on relating the activity of multiple independent accumulators to choice, for example by considering a simple model of the ADS where choice is determined by a random accumulator on each trial. Developing a more elaborate probabilistic relationship between multiple accumulators and choice, for example a majority rule procedure for determining choice, within our framework is likely to lead to interesting scientific conclusions.

Acknowledgments

This work was supported by Simons Collaboration on the Global Brain (SCGB AWD543027 and AWD542593), and NIH-NINDS BRAIN Initiative Award (5U19NS104648-02). We thank Michael Yartsev, Tim Hanks, and Charles Kopec for providing the neural data analyzed here. We thank members of the Brody and Pillow labs for comments on this work. B.D. would like to thank Carol Mason, Paul DiMaggio, and Betsy Levy Paluck for their support while this study was conducted.

References

Aoi, Mikio C., Valerio Mante, and Jonathan W. Pillow. 2020. "Prefrontal Cortex Exhibits Multidimensional Dynamic Encoding during Decision-Making." *Nature Neuroscience* 23 (11): 1410–20.

Beck, Jeffrey M., Wei Ji Ma, Roozbeh Kiani, Tim Hanks, Anne K. Churchland, Jamie Roitman, Michael N. Shadlen, Peter E. Latham, and Alexandre Pouget. 2008. "Probabilistic Population Codes for Bayesian Decision Making." *Neuron* 60 (6): 1142–52.

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

Churchland, Mark M., Byron M. Yu, John P. Cunningham, Leo P. Sugrue, Marlene R. Cohen, Greg S. Corrado, William T. Newsome, et al. 2010. "Stimulus Onset Quenches Neural Variability: A Widespread Cortical Phenomenon." *Nature Neuroscience* 13 (3): 369–78.

Ding, L., and J. I. Gold. 2013. "The Basal Ganglia's Contributions to Perceptual Decision Making." *Neuron* 79 (4). https://doi.org/10.1016/j.neuron.2013.07.042.

Ding, Long, and Joshua I. Gold. 2010. "Caudate Encodes Multiple Computations for Perceptual Decisions." *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 30 (47): 15747–59.

Erlich, Jeffrey C., Max Bialek, and Carlos D. Brody. 2011. "A Cortical Substrate for Memory-Guided Orienting in the Rat." *Neuron* 72 (2): 330–43.

Erlich, Jeffrey C., Bingni W. Brunton, Chunyu A. Duan, Timothy D. Hanks, and Carlos D. Brody. 2015. "Distinct Effects of Prefrontal and Parietal Cortex Inactivations on an Accumulation of Evidence Task in the Rat." *eLife* 4 (April). https://doi.org/10.7554/eLife.05457.

Genkin, Mikhail, Owen Hughes, and Tatiana A. Engel. 2020. "Learning Non-Stationary Langevin Dynamics from Stochastic Observations of Latent Trajectories." *arXiv [stat.ML]*. arXiv. http://arxiv.org/abs/2012.14944.

Gold, J. I., and M. N. Shadlen. 2000. "Representation of a Perceptual Decision in Developing Oculomotor Commands." *Nature* 404 (6776): 390–94.

Gold, Joshua I., and Michael N. Shadlen. 2007. "The Neural Basis of Decision Making." *Annual Review of Neuroscience* 30: 535–74.

Hanks, Timothy D., Charles D. Kopec, Bingni W. Brunton, Chunyu A. Duan, Jeffrey C. Erlich, and Carlos D. Brody. 2015. "Distinct Relationships of Parietal and Prefrontal Cortices to Evidence Accumulation." *Nature* 520 (7546): 220–23.

Huk, Alexander C., and Michael N. Shadlen. 2005. "Neural Activity in Macaque Parietal Cortex Reflects Temporal Integration of Visual Motion Signals during Perceptual Decision Making." *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 25 (45): 10420–36.

Kim, J. N., and M. N. Shadlen. 1999. "Neural Correlates of a Decision in the Dorsolateral Prefrontal Cortex of the Macaque." *Nature Neuroscience* 2 (2). https://doi.org/10.1038/5739.

Latimer, Kenneth W., and David J. Freedman. 2021. "Low-Dimensional Encoding of Decisions in Parietal Cortex Reflects Long-Term Training History." *bioRxiv*. https://doi.org/10.1101/2021.10.07.463576.

Latimer, Kenneth W., Jacob L. Yates, Miriam L. R. Meister, Alexander C. Huk, and Jonathan W. Pillow. 2015. "NEURONAL MODELING. Single-Trial Spike Trains in Parietal Cortex Reveal Discrete Steps during Decision-Making." *Science* 349 (6244): 184–87.

Litwin-Kumar, Ashok, Kameron Decker Harris, Richard Axel, Haim Sompolinsky, and L. F. Abbott. 2017. "Optimal Degrees of Synaptic Connectivity." *Neuron* 93 (5): 1153–64.e7.

Luo, Thomas Zhihao, Adrian Gopnik Bondy, Diksha Gupta, Verity Alexander Elliott, Charles D.

Kopec, and Carlos D. Brody. 2020. "An Approach for Long-Term, Multi-Probe Neuropixels Recordings in Unrestrained Rats." *eLife* 9 (October). https://doi.org/10.7554/eLife.59716.

Mante, Valerio, David Sussillo, Krishna V. Shenoy, and William T. Newsome. 2013. "Context-Dependent Computation by Recurrent Dynamics in Prefrontal Cortex." *Nature* 503 (7474): 78–84.

Park, Il Memming, Miriam L. R. Meister, Alexander C. Huk, and Jonathan W. Pillow. 2014. "Encoding and Decoding in Parietal Cortex during Sensorimotor Decision-Making." *Nature Neuroscience* 17 (10): 1395–1403.

Peixoto, Diogo, Jessica R. Verhein, Roozbeh Kiani, Jonathan C. Kao, Paul Nuyujukian, Chandramouli Chandrasekaran, Julian Brown, et al. 2021. "Decoding and Perturbing Decision States in Real Time." *Nature*, January. https://doi.org/10.1038/s41586-020-03181-9.

Piet, Alex T., Ahmed El Hady, and Carlos D. Brody. 2018. "Rats Adopt the Optimal Timescale for Evidence Integration in a Dynamic Environment." *Nature Communications* 9 (1): 4265.

Pillow, Jonathan W., Jonathon Shlens, Liam Paninski, Alexander Sher, Alan M. Litke, E. J. Chichilnisky, and Eero P. Simoncelli. 2008. "Spatio-Temporal Correlations and Visual Signalling in a Complete Neuronal Population." *Nature* 454 (7207): 995–99.

