



Universiteit
Leiden
The Netherlands

Mice alternate between discrete strategies during perceptual decision-making

Ashwood, Z.C.; Roy, N.A.; Stone, I.R.; The International Brain Laboratory; Urai, A.E.; Churchland, A.K.; ... ; Pillow, J.W.

Citation

Ashwood, Z. C., Roy, N. A., Stone, I. R., Urai, A. E., Churchland, A. K., Pouget, A., & Pillow, J. W. (2022). Mice alternate between discrete strategies during perceptual decision-making. *Nature Neuroscience*, 25, 201-212. doi:10.1038/s41593-021-01007-z

Version: Publisher's Version

License: [Licensed under Article 25fa Copyright Act/Law \(Amendment Taverne\)](#)

Downloaded from: <https://hdl.handle.net/1887/3443847>

Note: To cite this publication please use the final published version (if applicable).



Mice alternate between discrete strategies during perceptual decision-making

Zoe C. Ashwood^{1,2}✉, Nicholas A. Roy², Iris R. Stone^{1,2}, The International Brain Laboratory^{*},
Anne E. Urai³, Anne K. Churchland⁴, Alexandre Pouget⁵ and Jonathan W. Pillow^{1,2,6}✉

Classical models of perceptual decision-making assume that subjects use a single, consistent strategy to form decisions, or that decision-making strategies evolve slowly over time. Here we present new analyses suggesting that this common view is incorrect. We analyzed data from mouse and human decision-making experiments and found that choice behavior relies on an interplay among multiple interleaved strategies. These strategies, characterized by states in a hidden Markov model, persist for tens to hundreds of trials before switching, and often switch multiple times within a session. The identified decision-making strategies were highly consistent across mice and comprised a single ‘engaged’ state, in which decisions relied heavily on the sensory stimulus, and several biased states in which errors frequently occurred. These results provide a powerful alternate explanation for ‘lapses’ often observed in rodent behavioral experiments, and suggest that standard measures of performance mask the presence of major changes in strategy across trials.

Understanding the computations performed in the brain will require a comprehensive characterization of behavior^{1–3}. This realization has fueled a recent surge in methods devoted to the measurement, quantification and modeling of natural behaviors^{4–8}. Historically, studies of perceptual decision-making behavior have tended to rely on models derived from signal detection theory^{9,10} or evidence accumulation^{11–13}. More recently, approaches based on reinforcement learning have also been used to model the effects of context, reward and trial history on perceptual decision-making behavior^{14–18}. In all cases, however, these approaches describe decision-making in terms of a single strategy that does not change abruptly across trials or sessions.

One puzzling aspect of sensory decision-making behavior is the presence of so-called ‘lapses’, in which an observer makes an error despite the availability of strong sensory evidence. The term itself suggests an error that arises from a momentary lapse in attention or memory, as opposed to an inability to perceive the sensory stimulus. Lapses arise in all species but are surprisingly frequent in rodent experiments, where lapses can comprise up to 10–20% of all trials^{19–21}.

The standard approach for modeling lapses involves augmenting the classic psychometric curve with a ‘lapse parameter’, which characterizes the probability that the observer simply ignores the stimulus on any given trial^{22–24}. This model can be conceived as a mixture model²⁵ in which, on every trial, the animal flips a biased coin to determine whether or not to pay attention to the stimulus when making its choice. Previous literature has offered a variety of explanations for lapses, including inattention, motor error and incomplete knowledge of the task^{22,24,26}, and recent work has argued that they reflect an active process of uncertainty-guided exploration¹⁷. However, a common thread to these explanations is that lapses arise independently across trials, in a way that does not depend on the time course of other lapses.

Here we show that lapses do not arise independently but, rather, depend heavily on latent states that underlie decision-making

behavior. We use a modeling framework based on hidden Markov models (HMMs) to show that mice rely on discrete decision-making strategies that persist for tens to hundreds of trials. Although the classic lapse model corresponds to a special case in our framework, the model that best described the empirical choice data of real mice had one state corresponding to an ‘engaged’ strategy, in which the animal’s choices were strongly influenced by the sensory stimulus, and other states that corresponded to biased or weakly stimulus-dependent strategies. Our analyses show that lapses arise primarily during long sequences of trials when the animal is in a biased or disengaged state. Conversely, we found that animals with high apparent lapse rates might, nevertheless, be capable of high-accuracy performance for extended blocks of trials. We applied our modeling framework to datasets from two different mouse decision-making experiments^{19,20} and to one dataset from a decision-making experiment in humans²⁷. Taken together, these results shed substantial new light on the factors governing sensory decision-making and provide a powerful set of tools for identifying hidden states in behavioral data.

Results

Classic lapse model for sensory decision-making. A common approach for analyzing data from two-choice perceptual decision-making experiments involves the psychometric curve, which describes the probability that the animal chooses one option (for example, ‘rightward’) as a function of the stimulus value^{22–24}. The psychometric curve is commonly parameterized as a sigmoidal function that depends on a linear function of the stimulus plus an offset or bias. This sigmoid rises from a minimum value of γ , to a maximal value of $1 - \gamma$, where γ , and γ_l denote ‘lapse’ parameters, which describe the probability of making a rightward or leftward choice independent of the stimulus value. Thus, the probability of a rightward choice is always at least γ , and it cannot be greater than $1 - \gamma_l$. In what follows, we will refer

¹Department of Computer Science, Princeton University, Princeton, NJ, USA. ²Princeton Neuroscience Institute, Princeton, NJ, USA. ³Cognitive Psychology Unit, Leiden University, Leiden, Netherlands. ⁴David Geffen School of Medicine, The University of California, Los Angeles, Los Angeles, CA, USA. ⁵Faculty of Medicine & Department of Basic Neurosciences, University of Geneva, Geneva, Switzerland. ⁶Department of Psychology, Princeton University, Princeton, NJ, USA. *A list of authors and their affiliations appears at the end of the paper. ✉e-mail: zashwood@princeton.edu; jwpillow@princeton.edu

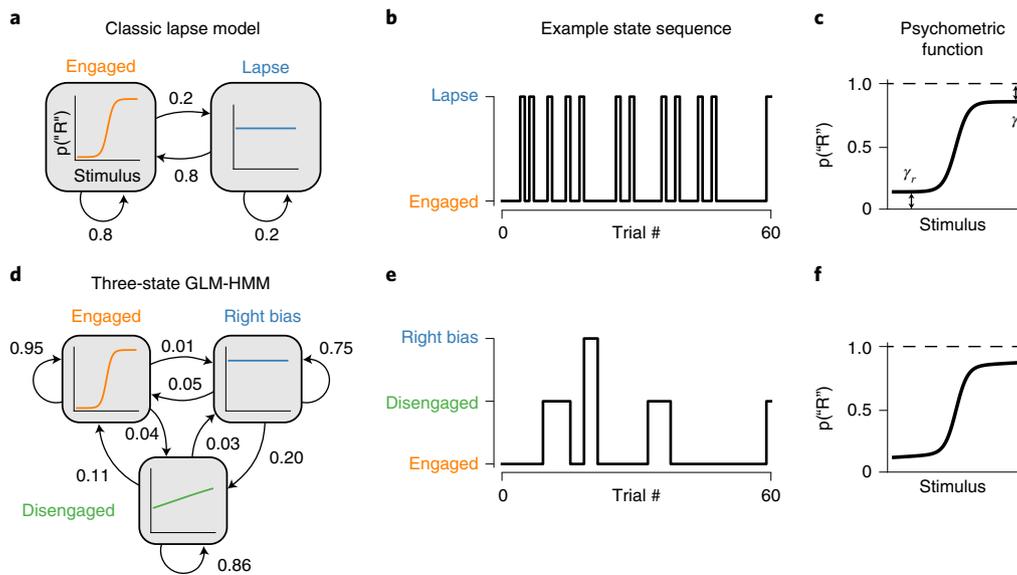


Fig. 1 | The GLM-HMM generalizes the classic lapse model. a, The classic lapse model formulated as a two-state GLM-HMM. Each box represents a GLM describing how the probability of a binary choice depends on the stimulus in the corresponding state: ‘engaged’ (left) or ‘lapse’ (right). Arrows between boxes indicate the transition probabilities between states. Note that the probability of switching to the engaged state at the next trial is always 0.8 (it is independent of state at the current trial), and, similarly, there is always a 0.2 probability of entering the lapse state on each trial. **b**, An example (simulated) sequence for the animal’s internal state when the transitions between states are governed by the probabilities in **a**. Notice that the lapse state tends to last for only a single trial at a time. **c**, Psychometric function arising from the model shown in **a**, depicting the probability of a rightward choice as a function of the stimulus. The parameters γ_r and γ_l denote the probability of a rightward and leftward lapse, respectively. As specified by the transition probabilities in **a**, the total lapse probability for this model is $\gamma_r + \gamma_l = 0.2$. **d**, Example three-state GLM-HMM, with three different GLMs corresponding to different decision-making strategies (labeled ‘engaged’, ‘disengaged’ and ‘right-biased’). Note that these are just example states for the three-state GLM-HMM. In reality, we will learn the states that best describe each animal’s choice data using the process described in the Methods section. The high self-transition probabilities of 0.95, 0.86 and 0.75 ensure that these states typically persist for many trials in a row. **e**, An example sequence for the animal’s internal state sampled from the GLM-HMM shown in **d**. **f**, The psychometric curve arising from the model shown in **d**, which corresponds to a weighted average of the psychometric functions associated with each state. Note that, although the decision-making models shown in **a** and **d** are vastly different, the resulting psychometric curves are nearly identical, meaning that the psychometric curve alone cannot be used to distinguish them.

to this as the ‘classic lapse model of choice behavior’, which can be written as:

$$p(y_t = 1 | \mathbf{x}_t) = (1 - (\gamma_r + \gamma_l)) \frac{1}{1 + e^{-\mathbf{x}_t \cdot \mathbf{w}}} + \gamma_r, \quad (1)$$

where $y_t \in \{0, 1\}$ represents the choice (left or right) that an animal makes at trial t ; $\mathbf{x}_t \in \mathbb{R}^M$ is a vector of covariates; and $\mathbf{w} \in \mathbb{R}^M$ is a vector of weights that describes how much each covariate influences the animal’s choice. Note that \mathbf{x}_t includes both the stimulus and a constant ‘1’ element to capture the bias or offset, but it may also include other covariates that empirically influence choice, such as previous choices, stimuli and rewards^{28–30}.

Although the classic lapse model can be viewed as defining a particular sigmoid-shaped curve relating the stimulus strength to behavior (Fig. 1c), it can equally be viewed as a mixture model²⁵. In this interpretation, we regard the animal as having an internal state z_t that takes on one of two different values on each trial, namely ‘engaged’ or ‘lapse’. If the animal is engaged, it makes its choice according to the classic sigmoid curve (which saturates at 0 and 1). If lapsing, it ignores the stimulus and makes its choice based on only the relative probabilities of a left and right lapse. Mathematically, this can be written as:

$$p(y_t = 1 | \mathbf{x}_t) = \begin{cases} \frac{1}{1 + e^{-\mathbf{x}_t \cdot \mathbf{w}}}, & z_t = \text{“engaged”} \\ \frac{\gamma_r}{\gamma_r + \gamma_l}, & z_t = \text{“lapse”}, \end{cases} \quad (2)$$

where $p(z_t = \text{“lapse”}) = (\gamma_r + \gamma_l)$ and $p(z_t = \text{“engaged”}) = 1 - (\gamma_r + \gamma_l)$. In this interpretation, the animal flips a biased coin, with fixed probability $(\gamma_r + \gamma_l)$, on each trial, and then adopts one of two strategies based on the outcome—a strategy that depends on the stimulus versus one that ignores it. Note that the animal can make a correct choice in the lapse state through pure luck; we use ‘lapse’ here simply to indicate that the animal is not relying on the stimulus when making its decision.

Viewing the classic lapse model as a mixture model highlights some of its limitations. First, it assumes that animals switch between only two decision-making strategies. Second, it assumes that lapses occur independently in time, according to an independent Bernoulli random variable on each trial. Finally, the model assumes that choices in the ‘lapse’ state are fully independent of the stimulus, neglecting the possibility that they are still weakly stimulus dependent²⁴ or are influenced by other covariates such as reward or choice history^{28,29}. These limitations motivate us to consider a more general family of models, which includes the classic lapse model as a special case.

A model for decision-making with multiple strategies.

Recognizing the limitations of the classic lapse model, we propose to analyze perceptual decision-making behavior using a framework based on HMMs with Bernoulli generalized linear model (GLM) observations^{6,31}. The resulting ‘GLM-HMM’ framework, also known as an input–output HMM³², allows for an arbitrary number of states, which can persist for an extended number of trials and exhibit different dependencies on the stimulus and other covariates.

A GLM-HMM has two basic pieces: an HMM governing the distribution over latent states and a set of state-specific GLMs, specifying the decision-making strategy employed in each state (Fig. 1). For a GLM-HMM with K latent states, the HMM has a $K \times K$ transition matrix A specifying the probability of transitioning from any state to any other,

$$p(z_t = k | z_{t-1} = j) = A_{jk}, \quad (3)$$

where z_{t-1} and z_t indicate the latent state at trials $t-1$ and t , respectively. The ‘Markov’ property of the HMM is that the state on any trial depends only on the state from the previous trial, and the ‘hidden’ property refers to the fact that states are latent or hidden from external observers. For completeness, the HMM also has a distribution over initial states, given by a K -element vector π whose elements sum to 1, giving $p(z_1 = k) = \pi_k$.

To describe the state-dependent mapping from inputs to decisions, the GLM-HMM contains K independent Bernoulli GLMs, each defined by a weight vector specifying how inputs are integrated in that particular state. The probability of a rightward choice ($y_t = 1$) given the input vector \mathbf{x}_t and the latent state z_t is given by

$$p(y_t = 1 | \mathbf{x}_t, z_t = k) = \frac{1}{1 + e^{-\mathbf{x}_t \cdot \mathbf{w}_k}}, \quad (4)$$

where $\mathbf{w}_k \in \mathbb{R}^M$ denotes the GLM weights for latent state $k \in \{1, \dots, K\}$. The full set of parameters for a GLM-HMM, $\Theta \equiv \{\pi, A, \{\mathbf{w}_k\}_{k=1}^K\}$, is learned directly from each animal’s choice data via the expectation–maximization (EM) algorithm described in the Methods section.

It is worth noting that the classic lapse model described in Eq. (1) and Eq. (2) corresponds to a restricted two-state GLM-HMM. If we consider state 1 to be ‘engaged’ and state 2 to be the ‘lapse’ state, then the state-1 GLM has weights $\mathbf{w}_1 = \mathbf{w}$, and the state-2 GLM has all weights set to 0 except the bias weight, which is equal to $-\log(\gamma_l/\gamma_r)$. The transition matrix has identical rows, with probability $1 - (\gamma_l + \gamma_r)$ of going into state 1 and probability $(\gamma_l + \gamma_r)$ of going into state 2 at the next trial, regardless of the current state. This ensures that the probability of a lapse on any given trial is stimulus independent and does not depend on the previous trial’s state. Fig. 1a–c shows an illustration of the classic lapse model formulated as a two-state GLM-HMM.

However, there is no general reason to limit our analyses to this restricted form of the GLM-HMM. By allowing the model to have more than two states, multiple states with non-zero stimulus weights and transition probabilities that depend on the current state, we obtain a model family with a far richer set of dynamic decision-making behaviors. Figure 1d shows an example GLM-HMM with three latent states, all of which have high probability of persisting for multiple trials. Intriguingly, the psychometric curve arising from this model (Fig. 1f) is indistinguishable from that of the classic lapse model. Thus, multiple generative processes can result in identical psychometric curves, and we must look beyond the psychometric curve if we want to gain insight into the dynamics of decision-making across trials.

Mice switch between multiple strategies. To examine whether animals employ multiple strategies during decision-making, we fit the GLM-HMM to behavioral data from two binary perceptual decision-making tasks. First, we fit the GLM-HMM to choice data from 37 mice performing a visual detection decision-making task developed in ref. ³³ and adopted by the International Brain Laboratory (IBL)¹⁹. During the task, a sinusoidal grating with contrast between 0% and 100% appeared either on the left or right side of the screen (Fig. 2a). The mouse had to indicate this side by turning a wheel. If the mouse turned the wheel in the correct direction, it

received a water reward; if incorrect, it received a noise burst and an additional 1-s timeout. During the first 90 trials of each session, the stimulus appeared randomly on the left or right side of the screen with probability 0.5. Subsequent trials were generated in blocks in which the stimulus appeared on one side with probability 0.8, alternating randomly every 20–100 trials. We analyzed data from animals with at least 3,000 trials of data (across multiple sessions) after they had successfully learned the task (Supplementary Figs. 1 and 2). For each animal, we considered only the data from the first 90 trials of each session, when the stimulus was equally likely to appear on the left or right of the screen.

We modeled the animals’ decision-making strategies using a GLM-HMM with four inputs: (1) the (signed) stimulus contrast, where positive values indicate a right-side grating and negative values indicate a left-side grating; (2) a constant offset or bias; (3) the animal’s choice on the previous trial; and (4) the stimulus side on the previous trial. A large weight on the animal’s previous choice gives rise to a strategy known as ‘persistence’ in which the animal makes the same choice many times in a row, regardless of whether it receives a reward. A large weight on the previous stimulus side, which we refer to as the ‘win-stay, lose-switch’ regressor, gives rise to the well-known strategy in which the animal repeats a choice if it was rewarded and switches choices if it was not. Note that, for the IBL task in question, bias and trial history dependencies were suboptimal, meaning that the maximal reward strategy was to have a large weight on the stimulus and zero weights on the other three inputs.

To determine the number of different strategies underlying decision-making behavior, we fit GLM-HMMs with varying numbers of latent states. Note that the one-state model is simply a standard Bernoulli GLM, whereas the classic lapse model (Eq. (1) and Eq. (2)) is the constrained two-state model mentioned previously. We found that a three-state GLM-HMM substantially outperformed models with fewer states, including the classic lapse model. The states of the fitted model were readily interpretable and tended to persist for many trials in a row.

Figures 2 and 3 show results for an example mouse. For this animal, the multi-state GLM-HMM outperformed both the standard (one-state) GLM and the classic lapse model, both in test log-likelihood and percent correct, with the improvement approximately leveling off at three latent states (Fig. 2b,c). Because the test set for this mouse contained 900 trials, test set log-likelihood increases of 0.13 bits per trial and 0.09 bits per trial for the three-state model over the one-state model and classic lapse models, respectively, meant that the data were $(2^{0.13})^{900} \approx 1.7 \times 10^{35}$ and $(2^{0.09})^{900} \approx 2.4 \times 10^{24}$ times more likely under the three-state model. Note that, whereas the number of model parameters increased with the number of states in the GLM-HMM (and a two-state GLM-HMM has more parameters than both the classic lapse model or single-state GLM), the log-likelihood of the test dataset was not guaranteed to increase with more states³⁴. Indeed, as the number of states increased, the GLM-HMM could have started to overfit the training dataset, causing it to fit the test dataset poorly (Extended Data Fig. 1).

The transition matrix for the fitted three-state model describes the transition probabilities among three different states, each of which corresponds to a different decision-making strategy (Fig. 2d). Large entries along the diagonal of this matrix, ranging between 0.94 and 0.98, indicate a high probability of remaining in the same state for multiple trials. The other set of inferred parameters was the GLM weights, which define how the animal makes decisions in each state (Fig. 2e). One of these GLMs (‘state 1’) had a large weight on the stimulus and negligible weights on other inputs, giving rise to high-accuracy performance on the task (Fig. 2f). The other two GLMs (‘state 2’ and ‘state 3’), by comparison, had smaller weights on the stimulus and relatively large bias weights.

We can visualize the decision-making strategies associated with these states by plotting the corresponding psychometric curves

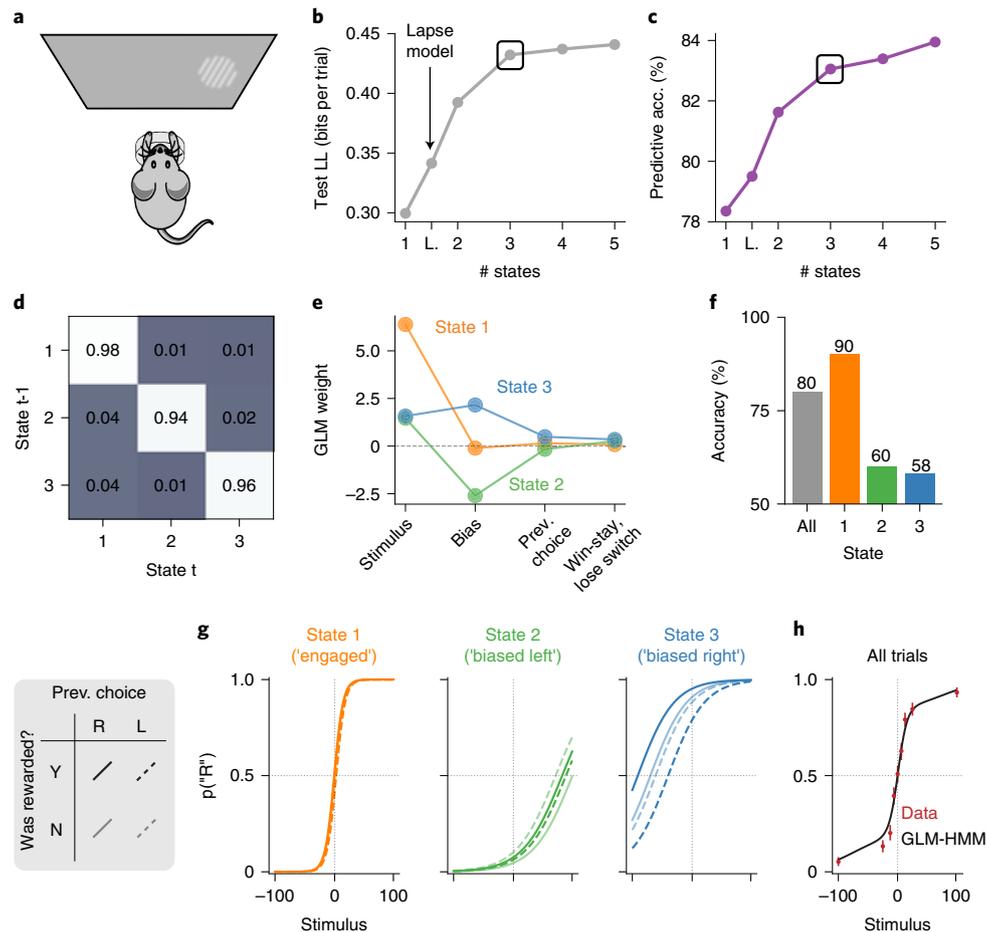


Fig. 2 | GLM-HMM analysis of choice behavior of an example IBL mouse. a, Schematic for visual decision-making task, which involved turning a wheel to indicate whether a Gabor patch stimulus appeared on the left or right side of the screen^{19,33}. **b**, Model comparison among GLM-HMMs with different numbers of latent states, as well as the classic lapse model (labeled 'L') using five-fold cross-validation. Test set log-likelihood is reported in units of bits per trial and was computed relative to a 'null' Bernoulli coin-flip model (see the Methods section). A black rectangle highlights the log-likelihood for the three-state model, which we used for all subsequent analyses. **c**, Test set predictive accuracy for each model, indicating the percent of held-out trials where the model successfully predicted the mouse's true choice. **d**, Inferred transition matrix for best-fitting three-state GLM-HMM for this mouse. Large entries along the diagonal indicate a high probability of remaining in the same state. **e**, Inferred GLM weights for the three-state model. State 1 weights have a large weight on the stimulus, indicating an 'engaged' or high-accuracy state. In states 2 and 3, the stimulus weight is small, and the bias weights give rise to large leftward (state 2) and rightward (state 3) biases. **f**, Overall accuracy of this mouse (gray) and accuracy for each of the three states. **g**, Psychometric curve for each state, conditioned on previous reward and previous choice. **h**, The standard psychometric curve for all choice data from this mouse, which can be seen to arise from the mixture of the three per-state curves shown in **g**. Using the fit GLM-HMM parameters for this animal and the true sequence of stimuli presented to the mouse, we generated a time series with the same number of trials as those that the example mouse had in its dataset. At each trial, regardless of the true stimulus presented, we calculated $p_i("R")$ for each of the nine possible stimuli by averaging the per-state psychometric curves of **g** and weighting by the appropriate row in the transition matrix (depending on the sampled latent state at the previous trial). Finally, we averaged the per-trial psychometric curves across all trials to obtain the curve that is shown in black, whereas the empirical choice data of the mouse are shown in red, as are 95% confidence intervals (n between 530 and 601, depending on stimulus value). acc., accuracy.

(Fig. 2g), which show the probability of a rightward choice as a function of the stimulus, conditioned on both previous choice and reward. The steep curve observed in state 1, which corresponds to the mouse achieving near-perfect performance on high-contrast stimuli, led us to adopt the name 'engaged' to describe this state. By comparison, the psychometric curves for states 2 and 3 reflected large leftward and rightward biases, respectively. They also had relatively large dependence on previous choice and reward, as indicated by the gap between solid and dashed lines. Although this mouse had an overall accuracy of 80%, it achieved 90% accuracy in the engaged state compared to only 60% and 58% accuracy in the two biased states (Fig. 2f).

To gain insight into the temporal structure of decision-making behavior, we used the fitted three-state model to compute the

posterior probability over the mouse's latent state across all trials (Fig. 3). The resulting state trajectories reflect our posterior beliefs about the animal's internal state on every trial, given the entire sequence of observed inputs and choices during a session (see Methods).

Contrary to the predictions of the classic lapse model, states tended to persist for many trials in a row. Remarkably, we found that the most probable state often had probability close to 1, indicating a high degree of confidence about the mouse's internal state given the observed data.

To quantify state occupancies, we assigned each trial to its most probable state and found that this example mouse spent approximately 69% of all trials in the engaged state (out of 5,040 total

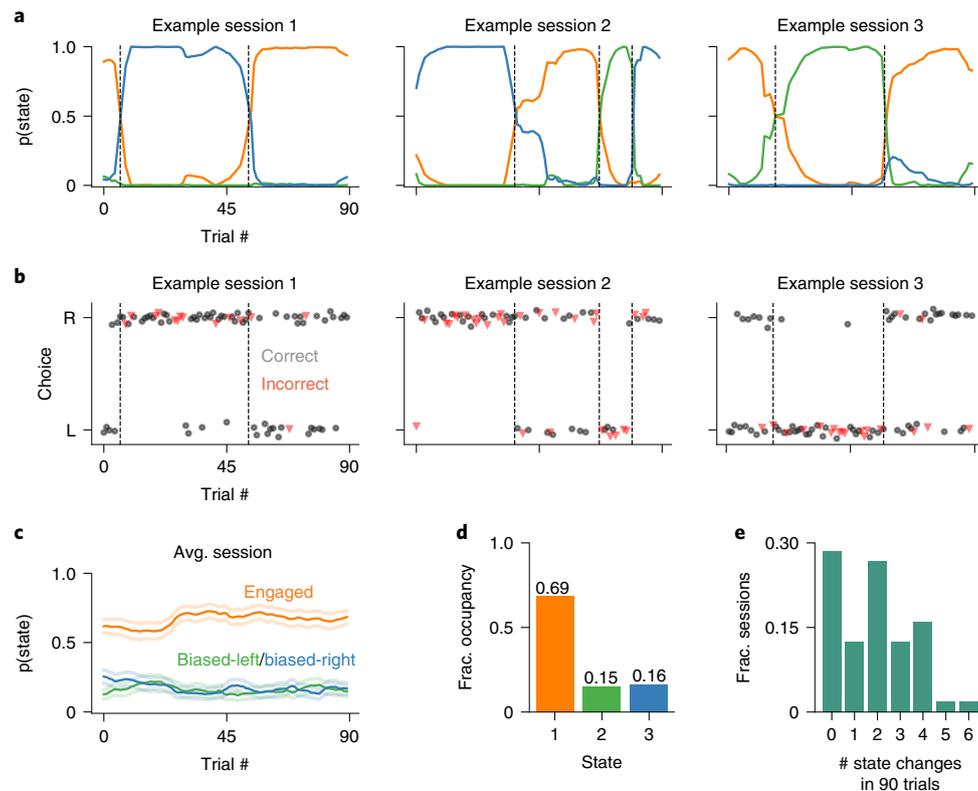


Fig. 3 | Inferred state dynamics for example IBL mouse. **a**, Posterior state probabilities for three example sessions, revealing high levels of certainty about the mouse's internal state and showing that states typically persisted for many trials in a row. **b**, Raw behavioral data corresponding to these three sessions. Each dot corresponds to a single trial; the x position indicates the trial within the session, whereas the y position indicates if the mouse went rightward or leftward on the trial (y positions are slightly jittered for visibility). The dot's color and shape indicates if the trial was an error or correct trial (red triangles correspond to errors, whereas black circles represent correct choices). All trials except 0% contrast trials (for which the correct answer is determined randomly) are shown here. The dashed vertical lines correspond to the location of state changes (obtained using the posterior probabilities shown in **a**). **c**, Average trajectories of state probabilities within a session, computed over 56 sessions. Error bars indicate ± 1 s.e.m. **d**, Fractional occupancies for the three states across all trials. For this analysis, we assigned each trial to its most likely state and then counted the fraction of trials assigned to each state. Overall, the mouse spent 69% of all trials in the engaged state and 31% of trials in one of the two biased states. **e**, Histogram showing the number of inferred state changes per session for all 56 sessions of data for this mouse. Only 29% of sessions involved the mouse persisting in a single state for the entire 90-trial session. Frac., fractional.

trials over 56 sessions) compared to 15% and 16% of trials in the biased-leftward and biased-rightward states (Fig. 3d). Moreover, the mouse changed state at least once within a session in roughly 71% of all 90-trial sessions and changed multiple times in 59% of sessions (Fig. 3e). This rules out the possibility that the states merely reflect the use of different strategies on different days. Rather, the mouse tended to remain in an engaged, high-performance state for tens of trials at a time, with lapses arising predominantly during interludes when it adopted a left-biased or right-biased strategy for multiple trials in a row. The multi-state GLM-HMM thus provides a very different portrait of mouse decision-making behavior than the basic GLM or lapse model.

State-based strategies are consistent across mice. To assess the generality of these findings, we fit the GLM-HMM to the choice data from 37 mice in the IBL dataset¹⁹ (Fig. 4). We found that the results shown for the example animal considered above were broadly consistent across animals. Specifically, we found that the three-state GLM-HMM strongly outperformed the basic GLM and classic lapse model in cross-validation for all 37 mice (Fig. 4a and Supplementary Fig. 3). On average, it predicted mouse choices with 4.2% higher accuracy than the basic GLM (which had an average prediction accuracy of 78%) and 2.8% higher accuracy than the

classic lapse model (Fig. 4b). Furthermore, for one animal, the improvement in prediction accuracy for the three-state GLM-HMM was as high as 12% relative to the basic GLM and 7% relative to the classic lapse model.

Although performance continued to improve slightly with four and even five latent states, we will focus our analyses on the three-state model for reasons of simplicity and interpretability. Supplementary Fig. 4 provides a full analysis of the four-state model, showing that it tended to divide the engaged state from the three-state model into two 'substates' that differed slightly in bias.

Fits of the three-state GLM-HMM exhibited remarkable consistency across mice, with most exhibiting states that could be classified as 'engaged', 'biased-left' and 'biased-right' (Fig. 4d) (see Methods section for details about the alignment of states across mice). Whereas we plotted inferred transition matrices for all 37 mice in Supplementary Fig. 5, here we used the diagonal elements of each matrix to compute an expected dwell time for each animal in each state (Fig. 4e). This revealed a median dwell time across animals of 24 trials for the engaged state versus 13 and 12 trials for the biased-left and biased-right states, respectively. This description of behavior departs markedly from the assumptions of the classic lapse model: for the classic lapse model and a lapse rate of 20%, the expected dwell time in the lapse state is just $1 / (1 - 0.2) = 1.25$ trials.

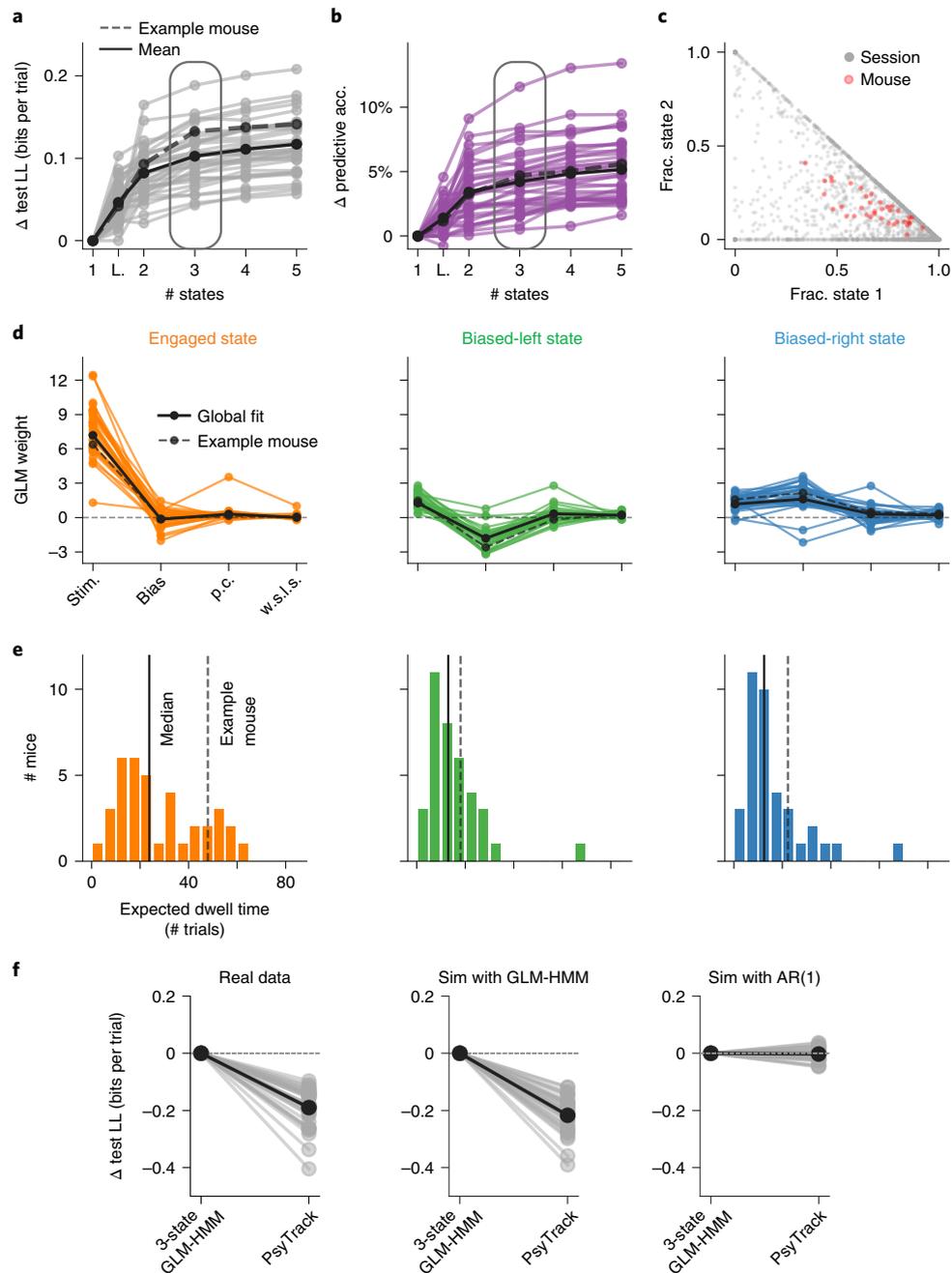


Fig. 4 | Model fits to full population of 37 IBL mice. **a**, Change in test log-likelihood as a function of number of states relative to a (one-state) GLM, for each mouse in the population. The classic lapse model, a restricted form of the two-state model, is labeled 'L'. Each trace represents a single mouse. Solid black indicates the mean across animals, and the dashed line indicates the example mouse from Figs. 2 and 3. The rounded rectangle highlights performance of the three-state model, which we selected for further analyses. **b**, Change in predictive accuracy relative to a basic GLM for each mouse, indicating the percentage improvement in predicting choice. **c**, Gray dots correspond to 2,017 individual sessions across all mice, indicating the fraction of trials spent in state 1 (engaged) and state 2 (biased left). Points at the vertices (1,0), (0,1) or (0,0) indicate sessions with no state changes, whereas points along the sides of the triangle indicate sessions that involved only two of the three states. Red dots correspond to the same fractional occupancies for each of the 37 mice, revealing that the engaged state predominated but that all mice spent time in all three states. **d**, Inferred GLM weights for each mouse, for each of the three states in the three-state model. The solid black curve represents a global fit using pooled data from all mice (Algorithm 1); the dashed line is the example mouse from Figs. 2 and 3. **e**, Histogram of expected dwell times across animals in each of the three states, calculated from the inferred transition matrix for each mouse. **f**, Mice have discrete, not continuous, decision-making states. Left: cross-validation performance of the three-state GLM-HMM compared to PsyTrack^{35,36} for all 37 mice studied (each individual line is a separate mouse; black is the mean across animals). Middle: As a sanity check, we simulated datasets from a three-state GLM-HMM, with the parameters for each simulation chosen as the best-fitting parameters for a single mouse. We then fit the simulated data both with PsyTrack and with the three-state GLM-HMM to check that the three-state GLM-HMM best described the data. Right: We did the opposite and fit PsyTrack to the animals' data and then generated data according to an AR(1) model with parameters specified using the PsyTrack fits (see Methods section for full details). By performing cross-validation on the simulated data, we confirmed that we could use model comparison to distinguish between discrete and continuous decision-making behavior in choice data. acc., accuracy; Frac., fractional; stim., stimulus; p.c., previous choice; w.s.l.s., win-stay lose-switch; Sim, simulated data.