Ratcliff, Roger, and Gail McKoon. 2008. "The Diffusion Decision Model: Theory and Data for Two-Choice Decision Tasks." *Neural Computation* 20 (4): 873–922.

Roitman, Jamie D., and Michael N. Shadlen. 2002. "Response of Neurons in the Lateral Intraparietal Area during a Combined Visual Discrimination Reaction Time Task." *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience* 22 (21): 9475–89.

Shadlen, M. N., and W. T. Newsome. 2001. "Neural Basis of a Perceptual Decision in the Parietal Cortex (area LIP) of the Rhesus Monkey." *Journal of Neurophysiology* 86 (4): 1916–36.

The International Brain Laboratory, Valeria Aguillon-Rodriguez, Dora E. Angelaki, Hannah M. Bayer, Niccolò Bonacchi, Matteo Carandini, Fanny Cazettes, et al. 2020. "Standardized and Reproducible Measurement of Decision-Making in Mice." *Cold Spring Harbor Laboratory*. https://doi.org/10.1101/2020.01.17.909838.

Yartsev, Michael M., Timothy D. Hanks, Alice Misun Yoon, and Carlos D. Brody. 2018. "Causal Contribution and Dynamical Encoding in the Striatum during Evidence Accumulation." *eLife* 7 (August). https://doi.org/10.7554/eLife.34929.

Zoltowski, David M., Kenneth W. Latimer, Jacob L. Yates, Alexander C. Huk, and Jonathan W. Pillow. 2019. "Discrete Stepping and Nonlinear Ramping Dynamics Underlie Spiking Responses of LIP Neurons during Decision-Making." *Neuron* 102 (6): 1249–58.e10.

Zoltowski, David M., Jonathan W. Pillow, and Scott W. Linderman. 2020. "Unifying and Generalizing Models of Neural Dynamics during Decision-Making." *arXiv [q-bio.NC]*. arXiv. http://arxiv.org/abs/2001.04571.

Methods

We model accumulated evidence as a one-dimensional drift diffusion model (DDM) with a symmetric absorbing boundary (Brunton, 2013). On a single behavioral trial, the evolution of the accumulated evidence, $a(t)$, is governed by

$$da = \lambda a dt + \sigma_a dW + \sigma_s dt(\eta' \delta_{t,t_R} C_R(t) - \eta' \delta_{t,t_L} C_L(t)) \,. \tag{1}$$

λ is the inverse of the drift time constant and $\sigma_a dW$ is a Wiener process with scaling σ_a. An absorbing boundary, B, if crossed, prevents $a(t)$ from evolving further (i.e. $da = 0$ if $a(t) > B$).

$\sigma_s \eta'$ are Gaussian variables with variance σ_s^2 and mean 1. δ_{t,t_L} and δ_{t,t_R} are the timing of left and right pulses respectively, and $C_L(t)$ and $C_R(t)$ are the magnitude that each left or right click, respectively, has at time t. The impact of each click is modulated by sensory adaptation, based on the following equation:

$$\frac{dC_\alpha}{dt} = \frac{1 - C_\alpha}{\tau_\phi} + (\phi - 1)(C_\alpha \delta_{t,t_\alpha}) \,, \tag{2}$$

where $\alpha = \{L, R\}$. We define the difference of the adapted click magnitude at time t as

$$\Delta(t) = \delta_{t,t_R} C_R(t) - \delta_{t,t_L} C_L(t) \,, \tag{3}$$

and the sum of the adapted click magnitude at time t as

$$\Sigma(t) = \delta_{t,t_R} C_R(t) + \delta_{t,t_L} C_L(t) \,. \tag{4}$$

By doing so, we can express Equation 1 as,

$$da = \lambda a dt + \Delta(t) dt + \sigma_a dW + \sigma_s \Sigma(t) \eta dt \,, \tag{5}$$

where η is a standard Normal. We define the collection of left and right clicks at time t as $\boldsymbol{\delta}_t$. The initial state of $a(t)$ is distributed normally with mean of 0 and variance of ${\sigma_i}^2$. We refer to all parameters that govern the dynamics of $a(t)$ as $\theta_a = \{\sigma_i, \lambda, B, \sigma_a, \sigma_s, \phi, \tau_\phi\}$

Computing the distribution of the latent state

The temporal dynamics of the probability distribution of $a(t)$, $P(a(t))$, can be expressed as a Fokker-Planck equation,

$$\frac{\partial P(a(t))}{\partial t} = \frac{\sigma_a^2 + \sigma_s^2 \Sigma(t)}{2} \frac{\partial^2 P}{\partial a^2} - \frac{\partial((\lambda a + \Delta(t))P)}{\partial a} \,. \tag{6}$$

We numerically compute the solution to Equation 6 by dividing $P(a(t))$ into a set of n discrete spatial bins, and determine how mass moves after a discrete temporal interval, Δt. The discrete time dynamics of $P(a_t)$ are Markov, and obey the following equation,

$$P(a_t) = M(\theta_a, \boldsymbol{\delta}_t) P(a_{t-1}) \,. \tag{7}$$

The transition matrix $M(\theta_a, \boldsymbol{\delta}_t)$ it is determined using methods established in Brunton (2013). Briefly, for each spatial bin, the deterministic effect of the dynamics on the mass is computed, and this is convolved with a discrete approximation to a Gaussian distribution with the appropriate variance and a finer spatial resolution than the initial spatial resolution described above, to determine the various locations of that mass at the next time bin. Because the location of each bit of mass after the Gaussian convolution is not likely to correspond to the spatial grid defined above, the mass is 'settled' into appropriate bins based on the distance of each bit of mass and the nearest two bins. Mass located in the first and last bin, corresponding to mass that has been captured by the boundary, cannot change locations, and the entries of $M(\theta_a, \boldsymbol{\delta}_t)$ that determines how the mass in these bins moves, reflects this.

$n = 53$ and $\Delta t = 10$ ms for all results presented here.

Relating $a(t)$ to observed data

On a single behavioral trial, the observed spike count of the nth neuron at time t is a Poisson random variable,

$$P(y_{n,t}|a_t, \theta_n) = (f_{\theta_n}(a_t)\Delta t)^{y_{n,t}} \exp(-f_{\theta_n}(a_t)\Delta t) \,, \tag{8}$$

where θ_n defines the function f for the nth neuron.