We analyzed the distribution of state dwell times inferred from data and found that they were well approximated by a geometric distribution, matching the theoretical distribution of data sampled from an HMM (Extended Data Fig. 2 and Supplementary Fig. 6).

We also examined the fraction of trials per session that mice spent in each of the three states (Fig. 4c). To do so, we used the fitted model parameters to compute the posterior probabilities over state and assigned each trial to its most likely state. The resulting ‘fractional occupancies’ revealed that the median mouse spent 69% of its time in the engaged state, with the best mice exceeding 90% engagement. Moreover, most sessions (83% of 2,017 sessions across all animals) involved a switch between two or more states; in only 17% of sessions did a mouse remain in the same state for an entire 90-trial session.

To further quantify the model’s ability to capture temporal structure of choice data, we examined choice run length statistics, where runs were defined as sequences of trials in which the mouse repeated the same choice. We found that the IBL mice exhibited longer runs than if they performed the task perfectly, and the fitted GLM-HMM reproduced this feature more accurately than the classic model (Extended Data Fig. 8).

Finally, to make sure that the states identified by the model were not a consequence of the IBL mice having been previously exposed to blocks of trials with a consistent side bias, we examined data from four mice that were never exposed to ‘bias blocks’. In Extended Data Fig. 3, we show that the retrieved states, dwell times and model comparison results for these animals were similar to those shown in the rest of the IBL population.

Data provide evidence for discrete, not continuous, states. The GLM-HMM describes perceptual decision-making in terms of discrete states that persist for many trials in a row before switching. However, the model’s marked improvement over classic models does not guarantee that the states underlying decision-making are best described as discrete. One could imagine, for example, that a state governing the animal’s degree of engagement drifts gradually over time and that the GLM-HMM simply divides these continuous changes into discrete clusters. To address this possibility, we fit the data with PsyTrack, a psychophysical model with continuous latent states^{35,36}. The PsyTrack model describes sensory decision-making using an identical Bernoulli GLM but with dynamic weights that drift according to a Gaussian random walk (see Methods section).

For all 37 mice in our dataset, the three-state GLM-HMM achieved substantially higher test log-likelihood than the PsyTrack model (Fig. 4f, left). Model selection also correctly identified simulated data from the GLM-HMM, whereas datasets simulated from a matched first-order autoregressive model had roughly equal log-likelihood under the two models (Fig. 4f, middle and right). Not only was the GLM-HMM favored over PsyTrack for IBL mice, but we show that it was favored over PsyTrack for the other mouse datasets that we studied in this work (Extended Data Fig. 4). Finally, to verify that the GLM-HMM’s improved performance compared to PsyTrack was not due merely to the fact that the GLM-HMM does better at accounting for lapses, we also show the average fits of both PsyTrack and the three-state GLM-HMM to the empirical choice data of the example mouse in Supplementary Fig. 7. Despite the absence of explicit lapse parameters in both of these models, both are able to capture the non-zero error rates of the mouse on easy trials.

Mice switch between multiple strategies in a second task. To ensure that our findings were not specific to the IBL task or training protocol, we examined a second mouse dataset with a different sensory decision-making task. Odoemene et al.²⁰ trained mice to report whether the flash rate of an LED was above or below 12 Hz by making a right or left nose poke (Fig. 5a). Once again, we found that the

multi-state GLM-HMM provided a far more accurate description of mouse decision-making than a basic GLM or the classic lapse model (Fig. 5b). Although the performance of the three-state and four-state models was similar, we focused on the four-state model because—in addition to having slightly higher test log-likelihood for most animals (see Supplementary Fig. 8 for individual curves)—the four-state model balanced simplicity and interpretability, with each state in the four-state model corresponding to a distinct behavioral strategy (see Supplementary Figs. 9 and 10 for a comparison to three-state and five-state fits). The four-state model exhibited an average improvement of 0.025 bits per trial over the classic lapse model, making the test dataset approximately 1×10^{18} times more likely under the GLM-HMM than the lapse model.

Figure 5d shows the inferred GLM weights associated with each state in the four-state GLM-HMM. Based on the qualitative differences in the state-specific psychometric curves (Fig. 5e), we labeled the four states as ‘engaged’, ‘biased left’, ‘biased right’ and ‘win-stay’. The combination of stimulus and choice history weights for this fourth state gave rise to a large separation between psychometric curves conditioned on previous reward; the resulting strategy could be described as ‘win-stay’ because the animal tended to repeat a choice if it was rewarded. (However, it did not tend to switch if a choice was unrewarded). Accuracy was highest in the engaged state (92%) and lowest in the biased-left (67%) and biased-right states (77%). Similarly to the IBL dataset, the identified states tended to persist for many trials in a row before switching (Fig. 5f and Extended Data Fig. 5b).

Finally, we used the fitted model to examine the temporal evolution of latent states within a session. Figure 5g shows the average posterior state probabilities over the first 200 trials in a session for two example mice and the average over all mice (Extended Data Fig. 6 shows average posterior state probabilities for each individual mouse in the cohort separately). These trajectories reveal that mice typically began a session in one of the two biased states and had a low probability of entering the engaged state within the first 50 trials: mice used these initial trials of a session to ‘warm up’ and then gradually improved their performance¹⁴. This represents a marked departure from the IBL mice, most of which had a high probability of engagement from the very first trial and had relatively flat average trajectories over the first 90 trials of a session (Fig. 3 and Supplementary Fig. 12). Note, however, that the warm-up effect was not the only form of state switching observed in the Odoemene et al. mice, as the animals switched states multiple times per session and continued switching beyond the first 100 trials (Extended Data Fig. 5). We also examined whether the effects of fatigue or satiety could be observed in the average state probabilities at the end of sessions, but we did not find consistent patterns across animals (Extended Data Fig. 7).

External correlates of engaged and disengaged states. One powerful feature of the GLM-HMM is the fact that it can be used to identify internal states from binary decision-making data alone. However, it is natural to ask whether these states manifest themselves in other observable aspects of mouse behavior. In other words, does a mouse behave differently when it is in the engaged state than in a biased state, above and beyond its increased probability of making a correct choice? To address this question, we examined response times and violation rates, two observable features that we did not incorporate into the models.

Previous literature has revealed that humans and monkeys commit more errors on long-duration trials than on short-duration trials^{37–39}. This motivated us to examine the distributions of response times for the engaged state compared to the disengaged (left-biased and right-biased) states.

In the IBL dataset, we found that engaged and disengaged response time distributions were statistically different for all

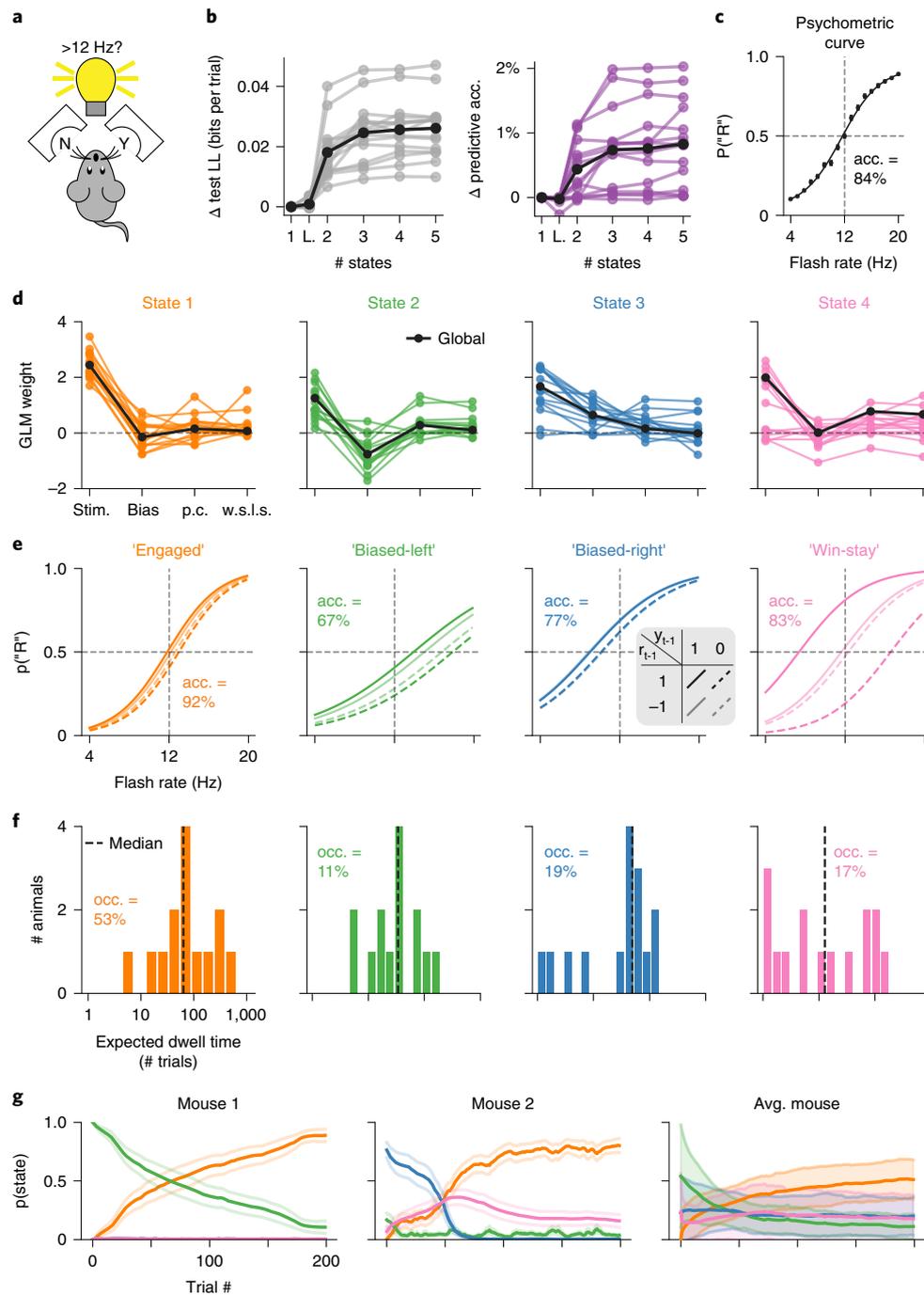


Fig. 5 | GLM-HMM application to second mouse dataset. **a**, Mice in the Odoemene et al.²⁰ task had to indicate if the flash rate of light pulses from an LED panel was above or below 12 Hz by nose poking to the right or left. **b**, Left: test set log-likelihood for 15 mice in this dataset. Black is the mean across animals. Right: change in test set predictive accuracy relative to a GLM for each mouse in the dataset. Black is the mean across all animals. **c**, The psychometric curve when all data from all mice are concatenated together. **d**, The retrieved weights for a four-state GLM-HMM for each mouse. The covariates influencing choice in each of the four states are the stimulus frequency, the mouse's bias, its choice on the previous trial ('p.c.') and the product of reward and choice on the previous trial ('w.s.l.s.'). Black indicates the global fit across all mice. **e**, The probability that the mouse went rightward in each of the four states as a function of the stimulus, previous choice and reward on the previous trial (each of the four lines corresponds to a different setting of previous choice and previous reward; the four lines are calculated using the global fit weights shown in **d**). We also report animals' task accuracy (labeled 'acc.') when in each of the four states. As discussed in the text, we label the fourth state the 'win-stay' state as the psychometric curve for this state is shifted so that the mouse is likely to repeat a choice only if the previous trial was rewarded. **f**, The expected dwell times for each mouse in each state, as obtained from the inferred transition matrices (for the full transition matrices for all mice, see Supplementary Fig. 11). The dashed black line indicates the median dwell time across animals. We report the fractional occupancies (labeled 'occ.') of the four states across all mice. For context, the median session length was 683 trials. **g**, Left and middle: average (across 20 sessions) posterior state probabilities for the first 200 trials within a session for two example mice; error bars are standard errors. Right: posterior state probabilities for first 200 trials when averaged across all mice. Error bars indicate standard deviation across mice. Stim., stimulus.

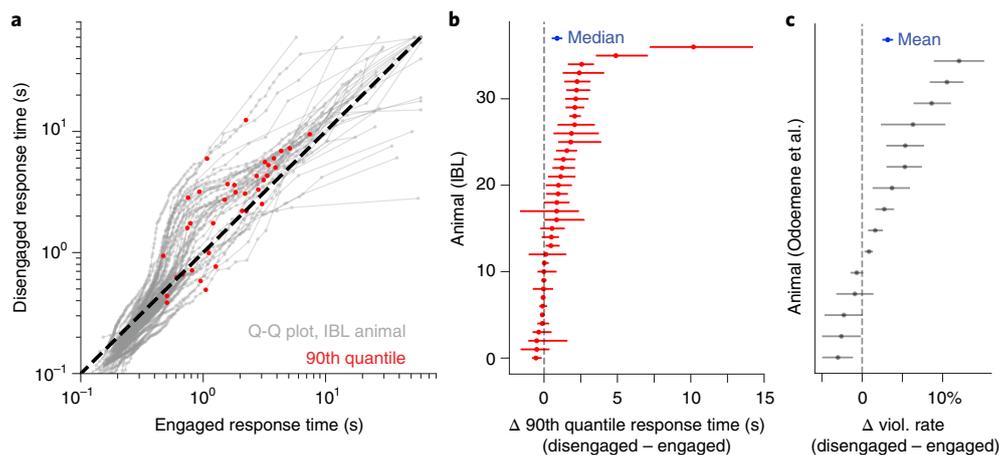


Fig. 6 | Behavioral correlates for GLM-HMM states. **a**, Q-Q plots for response time distributions associated with engaged and disengaged (biased leftward and rightward) states for IBL mice. Each curve is an individual animal, and the red dots indicate the 90th quantile response times. For all 37 mice, a Komogorov–Smirnov test indicated that the engaged and disengaged response time distributions were statistically different. Furthermore, as can readily be observed from the Q-Q plots, the longest response times typically occurred in the disengaged state. **b**, Difference in the 90th quantile response time for the engaged and disengaged states for each IBL animal as well as 95% bootstrap confidence intervals ($n = 2,000$). Blue is the median difference across mice (0.89 s) as well as the 95% bootstrap confidence interval. **c**, Violation rate differences for mice in the Odoemene et al.²⁰ dataset when in the engaged state (state 1) and disengaged states (states 2, 3 and 4). Error bars are 95% bootstrap confidence intervals ($n = 2,000$); blue indicates the mean difference in violation rate across all mice (3.2%).

37 mice (Komogorov–Smirnov tests reject the null hypothesis with $P < 0.05$ for all 37 mice; Fig. 6a). Examining the Q-Q plots, it was also clear that the most extreme response times for each animal were associated with the disengaged states (Fig. 6b). In the Odoemene et al.²⁰ dataset, we examined differences in the rate of violations—trials where the mouse failed to make a decision within the response period. The mean violation rate across all mice was 21% (much higher than the 1% violation rate observed in the IBL dataset), and violations were 3.2% more common in the disengaged states (2, 3, and 4) than in the engaged state (state 1).

State-dependent decision-making in human visual task. Finally, although the primary focus of this paper is identifying the discrete latent strategies underpinning mouse decision-making, the GLM-HMM framework is general and can be used to study choice behavior across species. We, thus, applied the GLM-HMM to the choice behavior of humans performing the visual motion discrimination task of ref.²⁷. Participants had to judge the difference in motion coherence between two consecutive random dot kinematograms: the first kinematogram always corresponded to a reference stimulus with 70% motion coherence, whereas the second stimulus could have more or less motion coherence (Fig. 7a). Participants indicated the stimulus that they perceived to have the greater motion coherence by pressing a button. Cross-validation revealed that the choice data of 24 of 27 humans was better explained by a two-state or three-state GLM-HMM compared to the classic lapse model (Fig. 7b,c). The mean improvement of the two-state GLM-HMM compared to the classic lapse model was 0.013 bits per trial, making a test dataset of 500 trials 90 times more likely to have been generated by a two-state GLM-HMM than by the classic lapse model.

Figure 7d shows the retrieved weights for each participant for the two-state model (we focus on the two-state model as this was the model favored by most participants), whereas Fig. 7e shows the probability of the participant selecting, in each of the two states (blue and green curves), that the second stimulus had more motion coherence than the first (choice = ‘more’) as a function of the relative motion coherence. The two states correspond to participants having a bias for selecting each of the two possible outcomes: when

using the blue state, participants were biased toward selecting ‘more’, whereas participants were biased toward selecting ‘less’ in the green state. Figure 7f shows the expected dwell times for all participants in each state, whereas Fig. 7g shows the empirical number of state changes detected in each session (across the entire participant population). The median number of state changes per session was six switches, indicating that it is not just mice that switch between strategies multiple times per session. Figure 7h illustrates the strategy changes across a session and shows the posterior state probabilities for three example sessions (for three different participants).

Discussion

In this work, we used the GLM-HMM framework to identify hidden states from perceptual decision-making data. Mouse behavior in two different perceptual decision-making tasks^{19,20} was far better described by a GLM-HMM with sustained engaged and biased states. Unlike the classic lapse model, these states alternated on the time scale of tens to hundreds of trials. We also applied the model to human psychophysical data²⁷ and found that human behavior was also better described by a GLM-HMM with sustained states that differed in bias. Although it might seem obvious that mice should exhibit fluctuations in task engagement over time, this is a very different account of behavior to that given by the classic lapse model, which assumes that error trials do not occur in blocks and are, instead, interspersed throughout a session. Furthermore, our method offers the practical benefit that it allows practitioners to segment trials according to the inferred behavioral strategy. Finally, it was not obvious that this alteration in choice policy over the course of a session should occur in a discrete rather than continuous manner. Comparing our results with those for PsyTrack—a model with continuous decision-making states—provided support for the view that mice use discrete rather than continuous decision-making states.

Although we found similarities in the strategies pursued by mice performing the different visual detection tasks, we also found some differences. For example, we found that mice performing the Odoemene et al.²⁰ task used an additional ‘win-stay’ strategy that was not observed in the IBL mice. One possible explanation for

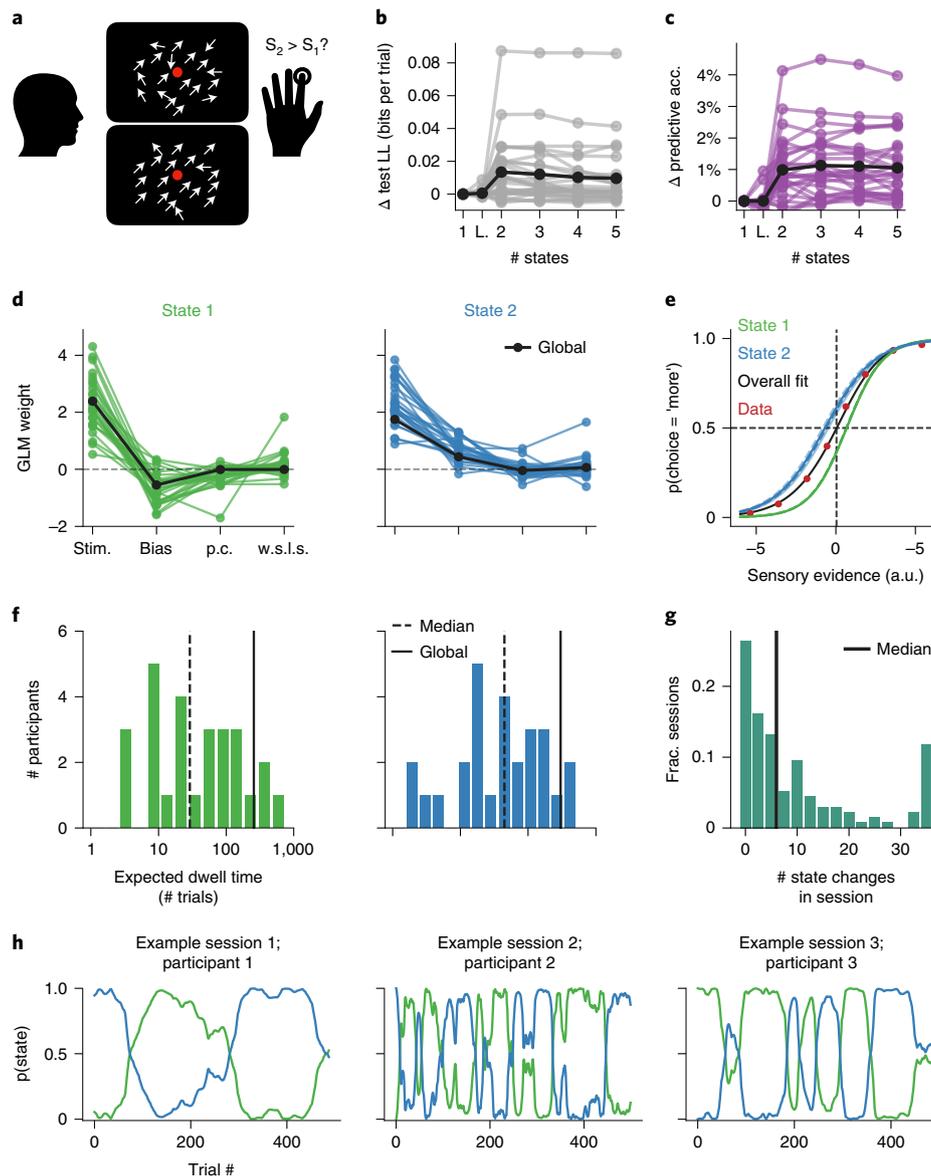


Fig. 7 | GLM-HMM application to human dataset. **a**, Twenty-seven human participants performed the motion coherence discrimination task of ref. ²⁷, where they used a button to indicate if there was more motion coherence in the first or second stimulus presented within a trial. **b**, Change in test set log-likelihood relative to a GLM for each of the 27 participants. Black is the mean across participants. **c**, Change in test set predictive accuracy relative to a GLM for each participant in the dataset. Black is the mean across participants. **d**, Retrieved weights for a two-state GLM-HMM for each participant. The covariates influencing choice in each of the four states are the difference in motion coherence among the two stimuli, the participant's bias, their choice on the previous trial ('p.c.') and the product of reward and choice on the previous trial ('w.s.l.s.'). Black indicates the global fit across all participants. **e**, Similarly to Figs. 2g and 5e, the blue and green curves give the probability that the participant went rightward in each of the two states as a function of the stimulus, previous choice and previous reward. That all blue curves are overlapping indicates the lack of dependence on previous choice and reward in this state, similarly for state 1. The black curve is the overall psychometric curve generated by the GLM-HMM. Specifically, using the global GLM-HMM parameters, we generated the same number of choices as those in the human dataset. At each trial, regardless of the true stimulus presented, we calculated $p_i("R")$ for each possible stimuli by averaging the per-state psychometric curves (blue and green) and weighting by the appropriate row in the transition matrix (depending on the sampled latent state at the previous trial). Finally, we averaged the per-trial psychometric curves across all trials to obtain the curve that is shown in black, whereas the red dots represent the empirical choice data. **f**, The expected dwell times for each participant in each state, as obtained from the inferred transition matrices. The black dashed line indicates the median dwell time across participants, whereas the black solid line indicates the global fit (Algorithm 1). **g**, Empirical number of state changes per session obtained using posterior state probabilities, such as those shown in **h**; median session length is 500 trials. Black indicates the median number of state changes across all sessions. **h**, Posterior state probabilities for three example sessions corresponding to three different participants. acc., accuracy; a.u., arbitrary units; Frac., fractional.

these strategy differences is that they arise due to the different shaping protocols associated with the different tasks. A follow-up study could investigate the role of shaping on GLM-HMM behavioral strategies.

The ability to infer the state or strategy employed by an animal on a trial-by-trial basis will be useful for characterizing differences in performance across sessions and across animals. It will also provide a powerful tool for neuroscientists studying the neural

mechanisms that support decision-making^{40–46}. It might be the case that different strategies rely on different circuits or different patterns of neural activity⁴⁷.

Although we found evidence of warm-up behavior at the start of sessions in the Odoemene et al.²⁰ mice, we were somewhat surprised to find few signatures of satiation or fatigue toward the end of a session. This might be due to the fact that sessions were typically of fixed duration and might have ended before mice had a chance to grow satiated or fatigued. One future direction will be to apply the GLM-HMM to experiments with longer sessions, where it might be useful for detecting changes in behavior reflecting satiety or fatigue.

Compared to the mice we studied, humans used two relatively engaged strategies with opposite biases when performing the perceptual decision-making task of ref.²⁷. Unfortunately, it is impossible to say if the difference in strategies observed across species were due to species differences or to differences in task protocol. An interesting follow-up experiment would be to apply the GLM-HMM to data from different species performing the same task.

In future work, we will aim to make explicit the connection between the ‘engaged’ and ‘disengaged’ strategies identified by our model and existing measures of arousal and engagement in the literature. Identifying the relationship between the GLM-HMM’s hidden states and pupil diameter, low-frequency LFP oscillations, spontaneous neuronal firing, noise correlations and the action of neuromodulators^{48–51} will be a priority.

That discrete states underpin mouse and human choice behavior might also call for new normative models to explain why subjects develop these states to begin with⁵². The existence of disengaged and engaged states could reflect explore–exploit behavior^{17,53,54} or optimal learning (for example, refs.^{55–58}), or it could simply indicate incomplete learning of the task. Alternatively, if the states uncovered by our model reflect a similar phenomenon to the internal and external modes of sensory processing observed in refs.^{59,60}, it might be that they arise due to the limited information processing capacity of the brain.

Another promising direction will be to replace the model’s fixed transition matrix with a GLM so as to allow external covariates to modulate the probability of state changes⁶. This would allow us to identify the factors that influence state changes (for example, a preponderance of unrewarded choices) and, potentially, seek to control such transitions.

Although there are many avenues for future research, we think that these results call for a critical rethinking of rodent and human perceptual decision-making behavior and the methods for analyzing it. Indeed, standard analysis methods do not take account of the possibility that an animal makes abrupt changes in decision-making strategy multiple times per session. We think that the ability to infer internal states from choice behavior will open up new directions for data analysis and provide new insights into a previously inaccessible dimension of perceptual decision-making behavior.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41593-021-01007-z>.

Received: 10 May 2021; Accepted: 17 December 2021;

Published online: 7 February 2022

References

- Gomez-Marín, A., Paton, J. J., Kampff, A. R., Costa, R. M. & Mainen, Z. F. Big behavioral data: psychology, ethology and the foundations of neuroscience. *Nat. Neurosci.* **17**, 1455–1462 (2014).

- Krakauer, J. W., Ghazanfar, A. A., Gomez-Marín, A., MacIver, M. A. & Poeppel, D. Neuroscience needs behavior: correcting a reductionist bias. *Neuron* **93**, 480–490 (2017).
- Berman, G. J. Measuring behavior across scales. *BMC Biol.* **16**, 23 (2018).
- Wiltchko, A. B. et al. Mapping sub-second structure in mouse behavior. *Neuron* **88**, 1121–1135 (2015).
- Sharma, A., Johnson, R., Engert, F. & Linderman, S. Point process latent variable models of larval zebrafish behavior. In: Bengio, S. et al. (eds) *Advances in Neural Information Processing Systems 31*, 10919–10930 (MIT Press, 2018).
- Calhoun, A. J., Pillow, J. W. & Murthy, M. Unsupervised identification of the internal states that shape natural behavior. *Nat. Neurosci.* **22**, 2040–2049 (2019).
- Batty, E. et al. BehaveNet: nonlinear embedding and Bayesian neural decoding of behavioral videos. In: Wallach, H. et al. (eds) *Advances in Neural Information Processing Systems 32*, 15706–15717 (MIT Press, 2019).
- Johnson, R. E. et al. Probabilistic models of larval zebrafish behavior reveal structure on many scales. *Curr. Biol.* **30**, 70–82 (2020).
- Green, D. M. & Swets, J. A. *Signal Detection Theory and Psychophysics* (Wiley, 1966).
- Klein, S. A. Measuring, estimating, and understanding the psychometric function: a commentary. *Percept. Psychophys.* **63**, 1421–1455 (2001).
- Bogacz, R., Brown, E., Moehlis, J., Holmes, P. & Cohen, J. D. The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. *Psychol. Rev.* **113**, 700–765 (2006).
- Gold, J. I. & Shadlen, M. N. The neural basis of decision making. *Annu. Rev. Neurosci.* **30**, 535–574 (2007).
- Ratcliff, R. & McKoon, G. The diffusion decision model: theory and data for two-choice decision tasks. *Neural Comput.* **20**, 873–922 (2008).
- Iigaya, K., Fonseca, M. S., Murakami, M., Mainen, Z. F. & Dayan, P. An effect of serotonergic stimulation on learning rates for rewards apparent after long intertrial intervals. *Nat. Commun.* **9**, 2477 (2018).
- Urai, A. E., de Gee, J. W., Tsetsos, K. & Donner, T. H. Choice history biases subsequent evidence accumulation. *eLife* **8**, e46331 (2019).
- Lak, A. et al. Reinforcement biases subsequent perceptual decisions when confidence is low, a widespread behavioral phenomenon. *eLife* **9**, e49834 (2020).
- Pisupati, S., Chartarifsky-Lynn, L., Khanal, A. & Churchland, A. K. Lapses in perceptual decisions reflect exploration. *eLife* **10**, e55490 (2021).
- Beron, C., Neufeld, S., Linderman, S. & Sabatini, B. Efficient and stochastic mouse action switching during probabilistic decision making. Preprint at <https://www.biorxiv.org/content/10.1101/2021.05.13.444094v1> (2021).
- The International Brain Laboratory et al. A standardized and reproducible method to measure decision-making in mice. *eLife* **10**, e63711 (2021).
- Odoemene, O., Pisupati, S., Nguyen, H. & Churchland, A. K. Visual evidence accumulation guides decision-making in unrestrained mice. *J. Neurosci.* **38**, 10143–10155 (2018).
- Pinto, L. et al. An accumulation-of-evidence task using visual pulses for mice navigating in virtual reality. *Front. Behav. Neurosci.* **12**, 36 (2018).
- Wichmann, F. A. & Hill, N. J. The psychometric function: I. Fitting, sampling, and goodness of fit. *Percept. Psychophys.* **63**, 1293–1313 (2001).
- Prins, N. The psychometric function: the lapse rate revisited. *J. Vis.* **12**, 25 (2012).
- Carandini, M. & Churchland, A. K. Probing perceptual decisions in rodents. *Nat. Neurosci.* **16**, 824–831 (2013).
- Lindsay, B. G. Mixture models: theory, geometry and applications. *NSF-CBMS Regional Conference Series in Probability and Statistics* **5**, i–163 <https://books.google.com/books?id=VFDzNhikFbQC&printsec=copyright#v=onepage&q&f=false> (1995).
- Law, C.-T. & Gold, J. I. Reinforcement learning can account for associative and perceptual learning on a visual-decision task. *Nat. Neurosci.* **12**, 655–663 (2009).
- Urai, A. E., Braun, A. & Donner, T. H. Pupil-linked arousal is driven by decision uncertainty and alters serial choice bias. *Nat. Commun.* **8**, 14637 (2017).
- Busse, L. et al. The detection of visual contrast in the behaving mouse. *J. Neurosci.* **31**, 11351–11361 (2011).
- Fründ, I., Wichmann, F. A. & Macke, J. H. Quantifying the effect of intertrial dependence on perceptual decisions. *J. Vis.* **14**, 9 (2014).
- Lueckmann, J.-M., Macke, J. H. & Nienborg, H. Can serial dependencies in choices and neural activity explain choice probabilities? *J. Neurosci.* **38**, 3495–3506 (2018).
- Escola, S., Fontanini, A., Katz, D. & Paninski, L. Hidden Markov models for the stimulus–response relationships of multistate neural systems. *Neural Comput.* **23**, 1071–1132 (2011).
- Bengio, Y. & Frasconi, P. An input output HMM architecture. In: Tesauro, G., Touretzky, D. S. & Leen, T. K. (eds) *Advances in Neural Information Processing Systems 7*, 427–434 (MIT Press, 1995).

33. Burgess, C. P. et al. High-yield methods for accurate two-alternative visual psychophysics in head-fixed mice. *Cell Rep.* **20**, 2513–2524 (2017).
34. Friedman, J., Hastie, T. & Tibshirani, R. *The Elements of Statistical Learning* (Springer, 2001).
35. Roy, N. A., Bak, J. H., Akrami, A., Brody, C. & Pillow, J. W. Efficient inference for time-varying behavior during learning. In: *Advances in Neural Information Processing Systems 31*, 5695–5705 (MIT Press, 2018).
36. Roy, N. A. et al. Extracting the dynamics of behavior in sensory decision-making experiments. *Neuron* **109**, 597–610 (2021).
37. Heitz, R. P. The speed–accuracy tradeoff: history, physiology, methodology, and behavior. *Front. Neurosci.* **8**, 150 (2014).
38. Ratcliff, R., Smith, P. L., Brown, S. D. & McKoon, G. Diffusion decision model: current issues and history. *Trends Cogn. Sci.* **20**, 260–281 (2016).
39. Shevinsky, C. A. & Reinagel, P. The interaction between elapsed time and decision accuracy differs between humans and rats. *Front. Neurosci.* **13**, 1211 (2019).
40. Ding, L. & Gold, J. I. Neural correlates of perceptual decision making before, during, and after decision commitment in monkey frontal eye field. *Cereb. Cortex* **22**, 1052–1067 (2012).
41. Ding, L. & Gold, J. I. The basal ganglia's contributions to perceptual decision making. *Neuron* **79**, 640–649 (2013).
42. Engel, T. A. et al. Selective modulation of cortical state during spatial attention. *Science* **354**, 1140–1144 (2016).
43. Zhou, Y. & Freedman, D. J. Posterior parietal cortex plays a causal role in perceptual and categorical decisions. *Science* **365**, 180–185 (2019).
44. Steinmetz, N. A., Zatzka-Haas, P., Carandini, M. & Harris, K. D. Distributed coding of choice, action and engagement across the mouse brain. *Nature* **576**, 266–273 (2019).
45. Zoltowski, D., Pillow, J. & Linderman, S. A general recurrent state space framework for modeling neural dynamics during decision-making. In: *Proceedings of the 37th International Conference on Machine Learning*, 11680–11691 (PMLR, 2020).
46. Hennig, J. A. et al. Learning is shaped by abrupt changes in neural engagement. *Nat. Neurosci.* **24**, 727–736 (2021).
47. Bolkan, S. S. et al. Strong and opponent contributions of dorsomedial striatal pathways to behavior depends on cognitive demands and task strategy. Preprint at <https://www.biorxiv.org/content/10.1101/2021.07.23.453573v1> (2021).
48. Carter, M. E. et al. Tuning arousal with optogenetic modulation of locus coeruleus neurons. *Nat. Neurosci.* **13**, 1526–1533 (2010).
49. Harris, K. D. & Thiele, A. Cortical state and attention. *Nat. Rev. Neurosci.* **12**, 509–523 (2011).
50. Reimer, J. et al. Pupil fluctuations track fast switching of cortical states during quiet wakefulness. *Neuron* **84**, 355–362 (2014).
51. Vinck, M., Batista-Brito, R., Knoblich, U. & Cardin, J. A. Arousal and locomotion make distinct contributions to cortical activity patterns and visual encoding. *Neuron* **86**, 740–754 (2015).
52. Ashwood, Z., Roy, N. A., Bak, J. H. & Pillow, J. W. Inferring learning rules from animal decision-making. In: *Advances in Neural Information Processing Systems 33*, 3442–3453 (MIT Press, 2020).
53. Daw, N. D., O'Doherty, J. P., Dayan, P., Seymour, B. & Dolan, R. J. Cortical substrates for exploratory decisions in humans. *Nature* **441**, 876–879 (2006).
54. Dayan, P. & Daw, N. D. Decision theory, reinforcement learning, and the brain. *Cogn. Affect. Behav. Neurosci.* **8**, 429–453 (2008).
55. Behrens, T. E. J., Woolrich, M. W., Walton, M. E. & Rushworth, M. F. S. Learning the value of information in an uncertain world. *Nat. Neurosci.* **10**, 1214–1221 (2007).
56. Fiser, J., Berkes, P., Orbán, G. & Lengyel, M. Statistically optimal perception and learning: from behavior to neural representations. *Trends Cogn. Sci.* **14**, 119–130 (2010).
57. Mendonça, A. G. et al. The impact of learning on perceptual decisions and its implication for speed-accuracy tradeoffs. *Nat. Commun.* **11**, 2757 (2020).
58. Gershman, S. J. Origin of perseveration in the trade-off between reward and complexity. *Cognition* **204**, 104394 (2020).
59. Weinhhammer, V., Chikermane, M. & Sterzer, P. Bistable perception alternates between internal and external modes of sensory processing. *iScience* **24**, 102234 (2021).
60. Weinhhammer, V. A., Stuke, H., Eckert, A.-L., Standvoss, K. & Sterzer, P. Humans and mice fluctuate between external and internal modes of sensory processing. Preprint at <https://www.biorxiv.org/content/10.1101/2021.08.20.457079v2> (2021).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2022

The International Brain Laboratory

Zoe C. Ashwood^{1,2}, Nicholas A. Roy², Anne E. Urai³, Anne K. Churchland⁴, Alexandre Pouget⁵ and Jonathan W. Pillow^{2,6}

Methods

We confirm that our research complies with all relevant ethical regulations.