We choose f_{θ_n} to be a softplus function, i.e. $\text{softplus}(x) = \log(1 + \exp(x))$. Each neuron has their own parameter θ_n that relates f_{θ_n} to a_t. $\theta_{\boldsymbol{y}} = \{\theta_1, \theta_2, \ldots \theta_N\}$ is the collection of all neural parameters for the population of N neurons.

To capture time-varying changes in firing rate that do not depend on $a(t)$, we approximate each neuron's trial-averaged firing rate as a function of time. This is done prior to fitting the full model, i.e. before learning θ_a and θ_y, etc.). We approximate the time-varying average rate with a set of six Gaussian radial basis functions

$$\theta^0_{n,t} = \sum_i^6 w^{RBF}_{in} N(\mu_i, \sigma^2_{RBF}) \,. \tag{9}$$

μ_i and σ^2_{RBF} are set depending on the maximum length of t, i.e. the longest trial. We learn w^{RBF}_{in} by assuming that $y_{n,t}$ is distributed Poisson, with an intensity function $\theta^0_{n,t}$, and maximize the likelihood.

With both parameters for each neuron – θ_n and $\theta^0_{n,t}$ – we define $f_{\theta_n}(a_t)$ as

$$f_{\theta_n}(a_t) = \text{softplus}(\theta_n a_t + \theta^0_{n,t}) \,. \tag{10}$$

On a single behavioral trial, with a probability $1 - \gamma$, the subject's choice, d, is a deterministic function of $a(t)$ at the end of the trial (time T), (Brunton, 2013); with probability γ the choice is made without considering $a(t)$. γ captures "lapses" in the subject's performance. For choices that depend on $a(t)$, if $a(T)$ is greater than a cutoff value c, $d = 1$, otherwise $d = 0$. We refer to the parameters relating $a(t)$ to the likelihood of a subjects choice as $\theta_d = \{c, \gamma\}$.

Thus the probability of the choice, given $a(t)$ and θ_d can be written as,

$$P(d|a_T, \theta_d) = \Big(\frac{\gamma}{2} + (1-\gamma)H(a_T - c)\Big)^d \Big(\frac{\gamma}{2} + (1-\gamma)(1 - H(a_T - c))\Big)^{(1-d)} \,, \tag{11}$$

where $H(\cdot)$ is the Heaviside function.

Relative binning of clicks and spikes

A minor, but key, implementation detail concerns defining the start and end times of the temporal bin edges that are used to bin the click inputs and the spikes trains.

Through numerical experimentation, we identified that our numerical procedure produces a systematic error in estimating the model parameters when the temporal bins for the clicks are aligned with the temporal bins for the spikes.

To circumvent this issue, we offset the bins for the spikes by $\Delta t/2$, so that the bin edges for spikes at time t surrounds the forward bin edge of the clicks by $+/-\Delta t/2$, similarly to the central difference formulation of a finite difference approximation to a differential equation.

Inferring model parameters with maximum likelihood

We refer to the set of all parameters for the model fit to neural and choice data as $\Theta = \{\theta_a, \theta_{\boldsymbol{y}}, \theta_d\}$. Given the Markov dynamics described above, the relationship between $a(t)$ and the observed data, and the model parameters, we can write out the likelihood of the spike train data $\boldsymbol{Y}$ from N neurons and T time bins, the behavioral choice d, and the latent variable $\boldsymbol{a}$ for T time bins as

$$P(\boldsymbol{a}, \boldsymbol{Y}, d|\Theta) = P(a_0|\theta_a) \prod_{t=1}^{T} P(a_t|a_{t-1}, \theta_a, \boldsymbol{\delta}_t) \prod_{n=1}^{N} P(y_{n,t}|a_t, \theta_n)\Big(P\left(d|a_T, \theta_d\right)\Big) \,. \tag{12}$$

We compute the likelihood of the data by integrating over $\boldsymbol{a}$

$$P(\boldsymbol{Y}, d|\Theta) = \sum_{\boldsymbol{a}} P(\boldsymbol{a}, \boldsymbol{Y}, d|\Theta) \,. \tag{13}$$

Because of the way in which we compute $P(a_t|a_{t-1}, \theta_a, \boldsymbol{\delta}_t)$ (see above) computing the log-likelihood of the data can be done with a single forward pass over the data using the 'forward-backward' algorithm method for Hidden Markov models (Bishop, 2006).

We maximize the sum over K behavioral trials of the logarithm of this quantity with respect to Θ via gradient ascent. To compute the gradient of $\sum^K \log P(\boldsymbol{Y}_k, d_k|\Theta)$ with respect to Θ we use a standard automatic differentiation package (Revels, 2016). We refer to the set of parameters that maximizes the likelihood as $\hat{\Theta}$.

We note that all K trials for many of the models we fit were not recorded on the same behavioral session, and therefore, all N neurons are not recorded for every trial. For example, neurons 1-3 might be recorded on trials 1-500, while neurons 4-6 might be recorded on trials 501-1000. Although our notation does not reflect this in order to keep the notation simple, only neurons recorded on a trial contribute to the likelihood on that trial.

Bounded optimization

Several model parameters are only defined within a restricted domain; for example, all variances parameters, such as σ_a^2, are only defined on the positive real axis. Alternatively, other parameters, although defined on a more expansive domain, have values that correspond to models that are not very likely; for example, although B is defined on the positive real axis, values much greater than 40 are not likely to be exhibited in the data, given the specifics of the stimulus, where greater than 40 clicks were rare.

For both of these reasons, we define the following domain over which optimization of the parameters was performed:

- $1e^{-3} \leq \sigma_i^2 \leq 100$
- $8 \leq B \leq 40$
- $-5 \leq \lambda \leq 5$
- $1e^{-3} \leq \sigma_a^2 \leq 400$
- $1e^{-3} \leq \sigma_s^2 \leq 10$
- $1e^{-3} \leq \phi \leq 1.2$
- $5e^{-3} \leq \tau_\phi \leq 1$
- $-10 \leq c \leq 10$
- $0 \leq \gamma \leq 1$
- $-10 \leq \theta_n \leq 10 \;\; \forall n$

Response latency

Previous analyses have identified a response latency between the stimulus and the neural responses, and that this latency can be different in different brain regions (Hanks, 2015). To account for this, we shifted the time of the neural responses relative to the clicks based on these prior results. FOF and ADS responses had a latency of 60 ms, while PPC responses had a latency of 120 ms.

Data pre-selection

To select which neurons were used, a firing rate for each neuron was computed by summing spike counts over the duration of the stimulus period and dividing this by the length of the stimulus period. A two-sided t-test was applied, comparing the firing rate distribution on trials when the animal choose left and when the animal chose right. Neurons with a p-value less than 0.01 were included for analysis.