Inference of GLM-HMM parameters. *GLM-HMM objective function.* When fitting the GLM-HMM, the goal is to learn the transition matrix $A \in \mathbb{R}^{K \times K}$, the set of weights influencing choice in each state, $\{\mathbf{w}_k\}_{k=1}^K$ with $\mathbf{w}_k \in \mathbb{R}^M$, and the initial state distribution, $\boldsymbol{\pi} \in \mathbb{R}^K$. We fit this set of parameters, collectively labeled as Θ , to choice data using maximum a posteriori estimation. This was practically implemented by the EM algorithm⁶¹.

The EM algorithm has been previously adapted to fit HMMs with external inputs^{61,32}. However, there are some application-specific choices in the implementation, so, for completeness, we describe the full procedure here.

The EM algorithm seeks to maximize the log-posterior of the parameters given the choice and input data, $\mathcal{D} \equiv \{\mathbf{y}, \{\mathbf{x}_t\}_{t=1}^T\}$. The log-posterior is given, up to an unknown constant, by:

$$\log p(\Theta|\mathcal{D}) = \sum_{s=1}^S \log \left(\sum_{\mathbf{z}_s} p(\mathbf{y}_s, \mathbf{z}_s | \{\mathbf{x}_{s,t}\}_{t=1}^{T_s}, \Theta) \right) + \log p(\Theta), \quad (5)$$

where s indexes the session in which the data were collected (out of S total sessions), and the sum over \mathbf{z}_s is over all K^{T_s} possible latent state paths in the T_s trials of session s .

The prior distribution over the model parameters $p(\Theta)$ that we used was:

$$p(\Theta) \equiv p(\{\mathbf{w}_k\})p(A)p(\boldsymbol{\pi}) = \left(\prod_{j=1}^K \mathcal{N}(\mathbf{w}_j | 0, \sigma^2 I) \right) \left(\prod_{j=1}^K \text{Dirichlet}(A_{j|\alpha}) \right) \text{Dirichlet}(\boldsymbol{\pi} | \boldsymbol{\alpha}_\pi) \quad (6)$$

The prior over the GLM weight vectors \mathbf{w}_k was, thus, an independent, zero-mean Gaussian with variance σ^2 . Smaller values of σ^2 have the effect of shrinking the fitted weights toward 0, whereas $\sigma^2 = \infty$ corresponds to a flat prior. For the transition matrix, we placed an independent Dirichlet prior over each row $A_{j\cdot}$, which is a natural choice for vectors on the unit $(K-1)$ simplex (that is, all elements of the vector must be non-negative and sum to 1). The Dirichlet distribution is controlled by a shape parameter α : $p(A_{j|\alpha}) = \frac{1}{B(\boldsymbol{\alpha})} \prod_{i=1}^K (A_{ji})^{\alpha-1}$. A value of $\alpha=1$ corresponds to a flat prior, whereas large values of α have the effect of spreading probability mass equally over states. We also placed a Dirichlet prior over the initial state distribution $\boldsymbol{\pi}$.

To select the hyperparameters σ and α governing the prior, we performed a grid search for $\sigma \in \{0.5, 0.75, 1, 2, 3\}$ and $\alpha \in \{1, 2\}$ and selected the hyperparameters that resulted in the best performance on a held-out validation set. For IBL mice, the prior hyperparameters selected were $\sigma=2$ and $\alpha=2$. For mice in the Odoemene et al.²⁰ dataset, the best hyperparameters were $\sigma=0.75$ and $\alpha=2$. Finally, for the initial state distribution, we set $\alpha_\pi=1$.

EM algorithm. We used the EM algorithm^{61,62} to maximize the log-posterior given in Eq. (5) with respect to the GLM-HMM parameters. As the sum given in Eq. (5) involves an exponential number of terms— $O(K^{T_s})$ terms to be specific—we do not maximize this expression directly. Instead, the EM algorithm provides an efficient way to compute this term using a single forward and backward pass over the data. During the E-step of the EM algorithm, we compute the ‘expected complete data log-likelihood’ (ECLL), which is a lower bound on the right-hand side of Eq. (5)^{61–63}. Then, during the ‘maximization’ or M-step of the algorithm, we maximize the ECLL with respect to the model parameters, Θ . It can be shown that this procedure has the effect of always improving the log-posterior in each step of the algorithm⁶³.

Concretely, the ECLL can be written as a sum over the ECLLs for each session and the log-prior. The ECLL for a single session s can be written

$$\begin{aligned} \text{ECLL}_s(\Theta) &\equiv \sum_{\mathbf{z}_s} p(\mathbf{z}_s | \mathcal{D}_s, \Theta^{\text{old}}) \log p(\mathbf{y}_s, \mathbf{z}_s | \Theta, \{\mathbf{x}_{s,t}\}_{t=1}^{T_s}) \\ &= \sum_{k=1}^K \gamma_{s,1,k} \log \pi_k + \sum_{t=1}^{T_s} \sum_{j=1}^K \sum_{k=1}^K \xi_{s,t,j,k} \log A_{jk} \\ &\quad + \sum_{t=1}^{T_s} \sum_{k=1}^K \gamma_{s,t,k} \log p(y_{s,t} | z_{s,t} = k, \mathbf{x}_{s,t}, \mathbf{w}_k) \end{aligned} \quad (7)$$

To get to the second line, we substituted the definition of the joint distribution for the GLM-HMM:

$$p(\mathbf{y}_s, \mathbf{z}_s | \{\mathbf{x}_{s,t}\}_{t=1}^{T_s}, \Theta) \equiv p(z_1) p(y_1 | z_1, \mathbf{x}_1) \prod_{t=2}^{T_s} p(z_t | z_{t-1}) p(y_t | z_t, \mathbf{x}_t). \quad (8)$$

In Eq. (7), $p(y_{s,t} | z_{s,t} = k, \mathbf{x}_{s,t}, \mathbf{w}_k)$ is the Bernoulli GLM distribution given by Eq. (4). Finally,

$$\gamma_{s,t,k} \equiv p(z_{s,t} = k | \mathcal{D}_s, \Theta^{\text{old}}) \quad (9)$$

is the posterior state probability at trial t (in session s) for state k , whereas

$$\xi_{s,t,j,k} \equiv p(z_{s,t+1} = k, z_{s,t} = j | \mathcal{D}_s, \Theta^{\text{old}}) \quad (10)$$

is the joint posterior state distribution for two consecutive latents.

While the formula for the log-posterior (Eq. (5)) involved the sum over all possible state assignments at each trial, by taking advantage of the structure of the joint probability distribution for the GLM-HMM (Eq. (8)), this sum was implemented efficiently, and Eq. (7) involves summing over at most $O(K^2 T_s)$ elements.

The single and joint posterior state probabilities, $\gamma_{s,t,k}$ and $\xi_{s,t,j,k}$, are estimated via the forward–backward algorithm, which we describe below.

Expectation step: forward–backward algorithm. During the E-step of the EM algorithm, the single and joint posterior state probabilities for all trials and states, $\{\gamma_{s,t,k}\}$ and $\{\xi_{s,t,j,k}\}$, are estimated using the forward–backward algorithm⁶⁴ at the current setting of the GLM-HMM parameters, Θ^{old} .

The forward–backward algorithm makes use of recursion and memoization to allow these posterior probabilities to be calculated efficiently, with the forward and backward passes of the algorithm each requiring just a single pass through all trials within a session.

The goal of the forward pass is to obtain, for each trial t within session s and each state k , the quantity

$$a_{s,t,k} \equiv p(\mathbf{y}_{s,[1:t]}, z_{s,t} | \{\mathbf{x}_{s,t'}\}_{t'=1}^t), \quad (11)$$

which represents the posterior probability of the choice data up until trial t and the latent state at trial t being state k .

The posterior probability associated with trial 1, $a_{s,1,k}$, can be calculated as follows:

$$a_{s,1,k} = \boldsymbol{\pi}_k p(y_{s,1} | z_{s,1} = k, \mathbf{x}_{s,1}, \mathbf{w}_k), \quad (12)$$

where $p(y_{s,1} | z_{s,1} = k, \mathbf{x}_{s,1}, \mathbf{w}_k)$ is the usual Bernoulli GLM distribution.

For trials $1 < t \leq T_s$, we can obtain these probabilities as follows:

$$a_{s,t,k} = \sum_{j=1}^K a_{s,t-1,j} A_{jk} p(y_{s,t} | z_{s,t} = k, \mathbf{x}_{s,t}, \mathbf{w}_k). \quad (13)$$

During the backward pass, the goal is to calculate the posterior probability of the choice data beyond the current trial, $b_{s,t,k}$, for each trial t within session s and for all states k :

$$b_{s,t,k} \equiv p(\mathbf{y}_{s,[t+1:T_s]} | z_{s,t} = k, \{\mathbf{x}_{s,t'}\}_{t'=t+1}^{T_s}). \quad (14)$$

Similarly to the forward pass, these quantities can be calculated efficiently by recognizing that

$$b_{s,T_s,k} = 1 \quad (15)$$

and, for $t \in \{T_s - 1, \dots, 1\}$:

$$b_{s,t,j} = \sum_{k=1}^K b_{s,t+1,k} A_{jk} p(y_{s,t+1} | z_{s,t+1} = k, \mathbf{x}_{s,t+1}, \mathbf{w}_k). \quad (16)$$

From the $a_{s,t,k}$ and $b_{s,t,k}$ quantities obtained via the forward–backward algorithm, we can form the single and joint posterior state probabilities, $\gamma_{s,t,k}$ and $\xi_{s,t,j,k}$, that we actually care about for forming the ECLL of Eq. (7) as follows:

$$\begin{aligned} \gamma_{s,t,k} &\equiv p(z_{s,t} = k | \mathbf{y}_s, \{\mathbf{x}_{s,t'}\}_{t'=1}^{T_s}, \Theta^{\text{old}}) \\ &= \frac{p(\mathbf{y}_{s,[0:t]} | z_{s,t} = k | \{\mathbf{x}_{s,t'}\}_{t'=1}^t, \Theta^{\text{old}}) p(\mathbf{y}_{s,[t+1:T_s]} | z_{s,t} = k, \{\mathbf{x}_{s,t'}\}_{t'=t+1}^{T_s}, \Theta^{\text{old}})}{p(\mathbf{y}_s | \{\mathbf{x}_{s,t'}\}_{t'=1}^{T_s}, \Theta^{\text{old}})} \\ &= \frac{a_{s,t,k} b_{s,t,k}}{\sum_{k=1}^K a_{s,t,k} b_{s,t,k}}. \end{aligned} \quad (17)$$

Similarly,

$$\xi_{s,t,j,k} = \frac{a_{s,t,j} A_{jk} b_{s,t+1,k} p(y_{s,t+1} | z_{t+1} = k, \mathbf{x}_{t+1}, \mathbf{w}_k)}{\sum_{k=1}^K a_{s,t,k} b_{s,t,k}} \quad (18)$$

where, once again, $p(y_{s,t+1} | z_{t+1} = k, \mathbf{x}_{t+1}, \mathbf{w}_k)$ is the Bernoulli GLM distribution.

Maximization step. After running the forward-backward algorithm, we can compute the total ECLL by summing over the per-session ECLLs (Eq. (7)) and adding the log-prior. During the M-step, we maximize the ECLL with respect to the GLM-HMM parameters, Θ . For the initial state distribution π and the transition matrix A , this results in the closed-form updates:

$$\pi_k^{\text{new}} = \frac{\sum_{s=1}^S \gamma_{s,1,k}}{\sum_{s=1}^S \sum_{j=1}^K \gamma_{s,1,j}} \quad (19)$$

$$A_{jk}^{\text{new}} = \frac{\alpha - 1 + \sum_{s=1}^S \sum_{t=2}^{T_s} \xi_{s,t,j,k}}{K(\alpha - 1) + \sum_{s=1}^S \sum_{t=2}^{T_s} \sum_{k=1}^K \xi_{s,t,j,k}}. \quad (20)$$

For the GLM weights, there is no such closed form update, but a Bernoulli GLM falls into the class of functions mapping external inputs to HMM emission probabilities considered in ref.³¹, so we know that the ECLL is concave in the GLM weights. As such, we can numerically find the GLM weights that maximize (not just locally but globally) the ECLL using the BFGS algorithm^{65–68} as implemented by the `scipy.optimize` function in Python⁶⁹.

Comparing states across animals and GLM-HMM parameter initialization. In Figs. 4 and 5, we show the results from fitting a single GLM-HMM to each animal; however, it is non-trivial to map the retrieved states across animals to one another. As such, we employed a multi-stage fitting procedure that allowed us to make this comparison, and we detail this procedure in Algorithm 1. In the first stage, we concatenated the data for all animals in a single dataset together (for the IBL dataset, this would be the data for all 37 animals). We then fit a GLM (a one-state GLM-HMM) to the concatenated data using maximum likelihood estimation. We used the fit GLM weights to initialize the GLM weights of a K -state GLM-HMM that we again fit to the concatenated dataset from all animals together (to obtain a ‘global fit’). We added Gaussian noise with $\sigma_{\text{init}} = 0.2$ to the GLM weights, so that the initialized states were distinct, and we initialized the transition matrix of the K -state GLM-HMM as $0.95 \times \mathbb{1} + \mathcal{N}(0, \Sigma_{\text{trans}})$ where $\Sigma_{\text{trans}} \in \mathbb{R}^{K^2 \times K^2}$ and $\Sigma_{\text{trans}} = 0.05 \times \mathbb{1}$. We then normalized this so that rows of the transition matrix added up to 1 and represented probabilities. Although the EM algorithm is guaranteed to converge to a local optimum in the log-probability landscape of Eq. (5), there is no guarantee that it will converge to the global optimum⁷⁰. Correspondingly, for each value of K , we fit the model 20 times using 20 different initializations.

In the next stage of the fitting procedure, we wanted to obtain a separate GLM-HMM fit for each animal, so we initialized a model for each animal with the GLM-HMM global fit parameters from all animals together (out of the 20 initializations, we chose the model that resulted in the best training set log-likelihood). We then ran the EM algorithm to convergence; it is these recovered parameters that are shown in Figs. 4 and 5. By initializing each individual animal’s model with the parameters from the fit to all animals together, it was no longer necessary for us to permute the retrieved states from each animal so as to map semantically similar states to one another.

Algorithm 1. Multi-stage GLM-HMM fitting procedure

```

1: Fit GLM (one-state GLM-HMM) to all data from all animals
2: Fit global GLM-HMM to all data from all animals:
3: for  $K \in \{2, \dots, 5\}$  do
4:   for  $\text{init} \in \{1, \dots, 20\}$  do
5:     Initialize  $K$ -state GLM-HMM using noisy GLM weights
6:     Run EM algorithm until convergence
7:   end for
8: end for
9: Fit separate GLM-HMM to each animal by initializing with global fit:
10: for each individual animal do
11:   for  $K \in \{2, \dots, 5\}$  do
12:     Initialize  $K$ -state GLM-HMM using best global
      GLM-HMM parameters for this  $K$ 
13:     Run EM algorithm until convergence
14:   end for
15: end for

```

We note that the initialization scheme detailed above is sufficiently robust so as to allow recovery of GLM-HMM parameters in various parameter regimes of interest. In particular, we simulated datasets from a GLM-HMM with the global fit parameters for both the IBL and Odoemene et al.²⁰ datasets as well as a global fit lapse model. We show the results of these recovery analyses in Extended Data Figs. 9 and 10.

Assessing model performance. Cross-validation. There are two ways in which to perform cross-validation when working with HMMs. First, it is possible to hold out entire sessions of choice data for assessing test set performance. That is, when fitting the model, the objective function in Eq. (5) and the ECLL in Eq. (7) are modified to include only 80% of sessions (because we use five-fold cross-validation throughout this work); and the log-likelihood of the held-out 20% of sessions is calculated using the fit parameters and a single run of the forward pass on the held-out sessions:

$$\text{LL}_{\text{test}} \equiv \sum_{s \in S \setminus S'} \log p \left(\mathbf{y}_s | \{ \mathbf{x}_{s,t} \}_{t=1}^T, \Theta' \right) = \log \sum_{s \in S'} \sum_{k=1}^K a_{s,T,k} \quad (21)$$

where $S \setminus S'$ is the set of held-out sessions, and Θ' is the set of GLM-HMM parameters obtained by fitting the model using the trials from S' .

The second method of performing cross-validation involves holding out 20% of trials within a session. When fitting the model, the third term in the ECLL is modified so as to exclude these trials and is now

$$\sum_{t \in T'} \sum_{k=1}^K \gamma_{s,t,k} \log p \left(y_{s,t,k} | z_{s,t} = k, \mathbf{x}_{s,t}, \mathbf{w}_k \right),$$

where T' is the set of trials to be used to fit the model. Furthermore, the calculation of the posterior state probabilities, $\gamma_{s,t,k}$ and $\xi_{s,t,j,k}$, is also modified so as to exclude the test set choice data. In particular, $\gamma_{s,t,k}$ is now $p \left(z_{s,t} | \{ y_{s,t'} \}_{t' \in T'}, \{ \mathbf{x}_{s,t'} \}_{t' \in T'}, \Theta^{\text{old}} \right)$ and, similarly, $\xi_{s,t,j,k}$ is now

$$p \left(z_{s,t+1} = k | z_{s,t} = j, \{ y_{s,t'} \}_{t' \in T'}, \{ \mathbf{x}_{s,t'} \}_{t' \in T'}, \Theta^{\text{old}} \right).$$

The method of calculating these modified posterior probabilities is as detailed in Eq. (17) and Eq. (18), but now the calculation of the forward and backward probabilities $a_{s,t,k}$ and $b_{s,t,k}$ in Eq. (12), Eq. (13), Eq. (15) and Eq. (16) is modified so that, on trials that are identified as test trials, the $p \left(y_{s,t} | z_{s,t} = k, \mathbf{x}_{s,t}, \mathbf{w}_k \right)$ term in these equations is replaced with 1. In Figs. 2, 4 and 5, we perform cross-validation by holding out entire sessions. We thought it would be more difficult to make good predictions on entire held-out sessions, compared to single trials within a session, as we thought that mice would exhibit more variability in behavior across sessions compared to within sessions. When we compare the performance of the GLM-HMM against the PsyTrack model of ref.³⁵ in Fig. 4f, we use the second method of cross-validation so as to use the same train and test sets as PsyTrack (PsyTrack cannot make predictions on entire held-out sessions).

Units for test set log-likelihood. In Figs. 2, 4, 5 and 7, as well as in some Supplementary Figures, we report the log-likelihood of different models on held-out sessions in units of bits per trial. This is calculated as follows:

$$\text{LL}_{\text{bits per trial}} = \frac{\text{LL}_{\text{test}} - \text{LL}_0}{n_{\text{test}} \log(2)} \quad (22)$$

where, for the GLM-HMM, LL_{test} is the log-likelihood of the entire test set as calculated in Eq. (21), and LL_0 is the log-likelihood of the same test set under a Bernoulli model of animal choice behavior. Specifically, this baseline model assumes that animals flip a coin on each trial so as to decide to go right, and the probability of going right is equal to the fraction of trials in the training set in which the animal chose to go to the right. n_{test} is the number of trials in the test set, and it is important to include because LL_{test} depends on the number of trials in the test set. Dividing by $\log(2)$ gives the log-likelihood the units of bits per trial. Clearly, larger values of the log-likelihood are better, with a value of 0 indicating that a model offers no improvement in prediction compared to the crude baseline model described above. However, even small values of test set log-likelihood, when reported in units of bits per trial, can indicate a large improvement in predictive power. For a test set size of $n_{\text{test}} = 500$, a log-likelihood value of 0.01 bits per trial indicates that the test data are 31.5 times more likely to have been generated with the GLM-HMM compared to the baseline model. For a test set of $n_{\text{test}} = 5,000$, and the same value of 0.01 bits per trial for the log-likelihood, the test set becomes 1×10^{15} times more likely under the GLM-HMM compared to the baseline model.

Predictive accuracy. In Figs. 2, 4 and 5, we also report the predictive accuracy of the GLM-HMM. When calculating the predictive accuracy, we employ a method similar to the second method of cross-validation described above. In particular, we hold out 20% of trials in a session and then obtain the posterior state probabilities for these trials, $t' \in \{T \setminus T'\}$, as $\gamma_{s,t',k} = p \left(z_{s,t'} | \{ y_{s,t'} \}_{t' \in T'}, \{ \mathbf{x}_{s,t'} \}_{t' \in T'}, \Theta \right)$, using the other 80% of trials (this latter set of trials being labeled T'). We then calculate the probability of the held-out choices being to go right as:

$$\begin{aligned} p_{R,s,t'} &\equiv p \left(y_{s,t'} = 1 | \{ y_{s,t'} \}_{t' \in T'}, \{ \mathbf{x}_{s,t'} \}_{t' \in T'}, \mathbf{x}_{s,t'}, \Theta \right) \\ &= \sum_{k=1}^K \left(\gamma_{s,t',k} \right) p \left(y_{s,t'} = 1 | z_{s,t'} = k, \mathbf{x}_{s,t'}, \mathbf{w}_k \right) \end{aligned} \quad (23)$$

We then calculate the predictive accuracy as:

$$\text{predictive accuracy} = \frac{\sum_{t'} \mathbb{1}_{\{T \setminus T'\}} \left(\left(\mathbb{1}_{\{y_{t'}=1\}} \mathbb{1}_{\{p_{R,t'} > 0.5\}} \right) + \left(\mathbb{1}_{\{y_{t'}=0\}} \mathbb{1}_{\{p_{R,t'} \leq 0.5\}} \right) \right)}{\mathbb{1}_{\{T \setminus T'\}}} \quad (24)$$

Comparison with PsyTrack Model of Roy et al. The PsyTrack model of Roy et al.³⁶ assumes that an animal makes its choice at trial t according to

$$p(y_t = 1 | \mathbf{x}_t) = \frac{1}{1 + e^{-\mathbf{w}_t \cdot \mathbf{x}_t}} \quad (25)$$

where \mathbf{w}_t evolves according to

$$\mathbf{w}_t = \mathbf{w}_{t-1} + \eta_t \quad (26)$$

where $\eta_t \sim \mathcal{N}(0, \text{diag}(\sigma_1^2, \dots, \sigma_M^2))$ and $\mathbf{w}_0 \sim \mathcal{N}(0, \text{diag}(\sigma_{\text{init},1}^2, \dots, \sigma_{\text{init},M}^2))$. Specifically, the animal is assumed to use a set of slowly changing weights to make its decision on each trial.

To perform model comparison with the PsyTrack model of refs.^{35,36}, we used the code provided at <https://github.com/nicholas-roy/psytrack> (version 1.3).

Simulating choice data with AR(1) model for weights. Although the GLM-HMM is a generative model that can be readily used to simulate choice data that resembles, in accuracy and in the resulting psychometric curves, the choice data of real animals, this is not true for the PsyTrack model of refs.^{35,36}. Indeed, specifying only the $\{\sigma_{\text{init},m}^2\}$ and $\{\sigma_m^2\}$ hyperparameters of that model and then generating weights according to Eq. (26) and choice data according to Eq. (25) will likely result in choice behavior that is vastly different from that of real animals (the PsyTrack model is under-constrained as a generative model). As such, so as to produce the right-most panel of Fig. 4f, we simulated smoothly evolving weights from an AR(1) model where the parameters of this model were obtained using the PsyTrack fits to real data. Specifically, we assumed that the probability of going rightward at time t was given by:

$$p(y_t = 1 | \mathbf{w}_t, \mathbf{x}_t) = \frac{1}{1 + e^{-\mathbf{w}_t \cdot \mathbf{x}_t}} \quad (27)$$

and we assumed that the weights evolved according to an AR(1) process as follows:

$$w_{m,t} = \alpha_m w_{m,t-1} + (1 - \alpha_m) \bar{w}_m^{\text{PsyTrack}} + \eta_{m,t} \quad (28)$$

where $w_{m,t}$ is the m th element of \mathbf{w}_t in Eq. (27), and $\bar{w}_m^{\text{PsyTrack}}$ is the average weight that an animal places on covariate m across all trials when fit with the PsyTrack model: $\bar{w}_m^{\text{PsyTrack}} = \frac{1}{T} \sum_t w_{t,m}^{\text{PsyTrack}}$. We obtained α_m by regressing $w_{t,m}^{\text{PsyTrack}}$ against $w_{t-1,m}^{\text{PsyTrack}}$ and taking the retrieved slope (after confirming that the retrieved slope had a magnitude of less than 1, so as to ensure that the weights did not diverge as $t \rightarrow \infty$). $\eta_{m,t} \sim \mathcal{N}(0, \tilde{\sigma}_m^2)$, where $\tilde{\sigma}_m^2 = (1 - \alpha_m^2) \sigma_m^2$ and σ_m^2 was obtained from the PsyTrack fit to an animal's real choice data. Finally, we set $\mathbf{w}_0 = 0$.

We simulated weight trajectories for thousands of trials for each animal, so that the AR(1) process reached the stationary regime for each covariate, and so that the mean, variance and autocovariance of each weight for each covariate were close to those returned by the PsyTrack fits to the real choice data.

Datasets studied. In this paper, we applied the GLM-HMM to three publicly available behavioral datasets associated with recent publications. First, we studied the data associated with ref.¹⁹ that is made available via Figshare at <https://doi.org/10.6084/m9.figshare.11636748>. We used the framework developed in ref.⁷¹ to access the data. We modeled the choice data for the 37 animals in this dataset, which had more than 30 sessions of data during the 'bias block' regime. We focused on this regime because of the fact that mice, when they have reached this regime, understand the rules of the task and exhibit stationary behavior (see Supplementary Fig. 1 for plots of accuracy against session identity for each animal, as well as Supplementary Fig. 2 for the psychometric curves for these animals for the trials studied). For each session, we subset to the first 90 trials of data because, during these trials, the stimulus was equally likely to appear on the left or right of the screen. After the first 90 trials, the structure of the task changed, and, for a block of trials, the stimulus appeared on the left with a probability of either 80% or 20%; the block identity switched multiple times throughout a session, so that 80% and 20% blocks were interleaved. We subset to the animals with more than 30 sessions of data because we were able to confidently recover GLM-HMM and lapse model generative parameters when we simulated datasets with this number of trials (Extended Data Figs. 9 and 10). As a sanity check, to make sure that the recovered states and transitions were not a consequence of the animals that we study having been exposed to bias blocks in earlier sessions, we obtained data for four animals that were never exposed to bias blocks (not included in the publicly released dataset) and fit the GLM-HMM to the choice data for these

animals. In Extended Data Fig. 3, we show that the retrieved states, dwell times and model comparison results for these animals look very similar to those shown in Fig. 4.

The second dataset that we studied was that associated with ref.²⁰, with the data being made available at <https://doi.org/10.14224/1.38944>. Once again, we studied sessions after animals had learned the task (Supplementary Figs. 13 and 14). For this dataset, the retrieved states were less distinct than those for the IBL dataset, and, as such, we required more trials to be able to recover the generative parameters in simulated data (Extended Data Fig. 9). We, thus, subset to the 15 animals with more than 20 sessions of data and 12,000 trials of data. Compared to the IBL dataset, where the violation rate across all animals' data was less than 1% of trials (where a violation is where the animal chose not to respond), the violation rate across the 15 animals that we studied from this second dataset was 21%. Thus, it was important to develop a principled method for dealing with violation trials. We treated violation trials as trials with missing choice data, and we handled these trials in a similar way to how we handled test data when performing the second type of cross-validation described above.

That is, we modified the third term of the ECLL given in Eq. (7) to exclude violation trials, and we modified the definition of the posterior state probabilities for these trials to be $\gamma_{s,t,k} = p(z_{s,t} | \{y_{s,t'}\}_{t' \in T'}, \{\mathbf{x}_{s,t'}\}_{t' \in T'}, \Theta^{\text{old}})$ and

$\xi_{s,t,j,k} = p(z_{s,t+1} = k | z_{s,t} = j, \{y_{s,t'}\}_{t' \in T'}, \{\mathbf{x}_{s,t'}\}_{t' \in T'}, \Theta^{\text{old}})$, where T' is the set of non-violation trials. The calculation of the forward and backward probabilities, $a_{s,t,k}$ and $b_{s,t,k}$, was modified so that, on violation trials, in Eq. (12), Eq. (13), Eq. (15) and Eq. (16), the $p(y_{s,t} | z_{s,t} = k, \mathbf{x}_{s,t}, \mathbf{w}_k)$ term was replaced with 1.

Finally, the third dataset that we studied was the human dataset associated with ref.²⁷, with the data being made available at <https://doi.org/10.6084/m9.figshare.4300043>. We applied the GLM-HMM to the data from all 27 participants in this dataset. This dataset has five sessions, each of approximately 500 trials, for each participant and does not include the initial session that was used to teach participants how to perform this task.

Forming the design matrix. Each of the models discussed in this paper (GLM-HMM, the classic lapse model and the PsyTrack model of Roy et al.³⁶) were fit using a design matrix of covariates, $X \in \mathbb{R}^{T \times M}$, where T was the number of trials of choice data for a particular animal. A single row in this matrix was the vector of covariates, $\mathbf{x}_t \in \mathbb{R}^M$, influencing the animal's choice at trial t . For all analyses presented in the text, unless specified otherwise, $M = 4$.

For all tasks, the first column in the design matrix was the z -scored stimulus intensity. For the IBL task, we calculated the stimulus intensity as the difference in the value of the visual contrast on the right side of the screen minus the visual contrast on the left side of the screen. This resulted in nine different values for the 'signed contrast': $\{-100, -25, -12.5, -6.25, 0, 6.25, 12.5, 25, 100\}$. We then z -scored this difference quantity across all trials. For the Odoemene et al.²⁰ task, we subtracted the 12-Hz threshold from the flash rate presented on each trial and then z -scored the resulting quantity. For the human dataset²⁷, we z -scored the relative motion strength presented on each trial.

For all trials, all animals and all tasks, the second column of the design matrix was set to 1, so as to enable us to capture the animal's innate bias for going rightward or leftward. The third column in the design matrix was the animal's choice on the previous trial $X_{t,3} \equiv 2y_{t-1} - 1$. Whereas $y_{t-1} \in \{0, 1\}$, $X_{t,3} \in \{-1, 1\}$. It is not strictly necessary to perform this scaling, but we did so to ensure that the range of values for $X_{t,1}$ and $X_{t,3}$ were more similar (which can be useful when performing parameter optimization). Finally, the fourth column in the design matrix was the win-stay, lose-switch covariate, which was calculated as $X_{t,4} \equiv r_{t-1} \times (2y_{t-1} - 1)$, where $r_{t-1} \in \{-1, 1\}$ was a binary variable indicating whether or not the animal was rewarded on the previous trial. Again, $X_{t,4} \in \{-1, 1\}$.

Response time data. In Fig. 6, we analyzed response times for IBL animals, as reaction time data were unavailable for most sessions. The mouse's response time is the time from the stimulus onset to the time at which it received feedback on its decision (receiving its reward or hearing an auditory cue to indicate an error trial). In comparison, a reaction time for the IBL task is the time from the stimulus appearing on the screen to the mouse moving the wheel for the first time in the direction of its decision. Response times ranged from a few hundred milliseconds to 10 s in the most extreme case, whereas typical reaction times (calculated on a few sessions for which we had reliable data) tended to be on the order of a few hundred milliseconds. Relative to other tasks, the IBL task is a task that results in short reaction times.

Statistics and reproducibility. We examined the choice data of animals in the IBL dataset with more than 3,000 trials of data. In the case of Odoemene et al. animals, we required 12,000 trials of data. As detailed in the Methods, as well as in the Supplementary Information, we determined that these numbers were sufficient because we found that we could successfully recover the generative parameters of a GLM-HMM in simulated data in the two different parameter regimes with these numbers of trials. All animals with the requisite number of trials of data were included in the analyses that we present in this paper. Each experiment presented in this paper was repeated in multiple animals (37 in the case of the

IBL dataset, 15 for the Odoemene et al. dataset and 27 in the case of the Urai et al. human dataset). The effects identified were largely consistent across subjects: the choice data for all 37 IBL animals, for 15 Odoemene et al. mice and for 24 of the 27 humans were better explained by a GLM-HMM than a classic lapse model. Analysis was performed with code that is freely available to promote replication. We randomized the allocation of sessions to folds for each animal when performing five-fold cross-validation. Blinding was not performed. All trials from animals with the requisite number of trials were analyzed; blinding was not necessary as cross-validation ensured that the results were not due to the particular choice of train/test split.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The raw data studied in this paper are publicly available. The IBL data associated with ref. ¹⁹ can be accessed at <https://doi.org/10.6084/m9.figshare.11636748>. The Odoemene et al. data associated with ref. ²⁰ can be accessed at <https://doi.org/10.14224/1.38944>. Finally, the human data associated with ref. ²⁷ can be accessed at <https://doi.org/10.6084/m9.figshare.4300043>.

Code availability

We contributed code to version 0.0.1 of the Bayesian State Space Modeling framework of ref. ⁷², and we used this code base to perform GLM-HMM inference. The code to analyze the resulting model fits and to produce the figures in this paper is available at <https://github.com/zashwood/glm-hmm>.

References

- Dempster, A. P., Laird, N. M. & Rubin, D. B. Maximum likelihood from incomplete data via the EM algorithm. *J. R. Stat. Soc. Series B Stat. Methodol.* **39**, 1–22 (1977).
- McLachlan, G. J. & Krishnan, T. *The EM Algorithm and Extensions* (Wiley, 2007).
- Bishop, C. M. *Pattern Recognition and Machine Learning* (Springer, 2006).
- Baum, L. E., Petrie, T., Soules, G. & Weiss, N. A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. *Ann. Math. Stat.* **41**, 164–171 (1970).
- Broyden, C. G. The convergence of a class of double-rank minimization algorithms 1. General considerations. *IMA J. Appl. Math.* **6**, 76–90 (1970).
- Fletcher, R. A new approach to variable metric algorithms. *Comput. J.* **13**, 317–322 (1970).
- Goldfarb, D. A family of variable-metric methods derived by variational means. *Math. Comput.* **24**, 23–26 (1970).
- Shanno, D. F. Conditioning of quasi-Newton methods for function minimization. *Math. Comput.* **24**, 647–656 (1970).
- Virtanen, P. et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* **17**, 261–272 (2020).
- Salakhutdinov, R., Roweis, S. T. & Ghahramani, Z. Optimization with EM and expectation-conjugate-gradient. In: *Proceedings of the 20th International Conference on Machine Learning (ICML-03)*, 672–679 (Association for Computing Machinery, 2003).
- The International Brain Laboratory et al. Data architecture for a large-scale neuroscience collaboration. Preprint at <https://www.biorxiv.org/content/10.1101/827873v3> (2020).
- Linderman, S., Antin, B., Zoltowski, D. & Glaser, J. SSM: Bayesian learning and inference for state space models <https://github.com/lindermanlab/ssm> (2020).
- Yu, S.-Z. Hidden semi-Markov models. *Artificial Intelligence* **174**, 215–243 (2010).

Acknowledgements

We are grateful to M. Wells, R. Terry, L. Funnell and the Cortexlab at University College London for providing us with the data for the four mice plotted in Extended Data Fig. 3. We are grateful to S. Linderman for developing the beautiful Bayesian State Space Modeling framework of ref. ⁷²; as described in our Methods section, we built our code on top of this framework. We thank members of the Pillow Laboratory, the International Brain Laboratory (IBL) and specifically the Behavior Analysis Working Group within the IBL for helpful feedback throughout the project. We thank P. Dayan, S. Bruijns and L. Paninski for acting as the IBL Review Board for this paper. We thank E. Dennis for feedback at various points during the project. We thank A. Russo and M. Whiteway for providing feedback on drafts of this manuscript. We thank H. Bayer for help and advice as we were preparing to submit this paper. Finally, we thank the anonymous reviewers for their insightful comments; our manuscript is greatly improved as a result of their input.

This work was supported by grants SCGB AWD543027 (J.W.P.) and SCGB AWD543011 (A.P.) from the Simons Collaboration on the Global Brain; grants NS104899 (J.W.P.), R01EB026946 (J.W.P.) and R01EY022979 (A.K.C.) from the National Institutes of Health BRAIN initiative; grant U19 NIH-NINDS BRAIN Initiative Award 5U19NS104648 (J.W.P.); grant 315230_197296 from the Swiss National Fund (A.P.); and Wellcome Trust grants 209558 (the IBL) and 216324 (the IBL). A.E.U. was supported by the German National Academy of Sciences Leopoldina and the International Brain Research Organization. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Author contributions

Conceptualization: Z.C.A. and J.W.P.; Methodology: Z.C.A. and J.W.P.; Additional Technical and Analysis Support: N.A.R., I.R.S., A.E.U., A.K.C. and A.P.; Implementation: Z.C.A.; Data Collection and Curation: A.E.U., A.K.C. and the International Brain Laboratory (IBL); Writing—Original Draft: Z.C.A. and J.W.P.; Writing—Review and Editing: Z.C.A., N.A.R., I.R.S., the IBL, A.E.U., A.K.C., A.P. and J.W.P.; Visualization: Z.C.A. and J.W.P.; Supervision: A.P. and J.W.P.; Project Administration: Z.C.A. and J.W.P.; Funding Acquisition: the IBL, A.K.C., A.P. and J.W.P.

Competing interests

The authors declare no competing interests.