Data grouping

We grouped together rats that had neural recordings from the same brain region (five FOF rats, three PPC rats, three ADS rats; see **Supp. Table 1** for information about the data) to improve our estimation of the model parameters for each region by using more data. For the PPC and ADS recordings, the majority of recorded neurons came from a single rat, providing further motivation for this modeling approach (**Supp. Table 1**). Although individual FOF rats had enough neurons to support fitting each rat alone, the maximum likelihood parameters for FOF rats fit individually were qualitatively similar (**Supp. Fig 7**), so we fit a single model to all of the FOF rats.

Specifics of data selection for each analyses

Our reports of the maximum likelihood parameters for each model are for models fit to the entire dataset. Each model was also fit using cross-validation (i.e. training on a subset of the data, while reserving data for testing) but the maximum likelihood parameters did not qualitatively change from those identified using the entire dataset, and the log-likelihood computed on test data using parameters identified with training data did not differ appreciably from the log-likelihood computed on those same trials using parameters identified with the entire dataset (**Supp Fig 8**). This consistency is likely due to the modest number of parameters in our models.

When we compute various quantities related to the data, such as peri-stimulus time histograms (PSTHs), the auto-, and cross-correlation functions, and psychometric functions, we likewise use the entire dataset. We did not find that we could accurately estimate the PSTH when only using a small, subset of the data (i.e. test data) due to the fact that our task lacks repeated stimulus conditions.

Additionally, when we stimulate data from a fit model (e.g. **Fig 2A**), we used the maximum likelihood parameters derived from model fits to the entire dataset, and used the stimuli of the entire dataset to generate these data. Again, because the maximum likelihood parameters did not qualitatively change when the model was fit to a subset of the data, we found it easier to focus our analyses on a single model.

The above statements apply to analyses in the following figures: **Fig. 2**, **Fig. 3**, **Fig. 4 C** & **D**, **Fig 5**, **Fig 6 B-F**, and **Supp. Figs 2, 3, 6, 7, 8**.

When comparing performance across models, cross-validation is necessary, and we did do so in those cases (e.g. **Fig 4B**, **Fig 6A** and **Supp. Figs 4, 5**). In these cases, we performed five-fold cross-validation by dividing the dataset into a training set that consisted of 80% of the data and a test set that consisted of 20% of the data. We fit each model using the training data of each fold, and computed the test log-likelihood using the test data and the parameters derived from the training

data. Test performance was averaged across the five folds. Again, we stress that the test performance on cross-validated data did not appreciably differ from that computed using a model trained to the entire dataset (**Supp Fig 8**).

We note, however, that even in cases when we performed cross-validation, we still computed an approximation to each neuron's trial-averaged firing rates, $\theta^0_{n,t}$, using all available data, prior to fitting the full model.

Fitting independent accumulator models

We refer to the set of all parameters for the model with independent accumulators per neuron as Θ_{ind}. The likelihood of the spike train data from the nth neuron $\boldsymbol{Y}_n$ for T time bins is

$$P(\boldsymbol{Y}_n|\Theta_{ind}) = \sum_{\boldsymbol{a_n}} P(a_{0,n}|\theta_a) \prod_{t=1}^{T} P(a_{n,t}|a_{n,t-1}, \theta_a, \boldsymbol{\delta}_t) P(y_{n,t}|a_{n,t}, \theta_n) \,. \tag{14}$$

The joint likelihood for the spike train data from all neurons is the product of the likelihood for each neuron:

$$P(\boldsymbol{Y}|\Theta_{ind}) = \prod_{n=1}^{N} P(\boldsymbol{Y}_n|\Theta_{ind}) \,. \tag{15}$$

Our primary interest in this analysis was capturing the neural responses, so we considered a simple model of choice for the independent accumulator model: on each trial, the choice is determined by randomly selecting one of the accumulators and using its value to determine the choice. The likelihood of the behavioral choice d under such a model is the average of the the n accumulators at time T:

$$P(d|\Theta_{ind}) = \frac{1}{N} \sum_{n=1}^{N} P(d|a_{n,T}, \theta_d) \,. \tag{16}$$

The full likelihood is the product of these terms:

$$P(\boldsymbol{Y}, d|\Theta_{ind}) = P(d|\Theta_{ind}) P(\boldsymbol{Y}|\Theta_{ind}) \,. \tag{17}$$

Fitting to choice data only

We refer to the set of all parameters for the model fit to neural data only as $\Theta_d = \{\theta_a, \theta_d\}$. The likelihood of the behavioral choice d is

$$P(d|\Theta_d) = \sum_{\boldsymbol{a}} P(a_0|\theta_a) \prod_{t=1}^{T} P(a_t|a_{t-1}, \theta_a, \boldsymbol{\delta}_t) \Big(P(d|a_T, \theta_d) \Big) \,. \tag{18}$$

Logistic regression model using spikes

In order to benchmark our method's ability to predict the animal's choice, we considered a basic logistic regression model that included stimulus information and neural activity (e.g. **Fig 6A** and **Supp. Fig 5**). For each trial, we computed the total number of spikes each neuron produced during the specified temporal window and the final cumulative click difference, and used them as regressors in a standard Bernoulli generalized linear model (i.e. logistic regression) to predict the animal's choice. A constant bias regressor was also included, as well as a single lapse parameter that scaled the minimum and maximum values of the logistic inverse link function.

Cross-validation was performed on this model as described above.

Null model of choice

In **Fig 6A**, we assess how well each of our fitted models can predict choice. We compare all models against a baseline model where each choice is a Bernoulli random variable with probability of making a right choice equal to the empirical fraction of choices made to the right.

Joint null model

To compare the improvement of the joint model in absolute terms (i.e. when not comparing two fitted models) we compute a null model of the spiking activity and choices (**Supp Fig 8B**). The null likelihood of the choice data is as described above. The null likelihood of the spike train data assumes that the time-varying expected firing rate of each neuron is equal to its estimated time-varying average firing rate, i.e. $f_{\theta_n}(t) = \theta^0_{n,t}$.

GLM & bounded GLM of neural responses

To validate the maximum likelihood parameters derived from our model, we fit a variant of a Poission GLM to the spiking responses (**Supp Fig 2**). As a regressor, we used the adapated, exponentially filtered click inputs, defined by the following dynamical equation

$$da = \lambda a dt + dt\Delta(t)\,, \tag{19}$$

where $\Delta(t)$ is defined as above. The expected firing rate of each neuron is defined as in the full model, by Equation 10.

For the bounded GLM model, the dynamics of $a(t)$ follow Equation 19, except that if $a(t)$ crosses B, $a(t)$ stops evolving (i.e. $da = 0$ if $a(t) > B$).

The parameters λ, B, ϕ, τ_ϕ, and θ_y that maximize the likelihood of the spike data were learned using gradient ascent. The null model described in **Supp Fig 2** is the same null model for neural responses as used for the joint null model, described above.

Computing PSTHs and auto- and cross-correlation functions on empirical and synthetic data

We computed a 'singe-trial' firing rate for each neuron by convolving its binned spikes with a Gaussian kernel of standard deviation 50 ms. We call this single-trial rate $r_{t,k,n}$ for the nth neuron on

the kth trial at time t. We divide all the trials into two equally-sized groups based on the cumulative click difference at the end of the trial and average $r_{t,k,n}$ based on these groupings. Because trials are not of equal duration, at time t we use whichever trials have data at that time. We refer to this average as $\bar{r}_{c,n,t}$ where the index c runs from 1 to 2.

We generated synthetic data from a model by using the maximum likelihood parameters to generate the expected firing rate of each neuron on each trial, i.e. $f_{t,k,n}$. We averaged this expected rate for each neuron on each trial over 20 different realization of the latent noise to reduce variation due to the latent process. We then grouped and averaged these average expected rates, as described above, to generate a synthetic PSTH, which we denote by $\bar{f}_{c,n,t}$.

We used the empirical and synthetic binned spikes counts to compute auto- and cross-correlation functions. Auto- and cross-correlation functions were normalized by the mean firing rates of the two neurons being used so they provided a measure of excess spike rate. The equation for these functions was,

$$R_{m,n}(\tau) = \frac{1}{m_m}\Big(\frac{1}{K}\sum_k \frac{1}{N_k(\tau)}\sum_t \frac{y_{n,k,t}}{\Delta t}\frac{y_{m,k,t-\tau}}{\Delta t}\Big) - m_n\,, \tag{20}$$

where t is over all bins for the kth trial, $y_{n,k,t}$ and $y_{m,k,t-\tau}$ are the binned spike train of neuron n and m at time t and $t-\tau$ respectively, and $N_k(\tau)$ is the number of bins such that both $y_{n,k,t}$ and $y_{m,k,t-\tau}$ are valid. m_n and m_m are the mean firing rates of the nth and m^{th} neuron respectively.

The auto-and cross-correlations of synthetic data are defined similarly, using the synthetic expected firing rates defined above. Synthetic binned spike counts are generted by sampling Poisson random variables using the expected firing rate, i.e.

$$y^{\text{syn}}_{n,k,t} \sim \text{Poiss}(f_{t,k,n}\Delta t)\,. \tag{21}$$

We have removed the 0th lag component from the auto-correlation plots in **Fig 4**.

Goodness-of-fit metrics

To compare empirical and synthetic PSTHs, we computed the coefficient of determination. Because fewer and fewer trials were included in computing the PSTH at large time values (because trials of great length were rare) we included PSTH values 200 ms before the stimulus onset up until 500 ms after stimulus onset in this calculation.

Based on the definitions of the empirical and synthetic PSTHs, the coefficient of determination is defined as:

$$R^2_n = 1 - \frac{\sum_c \sum_t (\bar{r}_{c,n,t} - \bar{f}_{c,n,t})^2}{\sum_c \sum_t (\bar{r}_{c,n,t} - <\bar{r}_{c,n,t}>_{c,t})}\,, \tag{22}$$

where $<\bar{r}_{c,n,t}>_{c,t}$ is the mean of $\bar{r}_{c,n,t}$ over trial groupings and times.

R^2 values comparing empirical and synthetic auto- and cross-correlation functions were defined similarly, using $R_{m,n}(\tau)$ and $R^{\text{syn}}_{m,n}(\tau)$ as defined above, where a sum is only performed over τ in this case. When computing R^2 values for auto- and cross-correlation functions, we considered values of τ between -800 and 800 ms.

Psychometric functions

We used a Bernoulli GLM (i.e. logistic regression) to compute psychometric functions for empirical and synthetic data. We generated synthetic data from a model by using the maximum likelihood parameters to generate the probability of a choice, and sampled the choice from a Bernoulli distribution.

For the GLM, for each trial, we computed the final click difference and used it as a regressor to predict the animal's choice. A constant bias regressor was also included, as well as a single lapse parameter that scaled the minimum and maximum values of the logistic inverse link function.

R^2 values comparing empirical and synthetic psychometric functions were defined as above, but using the psychometric functions whose domain was from the minimum final cumulative click difference to the maximum final cumulative click difference.

Choice decoding

We used two metrics to determine how well choice could be decoded from various models: choice prediction accuracy and test log likelihood.

Test likelihood was reported in bits per trial, i.e.

$$\Delta LL = \frac{LL_{\text{model}} - LL_{\text{null}}}{\log(2)K}\,, \tag{23}$$

where K is the number of trials in the test set and LL_{null} is the appropriate null model, as described above, or a second model with which to test against. Five-fold cross validation was performed, as described above.

Accuracy was determined, depending on the model, by computing the probability that the model predicted a right choice, given all available data (i.e. inputs, and spikes in a model that includes spikes). If the model had a greater than 0.5 probability of choosing right, we considered that a prediction of a rightward choice. The fraction of choices this prediction is correct is the accuracy.

Identifying putative changes of mind

Based on a recent study (Peixoto, 2021) we defined putative changes in mind in the following way. For each model and each trial, we computed the posterior distribution of $a(t)$ given all available data except for the choice. In the case of the choice only model, this means using only the stimulus, and is equivalent to the forward pass of the model. In the case of the joint model, this is equivalent to the posterior distribution of $a(t)$ given the spikes on that trial.

We then computed the expected value of the posterior distribution for each trial and identified moments when it crossed the decision threshold as determined for each model (i.e. the c parameter of

the choice likelihood). In order to qualify as a true change of mind, we required that the expected value remain on one side of the threshold for 50 ms, remain on the other side following the crossing for 50 ms, and achieve an absolute magnitude greater or equal to 2 at some point during that 100 ms window.

To confirm that these putative change of mind events had some relationship with the animal's behavior we performed linear regression between the time of the event relative to the end of the stimulus (i.e. how close to a decision the event occurred) and a surrogate measure of the animal's reaction time. In this task, the animal is required to fixate in the center poke for the duration of the stimulus, so it does not exhibit a true reaction time in the standard sense of the term as applied to decision-making tasks. However, following the end of the stimulus, it does take the animal some time to withdraw from the center port to makes its choice (see **Fig 1A**, bottom, upper two lines). We refer to the difference between the end of the stimulus and when the animal withdrew from the center port as the animal's reaction time, which we used in our linear regression analysis.

Estimating dimension

To estimate the effective dimension of groups of simultaneously recorded neurons, we computed a metric known as the 'participation ratio' (Litwin-Kumar, 2017). Single-trial firing rates were computed by convolving the spike trains with a Gaussian kernel (std=50 ms), and the covariance matrix of these rates was computed. The partipcation ratio is computed as:

$$\frac{(\sum_n^N \lambda_n)^2}{\sum_n^N \lambda_n^2}\,, \tag{24}$$

where $\boldsymbol{\lambda}$ are the Eigenvalues of the covariance matrix. If the firing rates are independent, the Eigenvalues will all be equal and the participation ratio will equal the number of neurons. If the firing rates are correlated such that some Eigenvalues are small (or perhaps even zero) the participation ratio will be less than the number of neurons.

Confidence intervals for maximum likelihood parameters

We use the Laplace approximation to the log-likelihood to compute confidence bounds for $\hat{\Theta}$ (and $\hat{\Theta}_{ind}$, and $\hat{\Theta}_d$ as shown in **Fig 3** and **Supp Fig 1, 3, 7, 8**).

We compute the Hessian (the matrix of second derivatives) of the log-likelihood at $\hat{\Theta}$ with automatic differentiation. The diagonal entries of the Hessian's inverse quantify the sharpness of the curvature of the log-likelihood surface, and therefore the uncertainty of the estimate of each parameter. We define the confidence bound as +/- two times the square root of each diagonal entry; approximating the log-likelihood surface as Gaussian, this would describe the range of parameters that would fall within approximately 95% of the log-likelihood volume. In cases where these confidence bounds extend beyond optimization bound that correspond to strict boundaries on the domain of a parameter (e.g. variance parameters being strictly positive), we truncate these intervals at the bounds.

In some cases, we found that Hessian was not positive semi-definite, a necessary condition to invert it. This most often occurred when a maximum likelihood parameter encroached upon a strict parameter boundary (e.g. variance parameters being strictly positive). We dealt with these scenarios

in one of two ways.

In some other cases, computing the Hessian in a transformed space (e.g. log space) where these parameters were free to take on any value, rectified the non-concavity. After computing confidence intervals in the transformed space, we mapped these values back into the standard space by the inverse transform.

In other cases, numerical line search along any eigenvector of the Hessian in the standard space corresponding to an eigenvalue with a negative value often confirmed that the convexity of the log-likelihood surface indicated by the negative eigenvalue was local, while more globally, the log-likelihood surface was concave. In light of this, we numerically computed the global concavity of the log-likelihood surface with a numerical line search and approximated this curve with a quadratic function. We replaced any negative eigenvalue of the Hessian with two times the coefficient of this quadratic approximation; the scalar multiplier two is used because entries of the Hessian (having no pre-multiplier) are likewise two times the second-order approximation of the log-likelihood function one would arrive at by Taylor series approximation, where the second-order term contains a 1/2 prefactor.

Code availability

All analysis code was written in the Julia programming language and available upon request.

References

Peixoto, D., Verhein, J.R., Kiani, R. et al. (2021). Decoding and perturbing decision states in real time. *Nature*, ttps://doi.org/10.1038/s41586-020-03181-9.

Litwin-Kumar, A., Harris, K.D., Axel, R., Sompolinsky, H., Abbott, L.F. (2017). Optimal Degrees of Synaptic Connectivity. *Neuron*, 93(5):1153-1164.

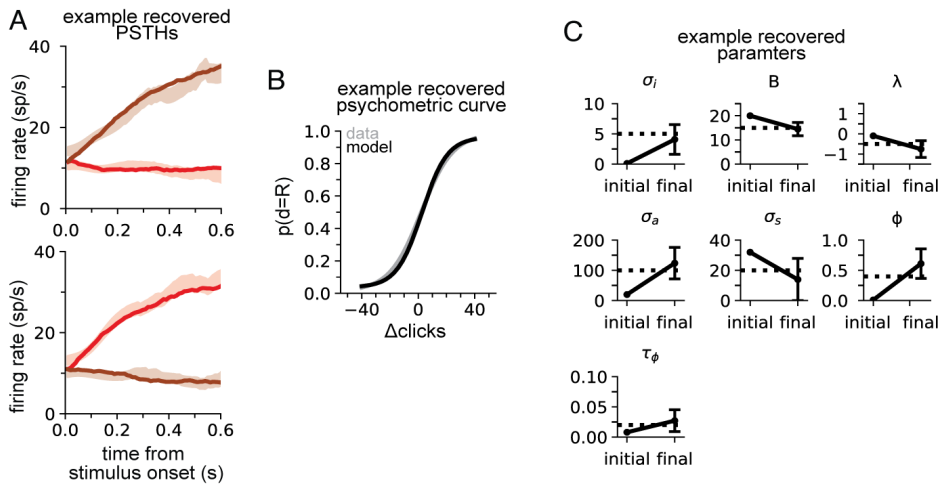
Brunton, B. W., Botvinick, M. M., Brody, C. D. (2013). Rats and Humans Can Optimally Accumulate Evidence for Decision-Making. *Science*, 340(6128):95–98.

Hanks, T. D., Kopec, C. D., Brunton, B. W., Duan, C.A., Erlich, J.C., Brody, C.D. (2015). Distinct Relationships of Parietal and Prefrontal Cortices to Evidence Accumulation. *Nature*, 520 (7546): 220–23.

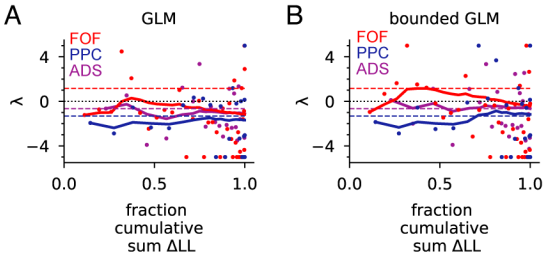
Revels, J., Lubin, M., Papamarko, T. (2016). Forward-Mode Automatic Differentiation in Julia. *arXiv*, http://arxiv.org/abs/1607.07892.

Bishop, C.M. (2006). Pattern Recognition and Machine Learning. *Springer-Verlag*, 0387310738.

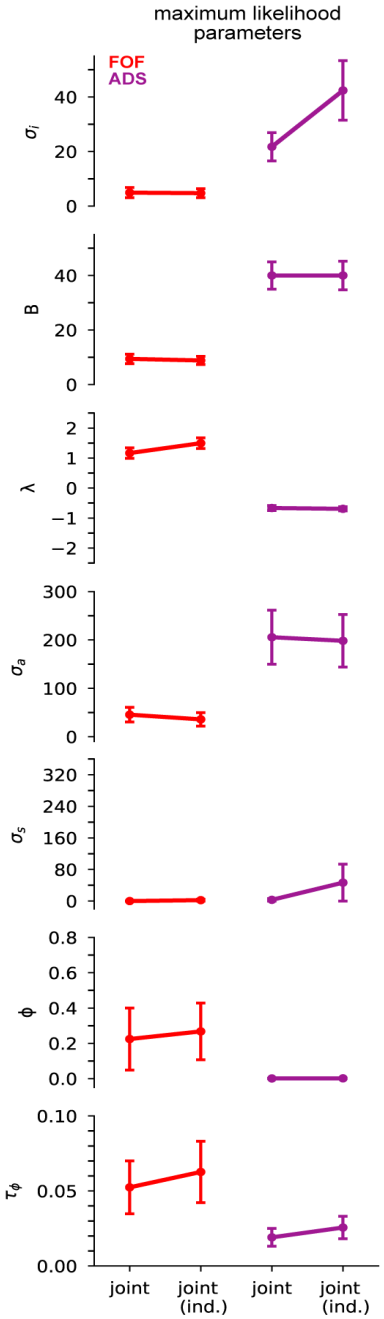
Supplemental Figures



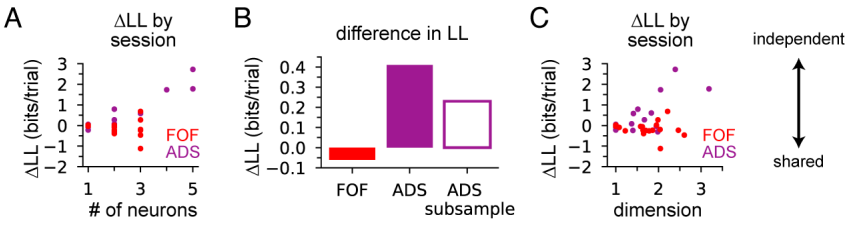
Supplemental Figure 1: Recovering the parameters of synthetic data. Synthetic data was generated with parameters $\sigma_i = 5$, B=15, $\lambda = -0.5$, $\sigma_a = 100$, $\sigma_s = 20$, $\phi = 0.4$, $\tau_\phi = 0.02$. Two synthetic 'sessions' were generated, with 400 trials and 3 neurons each. Softplus gain parameters were randomly generated between -2 and 2. c = 1, $\gamma = 0.05$. (A) PSTHs for two example neurons for synthetic data and simulated data after modeling fitting. (B) Psychometric curves for synthetic data and simulated data after modeling fitting. (C) Optimization was initialized at a random set of parameters ('initial'). Maximum likelihood parameters ('final') converged to within two standard deviations (error bars computed by Laplace approximation) of the parameters used to generate the data (dotted lines).



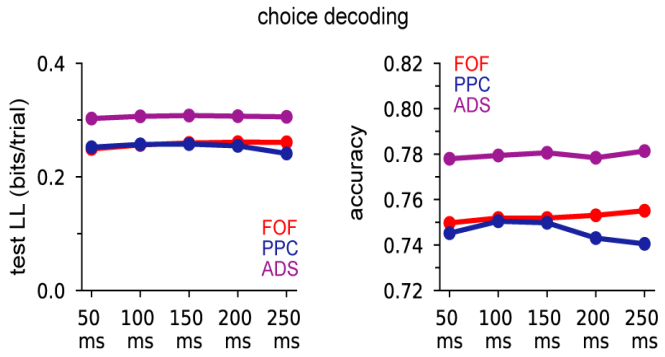
Supplemental Figure 2: GLM analysis of individual sessions. (A) Poisson GLM with a softplus nonlinearity was fitted with exponentially filtered clicks as the regressors (see Methods), using the same data as in **Fig 2** and **Fig 3**. Each dot is the maximum likelihood drift (λ) parameter for a session. Sessions are ordered (from left to right) based on the fraction of the cumulative sum (across all sessions for a brain region) of the change in log likelihood (LL) over the null model (see Methods for null model). For example, the leftmost dot for each brain region is the session with the largest change in LL. Dots on the right were from sessions with the smallest change in LL over the null. The colored lines are the cumulative mean of λweighted by that session's normalized change in LL. Dots on the far right have little change in LL and thus contribute to this mean only weakly. (B) GLM as in (A) but fit with a boundary, such that if the filtered clicks crossed a boundary *B*, the value of the regressors remained equal to *B* henceforth in a trial (see Methods). In each plot, the dashed colored lines are the values of λ from the full model fit (as in **Fig 2** and **Fig 3**).



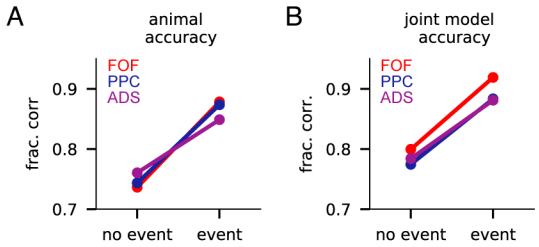
Supplemental Figure 3: Maximum likelihood parameters that govern *a(t)* for the joint (neural/choice) model and independent joint model. Error bars, computed by the Laplace approximation (Methods), are +/- 2 standard deviations. Parameters are σ_i: initial variance, B: accumulation bound, λ: drift, σ_a: accumulation noise variance, σ_s: click noise variance, ϕ: adaptation strength, τ_ϕ: adaptation timescale.



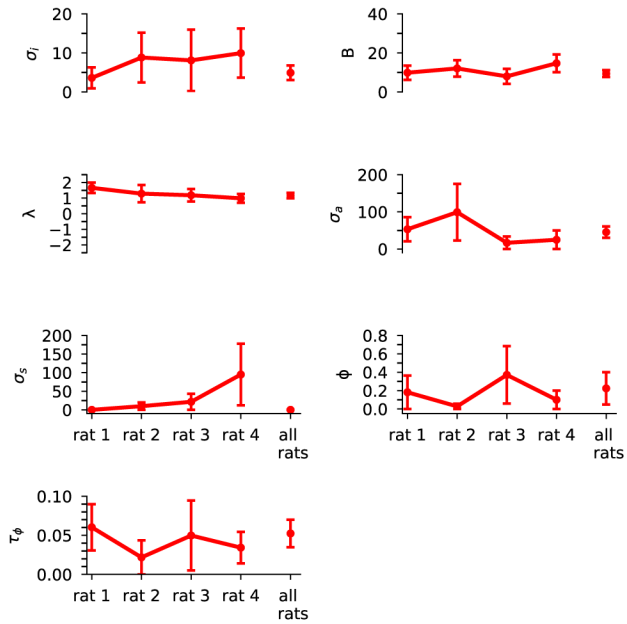
Supplemental Figure 4: ΔLL between the shared and independent accumulator model. (A) Difference in log likelihood for each session for FOF and ADS data plotted as a function of the number of neurons in each session. (B) When the number of neurons in each session for the ADS dataset was subsampled to match the maximum number of neurons in a FOF session (3 neurons) the ADS was still favored by an independent accumulator model (purple, no fill; averaged across 2 subsample permutations of the ADS recordings). (C) Same as (A) but plotted as a function of dimension, as computed by the participation ratio (see Methods). Sessions in the ADS with higher dimension strongly favored the independent accumulator model, leading to the net effect seen in **Figure 4B**.



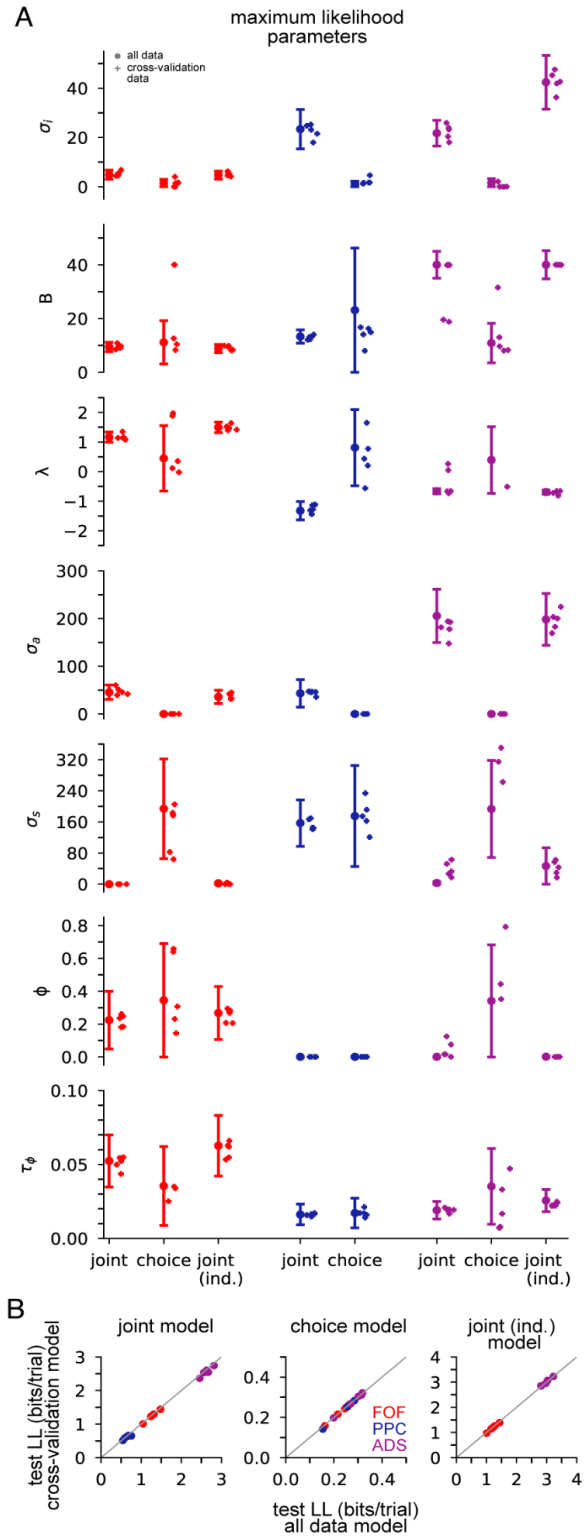
Supplemental Figure 5: GLM choice decoding (as in **Figure 6A**) using spikes in different time windows relative to stimulus offset.



Supplemental Figure 6: Accuracy on putative change-of-mind event trials and non-event trials. (A) Accuracy of the rat for data from each brain region for putative change-of-mind event trials and trials that lacked events ('no event'). (B) Same as (A) but for accuracy of the joint model for each brain region.



Supplemental Figure 7: Maximum likelihood parameters that govern *a(t)* for each FOF rat individually. Error bars, computed by the Laplace approximation (Methods), are +/- 2 standard deviations. Parameters are σ_i: initial variance, B: accumulation bound, λ: drift, σ_a: accumulation noise variance, σ_s: click noise variance, ϕ: adaptation strength, τ_ϕ: adaptation timescale.



Supplemental Figure 8: Comparison of maximum likelihood parameters that govern *a(t)* for three models, joint (neural/choice) model, choice model and independent joint model, when fit to all

data, or using cross-validation data. (A) Circles with error bars are for models fit to all data. Error bars for models fit to full data, computed by the Laplace approximation (Methods), are +/- 2 standard deviations. '+' marks are models (5 for each model type) fit to cross-validation data (5-fold). Parameters are σ_i: initial variance, B: accumulation bound, λ: drift, σ_a: accumulation noise variance, σ_s: click noise variance, ϕ: adaptation strength, τ_ϕ: adaptation timescale. (B) Test log likelihood for models fit to all data (i.e. using trials reserved as testing trials when cross-validation is done) plotted against test log likelihood for cross validation models, for each model type (joint, choice, joint (ind.)), for all three brain regions.

Rat	Region	Sessions	Neurons	Trials	Sessions with greater than 1 neuron	Max. # of simultaneously recorded neurons
B068	FOF	11	13	5859	2	2
T034	FOF	9	10	4138	1	2
T036	FOF	8	12	3026	4	2
T063	FOF	17	32	4002	9	3
T030	FOF	1	1	357	0	1
T035	PPC	15	16	5919	1	2
T011	PPC	7	7	2235	0	1
B053	PPC	2	2	883	0	1
T080	ADS	5	6	1731	1	2
T103	ADS	19	38	8332	9	5
E021	ADS	3	4	697	1	2

Supplemental Table 1: Number of neurons, sessions and trials for each rat.