Additional information

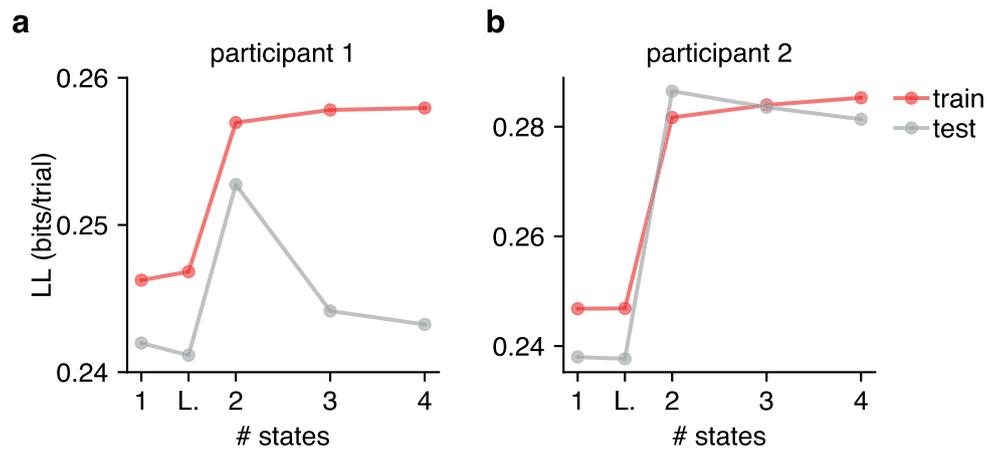
Extended data is available for this paper at <https://doi.org/10.1038/s41593-021-01007-z>.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41593-021-01007-z>.

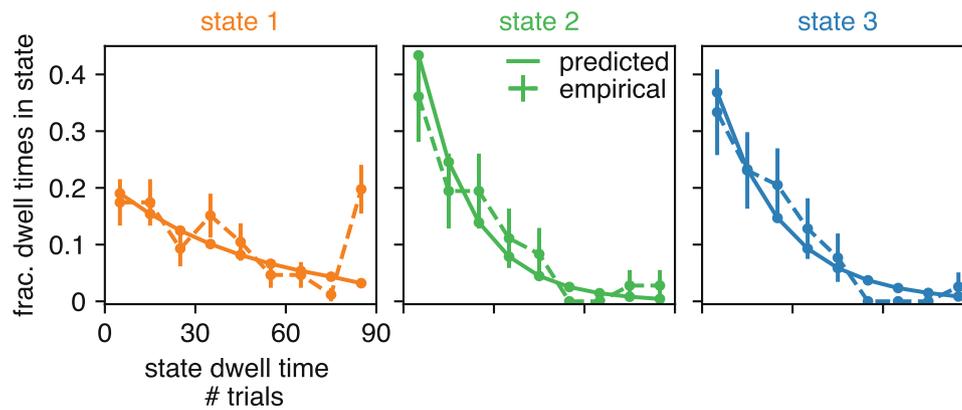
Correspondence and requests for materials should be addressed to Zoe C. Ashwood or Jonathan W. Pillow.

Peer review information *Nature Neuroscience* thanks Timothy Behrens and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

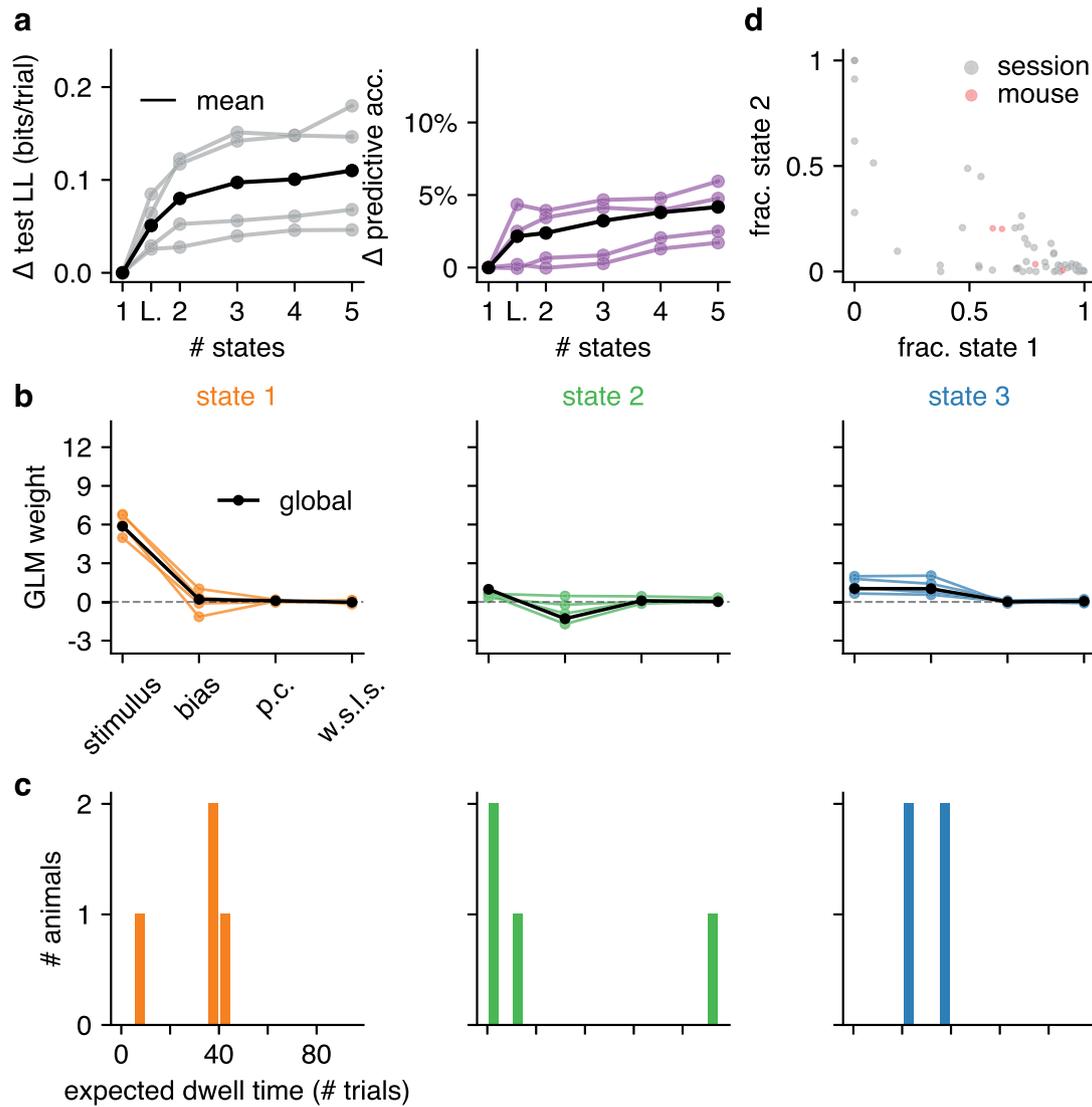
Reprints and permissions information is available at www.nature.com/reprints.



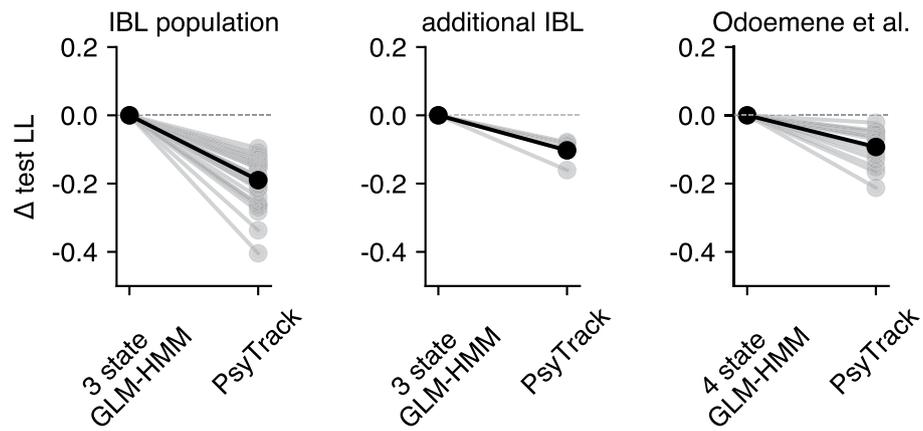
Extended Data Fig. 1 | Cross-validation for Model Selection. (a,b) Here we show the cross-validated train and test loglikelihood in units of bits per trial for two humans performing the task of²⁷ as a function of the number of states. While the number of parameters of the GLM-HMM increases as the number of states is increased (and a two state GLM-HMM has more parameters than both the classic lapse model and single state GLM), only the loglikelihood of the training dataset (red) is guaranteed to increase as the number of parameters increases. Indeed, as the number of states increases, the GLM-HMM may start to overfit the training dataset causing it to fit the test dataset poorly. This is what we see here when the grey curves begin to decrease in each of the two figures as the number of states increases. Thus, by observing the performance of the model on a test dataset, we can appropriately trade off predictive performance with model complexity.



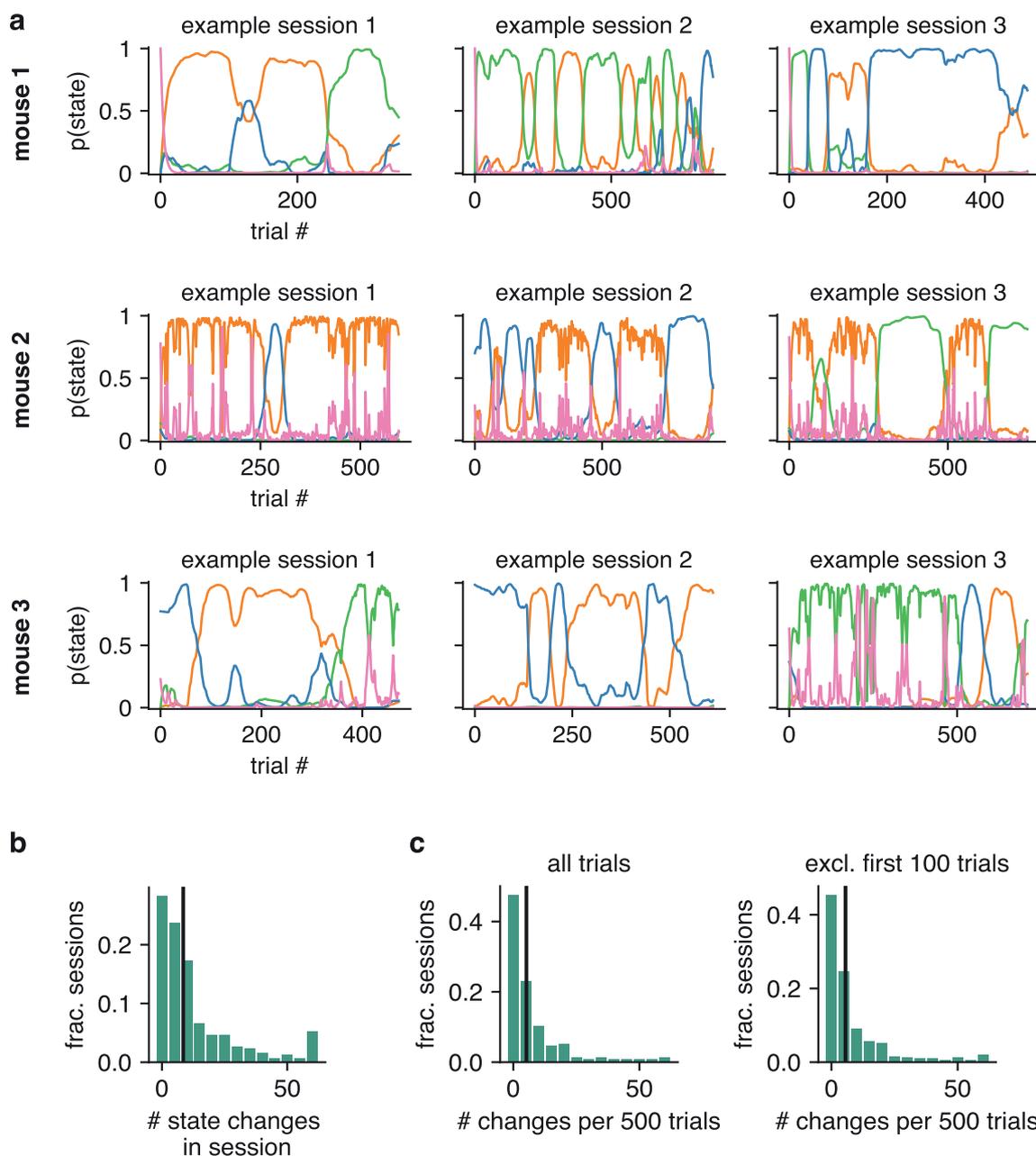
Extended Data Fig. 2 | Retrieved state dwell times are approximately geometrically distributed. With the solid line, we show the predicted dwell times (according to the retrieved transition matrix) in each of the three states for the example animal of Fig. 2 and Fig. 3. Predicted dwell times can be obtained from the transition matrix as $p(\text{dwell time} = t) = (1 - A_{kk})A_{kk}^{t-1}$ because state dwell times in the Hidden Markov Model are geometrically distributed. We then use the posterior state probabilities to assign states to trials in order to calculate the dwell times that are actually observed in the real data (shown with the dashed line); we also show the 68% confidence intervals associated with these empirical probabilities (n is between 36 and 86, depending on state). We find that the empirical dwell times for the biased leftward and rightward states seem to be geometrically distributed. For the engaged state, because there are some entire sessions (each session is 90 trials) during which the animal is engaged, we see that the empirical dwell times associated with this state are not as well described by a geometric distribution. A future direction may be to allow non-geometrically distributed state dwell times by replacing the Hidden Markov Model with e.g. the Hidden semi-Markov Model⁷³.



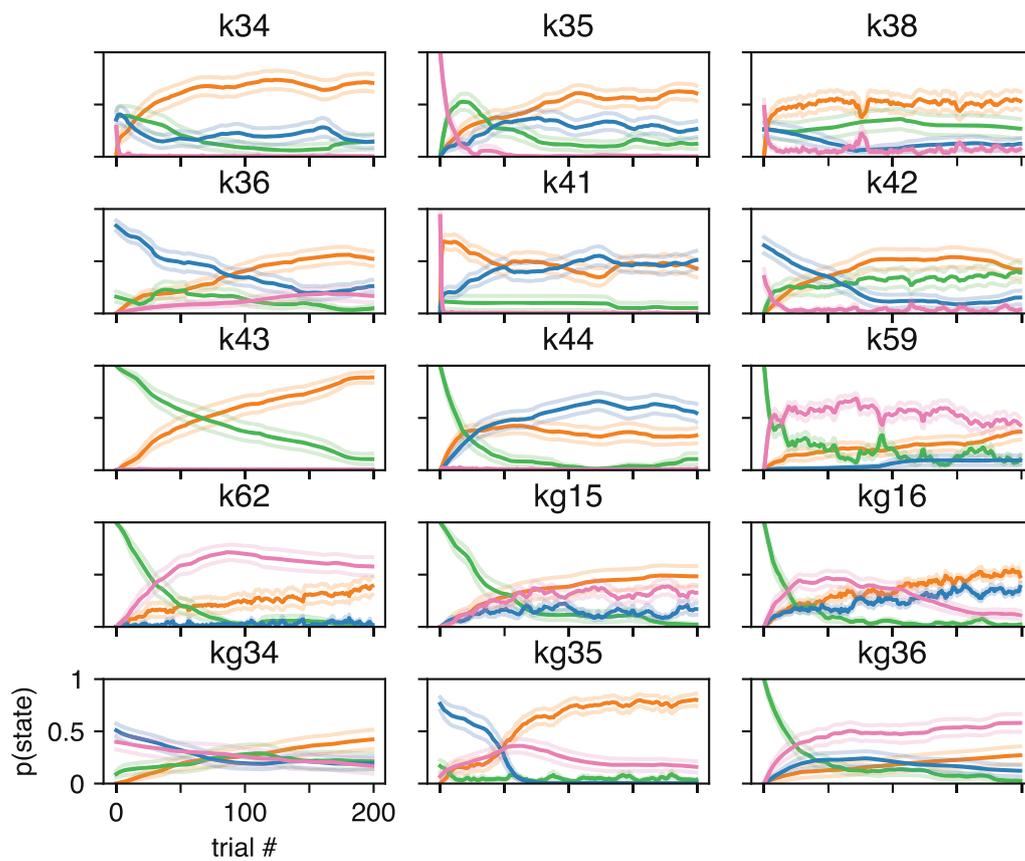
Extended Data Fig. 3 | GLM-HMM application to 4 mice not exposed to bias blocks in IBL task. We confirm that mice that have never been exposed to bias blocks in the IBL task continue to show state-dependent decision-making. This is a sister figure to Fig. 4, and each panel can be interpreted in the same way as in Fig. 4.



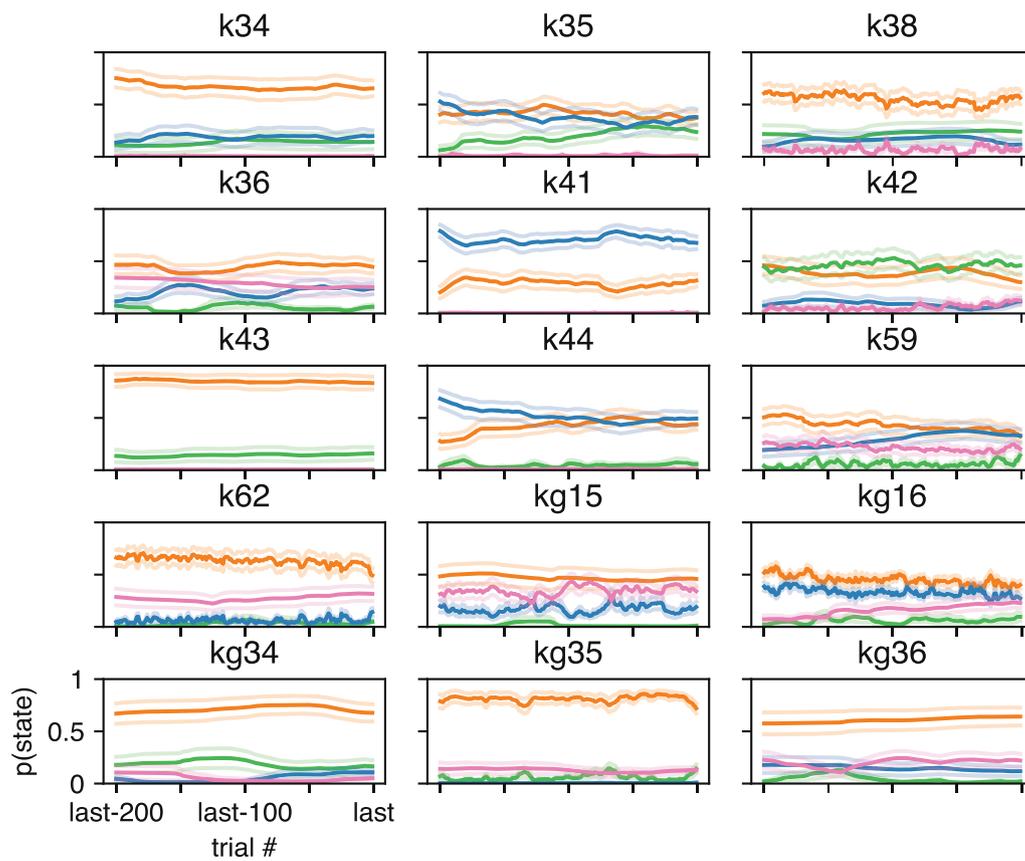
Extended Data Fig. 4 | Additional comparisons with PsyTrack model. (a) Copy of left panel from Fig. 4f showing the difference in test set loglikelihood for 37 IBL animals for the 3 state GLM-HMM compared to the PsyTrack model of^{35,36}. Black indicates the mean across animals. The 3 state GLM-HMM better explained the choice data of all 37 animals compared to the PsyTrack model with continuously evolving states. **(b)** Analogous figure for the 4 additional IBL animals studied in Fig. ED3. All 4 animals' data were better explained by the GLM-HMM compared to PsyTrack. **(c)** Same as in panel a for Odoemene et al. animals shown in Fig. 5, although comparison now utilizes 4 state GLM-HMM fits. All 15 animals' data were better explained by the GLM-HMM compared to PsyTrack.



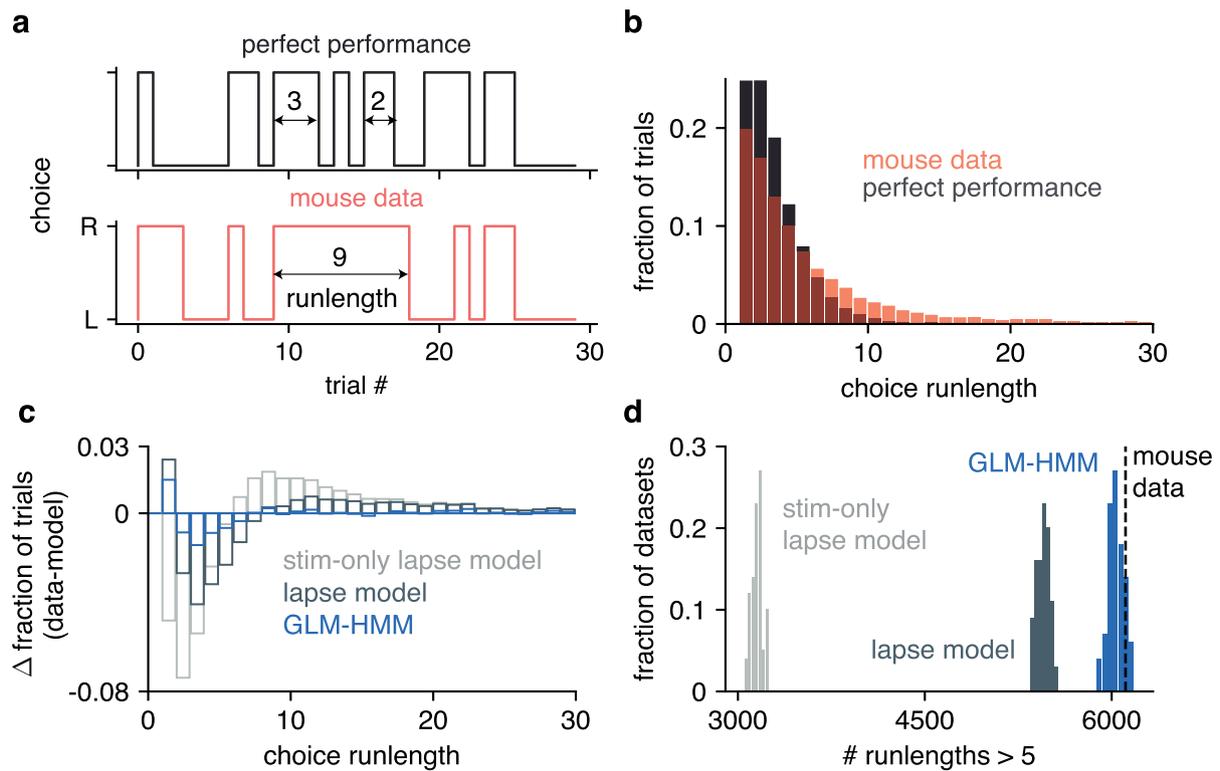
Extended Data Fig. 5 | State switching in Odoemene et al. task. (a) Posterior state probabilities for three example sessions for three example mice (different mice are shown in each row). **(b)** Histogram giving number of state changes (identified with posterior state probabilities) per session for all sessions across all animals. For visibility, state changes are censored above 60. **(c)** Different sessions have different numbers of trials, so we normalize the histogram of **b** to give the number of state changes per 500 trials for each session (the median session length is 683 trials). Again, for visibility, state changes are censored above 60. Left: we use all data from all trials to generate the normalized histogram. Right: we plot the normalized histogram when we exclude the first 100 trials of a session. As can be observed, the left and right normalized histograms are very similar (p -value = 0.96 using KS-test). While the GLM-HMM is able to capture ‘warm-up’ effects (as described in the main text), this test reveals that the GLM-HMM is able to capture more than this, and state switching occurs much later in the session too (as also indicated by the posterior state probabilities shown in **a**).



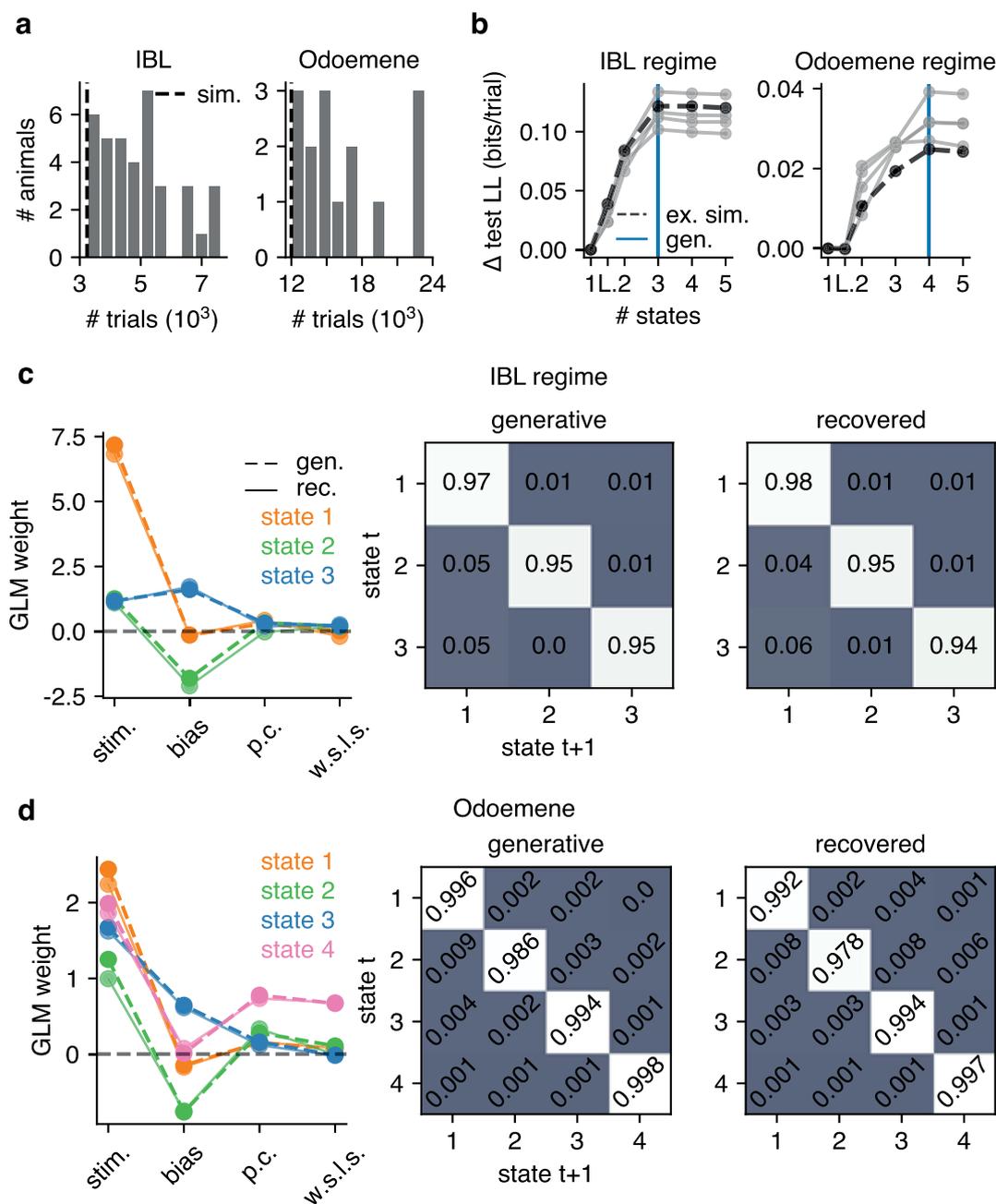
Extended Data Fig. 6 | GLM-HMM captures 'warm-up' effect for Odoemene et al. animals. Average (across 20 sessions) posterior state probabilities for the first 200 trials of a session for each animal in the Odoemene et al. dataset. Orange corresponds to the engaged state, green to the biased left, blue to the biased right and pink to the win-stay state from Fig. 5. Error bars represent standard errors.



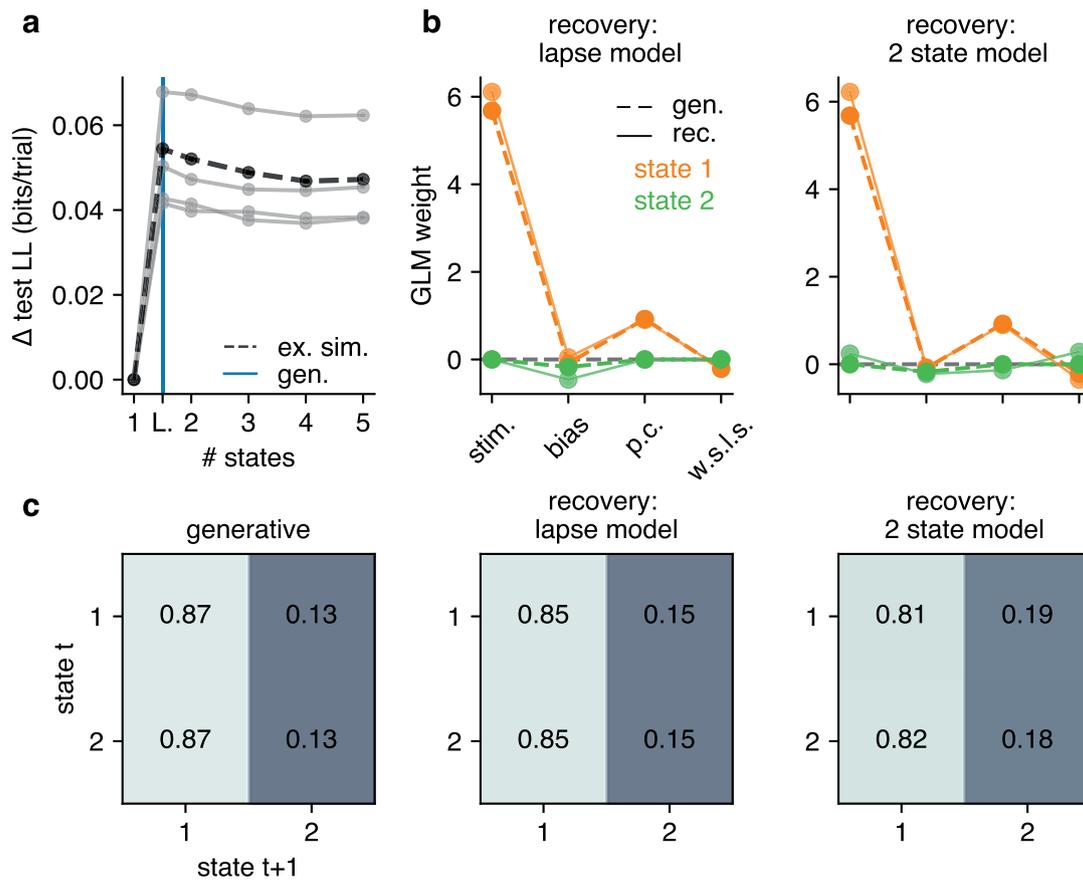
Extended Data Fig. 7 | GLM-HMM posterior state probabilities at end of session. Average (across 20 sessions) posterior state probabilities for the last 200 trials of a session for each animal in the Odoemene et al. dataset. Orange corresponds to the engaged state, green to the biased left, blue to the biased right and pink to the win-stay state from Fig. 5. Error bars represent standard errors.



Extended Data Fig. 8 | Simulated data from GLM-HMM captures statistics of real choice data. (a) Definition of choice run-length. Shown are the choices that an IBL mouse made over the course of 30 trials (red, bottom), as well as the choices it should have made during that same time course if the mouse performed the task perfectly (grey, top). Choice run-length is defined as the number of trials during which a mouse repeated the same decision (example choice run-lengths of 2, 3 and 9 trials are highlighted). **(b)** Red: fraction of trials in choice run-lengths of between 1 and 30 trials when calculated from all trials for all mice. Grey: distribution of choice run-lengths that would have been obtained if IBL mice performed the task perfectly. **(c)** Difference in choice run-length distribution for simulated data (from three different models) compared to the red distribution shown in (b). Models used to simulate data were a lapse model with only stimulus intensity and bias regressors, a lapse model that also included history regressors (previous choice and win-stay-lose-switch), and a 3 state GLM-HMM (also with history regressors). We simulated 100 example choice sequences from each model and calculated the mean histogram of choice run-lengths across the 100 simulations. This was then subtracted from the red histogram of (b). **(d)** Number of choice run-lengths with more than 5 trials for each model simulation used in (c). In the 181,530 trials of real choice data, there were 6111 run-lengths lasting more than 5 trials (as shown with the dashed line). When we simulated choice data according to each of the models shown in (c), we found that only the GLM-HMM could generate simulations with as many run-lengths lasting more than 5 trials as in the real data (15/100 simulations had 6111 or more run-lengths lasting more than 5 trials for the GLM-HMM compared to 0/100 for both of the lapse models).



Extended Data Fig. 9 | GLM-HMM Recovery Analysis 1. For dataset sizes comparable to those of real animals, we can recover the IBL and Odoemene global parameters in simulated data. **(a)** Dataset sizes for each of the 37 IBL animals studied (left) and 15 mice from Odoemene et al. (right). The dashed vertical line indicates the number of trials that we used in simulation data in panels b, c and d (3240 for the IBL parameter regime and 12000 for the Odoemene regime simulation). **(b)** Test set loglikelihood for each of 5 simulations is maximized at 3 states (blue vertical line) after we simulate according to the (IBL regime) parameters shown in panel c. Similarly, in the right panel, test set loglikelihood is maximized at 4 states when we simulate choice data with the (Odoemene regime) parameters shown in panel d. The thick black line marked as 'ex. sim.' (example simulation) indicates the simulation whose generative and recovered parameters we show in panels c and d. **(c)** Left: the generative and recovered GLM weights (for the simulation marked as 'ex. sim.' in panel b) when we simulate choice data in the IBL parameter regime. Middle and right: the generative and recovered transition matrices. **(d)** The generative and recovered parameters in the Odoemene et al. parameter regime.



Extended Data Fig. 10 | GLM-HMM Recovery Analysis 2: We can recover lapse behavior. (a) We simulate 5 datasets, each with 3240 trials, according to the best fitting lapse model for IBL animals. We then fit these simulated datasets with GLM-HMMs, as well as a lapse model (a constrained 2 state GLM-HMM). The test set loglikelihood is highest for the lapse model in all simulations, indicating that lapse behavior can be distinguished from the long-enduring multi-state behavior that best described the real data. The thick black line marked as 'ex. sim.' (example simulation) indicates the simulation whose generative and recovered parameters we show in panels b and c. (b) Left: the generative and recovered weights when recovery is with a lapse model. Right: the generative weights are the same as in the left panel, but we now recover with an unconstrained 2 state GLM-HMM (thus the stimulus, previous choice and w.s.l.s. weights for the second state can be non-zero) (c) The generative (left) transition matrix and the recovered transition matrices when we recover with a lapse model (middle) and an unconstrained 2 state GLM-HMM (right). While the lapse model and 2 state GLM-HMM results don't perfectly agree, if mice were truly lapsing, the transition matrix would not have the large entries on the diagonals that we observe in the real data.

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection We did not use any software for data collection (we analyzed three publicly available datasets that were already curated).

Data analysis We contributed code to version 0.0.1 of the Bayesian State Space Modeling framework (<https://github.com/lindermanlab/ssm>) of Linderman, Antin, Zoltowski and Glaser (2020). We used this code base to perform GLM-HMM inference. Our code to analyze the resulting model fits and to produce the figures in this paper is available at <https://github.com/zashwood/glm-hmm> (version 0.0.1). We used version 1.3 of the PsyTrack model of Roy et al. (<https://github.com/nicholas-roy/psytrack>) when performing model comparisons with this model. Finally, We used Python 3.8.5 and SciPy 1.5.2 throughout our work.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

The raw data studied in this paper is publicly available: the IBL data can be accessed at <https://doi.org/10.6084/m9.figshare.11636748>. The Odoemene et al. data can be accessed at <https://doi.org/10.14224/1.38944>. Finally, the human data associated with can be accessed at <https://doi.org/10.6084/m9.figshare.4300043>.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

Life sciences Behavioural & social sciences Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	We examined the choice data of animals in the IBL dataset with more than 3000 trials of data. In the case of Odoemene et al. animals, we required 12,000 trials of data. As detailed in the Methods section, as well as in the supplementary information, we determined that these numbers were sufficient because we found that we could successfully recover the generative parameters of a GLM-HMM in simulated data in the two different parameter regimes with these numbers of trials.
Data exclusions	As detailed above, we restricted ourselves to analyzing animals with sufficiently many trials (3000 or 12,000 depending on the dataset). All animals with the requisite number of trials of data were included in the analyses we present in the paper.
Replication	Each experiment presented in the paper was repeated in multiple animals (37 in the case of the IBL dataset, 15 for the Odoemene et al. dataset and 27 in the case of the Urai et al. human dataset). The effects identified were largely consistent across animals: the choice data for all 37 IBL animals, 15 Odoemene et al. mice and for 24 out of the 27 humans were better explained by a GLM-HMM compared to a classic lapse model. Analysis was performed with code that is freely available to promote replication.
Randomization	We randomized the allocation of sessions to folds for each animal when performing 5-fold cross-validation.
Blinding	Blinding was not performed. All trials from animals with the requisite number of trials were analyzed; blinding was not necessary as cross-validation ensured that the results were not due to the particular choice of train/test split.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern

Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging