Internal state is at least two dimensional

Jaweria Amjad

jaweria.amjad@ucl.ac.uk

UCL

Naoki Hiratani

Washington University

Peter Latham

UCL https://orcid.org/0000-0001-8713-9328

Article

Keywords:

Posted Date: February 11th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-7749081/v1

Additional Declarations: There is **NO** Competing Interest.

Internal state is at least two dimensional

Jaweria Amjad[1*], Naoki Hiratani[2] and Peter E. Latham[1]

[1]Gatsby Computational Neuroscience Unit, University College London, London, UK.
[2]Department of Neuroscience, Washington University in St Louis, St Louis, USA.

*Corresponding author(s). E-mail(s): jaweria.amjad@ucl.ac.uk;

Abstract

Normative models of behavior typically assume that the goal of animals is to optimize some measure of reward. It is well known, however that animals don't always conform to this ideal. Especially in tasks involving repeated trials, attention waxes and wanes: sometimes performance is near optimal; other times it's near chance. Such variation in behavior must be driven by changes in the internal state of the animal. To gain insight into this process, we examined three potential measures of internal state in a two-alternative forced choice task: the complexity of physical movements, the degree to which animals make use of the stimulus, and pupil diameter. Pupil diameter was almost independent of behavior, whereas the first two measures were strongly correlated with it – but in very different ways. This suggests that the internal state is at least two dimensional. In our experiments those dimensions corresponded to how much animals focus on the current stimulus, and how much they use their priors over task variables. These results add insight into internal states in general, and also provide a cautionary tale: internal state is not one-dimensional, and it may take several measurements to infer it accurately.

1 Introduction

In behavioral tasks in which animals perform the same action repeatedly, performance often varies widely: periods of near-optimal performance are interleaved with periods that resemble random guessing. A widely held explanation for this variability is that it reflects fluctuations in the animal's internal state, such as shifts in attention, motivation, or arousal. Here we ask: what is the nature of the internal state?

This is not a new question, and a number of approaches have been proposed. These fall into two main categories: ones based purely on behavior [1–5], and ones based on autonomic responses, such as pupil diameter [6–8]. Here, rather than considering only one approach, we consider two behavioral measures and one autonomic response, all within the context of the International Brain Laboratory [9]. The behavioral measures are a new measure we introduce based on complex movements, and the degree to which mice make use of the stimulus, as assessed by the GLM-HMM model [2, 3]; the autonomic response is pupil diameter.

Do the three methods measure the same thing? To answer that, we determined how the internal state, as inferred from the three methods, maps to various measures of performance on the task. Excluding pupil diameter, which was weakly correlated with behavior, we found that the mappings were very different, indicating that the methods measure different things. We argue that the differences arise not because one is right and the others are wrong, but because internal state has at least two axes: how much animals pay attention to the current trial, as measured by the slope of the psychometric curve, and how much they pay attention to past trials, as measured by their ability to infer the current context.

2 Results

2.1 International Brain Lab Task

We apply our analysis to the International Brain Lab (IBL) task [9] – a classic two-alternative forced choice task. Head-fixed mice are presented with visual stimuli consisting of black-and-white gratings, displayed either on the left or right side of a screen (Fig. 1a). The mice move the grating by turning a wheel. Movements that guide the stimulus to the center of their visual field, corresponding to a wheel rotation of 0.3 radians in the correct direction, are rewarded with 1–3 μL of glucose water; movements that displace the stimulus to the periphery, corresponding to a wheel rotation of 0.3 radians in the wrong direction, are classified as incorrect (typical wheel trajectories are shown in Fig. 1b). If neither a correct nor incorrect threshold is reached within 60 seconds, the trial is classified as a time-out.

The difficulty of each trial is determined by the contrast of the grating. Contrasts are drawn from the set $\pm\{0, 0.0625, 0.125, 0.25, 1.0\}$, with positive values indicating stimuli on the right and negative values indicating stimuli on the left. Each session starts with 90 unbiased trials, where stimuli appear equally often on the left and right. After that, trials are presented in biased blocks: the stimulus appears on one side with 80% probability and on the other side with 20% probability. Block lengths vary between 20 and 100 trials, sampled from a truncated exponential distribution. The block transitions are not signaled; mice must infer the current bias from past trials. Zero-contrast trials have no intrinsically correct choice. To encourage mice to rely on prior information, these trials are rewarded according to the block bias: in a right-biased block, choosing right yields reward with 80% probability and left with 20%, and vice versa in a left-biased block.

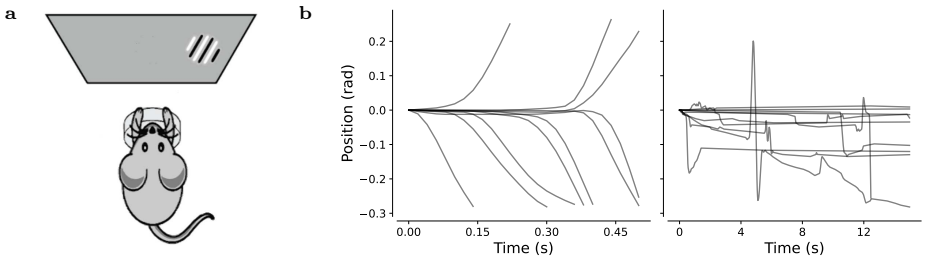

Fig. 1: The IBL task. **a**. A grating appears on the screen, and to receive a reward the mice have to move a wheel such that the grating moves to the center. **b**. Left: sample trajectories characteristic of the rapid, ballistic movements, where decisions are executed within a fraction of a second. Right: Longer, more variable trajectories, potentially indicative of disengagement in the task.

2.2 Potential measures of internal state

Here we describe the three potential measures of internal state that we use in our analysis. The first is based on wheel movements, which we compress into a relatively low dimensional latent variable. The second is based purely on the choices the animals make, under the assumption that the internal state changes slowly. The third is based on an autonomic response, pupil diameter.

2.2.1 Wheel movements

Intuitively, we expect an animal's internal state to affect its movements: when an animal is engaged its movements should be fast and deliberate; when not engaged movements should be slow and relatively random. Quantifying this expectation is, however, nontrivial, as movements are typically high-dimensional, and so not easy to categorize. Here we focus on wheel movements, which have the added complexity that they are variable length. To address the high-dimensionality and variability in length, we use a Variational Autoencoder (VAE) [10] to map the trajectories into a low dimensional latent space, enabling a tractable representation of their underlying structure.

Our VAE consists of an encoder that maps wheel trajectories, $\theta(t)$, to a Gaussian distribution, and a decoder that produces a reconstructed trajectory, $\hat{\theta}(t)$, from a sample drawn from that distribution (see Fig. 2). Details are given in Methods, Sec. M.3.1; here we provide a brief summary. The encoder consists of two stacked LSTM layers (which naturally handle time-dependent input), followed by a feedforward ReLU network; this produces the mean, $\boldsymbol{\mu}$, and the diagonal covariance, $\boldsymbol{\Sigma}$, of an 8-dimensional Gaussian latent distribution. The decoder starts with a latent vector sampled from the Gaussian distribution $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$; that is fed through two LSTM layers followed by a feedforward ReLU network.

The model parameters are adjusted to minimize the difference between the input and output sequence (the reconstruction loss), with a penalty to encourage latent variables to have high entropy. This leads to the loss

$$\mathcal{L} = \frac{1}{N}\sum_{n=1}^{N}\int_{0}^{T_n} dt\,\left(\theta_n(t) - \hat{\theta}_n(t)\right)^2 + \frac{\lambda}{2}\sum_{m=1}^{8}\left(\mu_m^2 + \Sigma_{mm} - \log(\Sigma_{mm}) - 1\right) \quad (1)$$

where n denotes trial, t denotes time within a trial, $\theta_n(t)$ is the input wheel trajectory and $\hat{\theta}_n(t)$ is the output of the VAE, T_n is the duration of trial n, and N is the number of trials. To handle variable length trajectories, time points in the input sequences beyond the trial length are padded with a number outside the range of wheel trajectories, and the reconstruction loss is masked so that only the actual trajectory contributes to the error; see Methods, Sec. M.3.1 for details. As shown in Methods, Sec. M.3.2, minimizing the loss is equivalent to maximizing the Evidence Lower Bound (ELBO) on the log likelihood of the data.

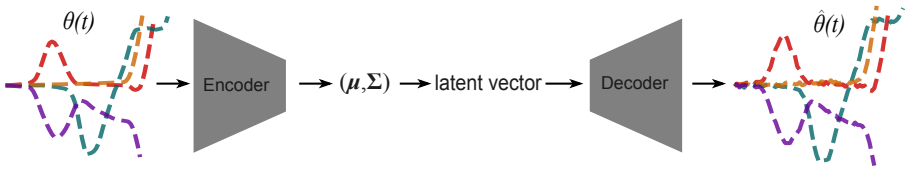

Fig. 2: Block diagram of the VAE model; see Methods, Sec. M.3.1 for additional details. Despite the stochastic mapping from $(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ to the latents, after training the output, $\hat{\theta}(t)$, is remarkably similar to the input, $\theta(t)$.

We trained our Variational Autoencoder on wheel trajectories using 2,415 sessions from 81 mice (see Methods, Sec. M.1 for details on data selection and inclusion criteria). After training, for each trial the VAE mapped the trajectory to a latent Gaussian distribution, as described above. To identify the most informative features, we performed principal component analysis on the latent representations, which we took to be the 16-dimensional vector consisting of $(\mu_i, \log \Sigma_{ii},\ i = 1, ..., 8)$. The first two principal components explained 89% of the total variance (62% and 27%, respectively); we used these components for all of our analysis. We rotated the principal components by 30° and flipped the x-axis to align the latent space with behavioral variables, maximizing the correlation of one axis with choice direction while minimizing its correlation with feedback.

Figure 3a shows the distribution in the space of the first two (rotated and flipped) principal components; wheel trajectories sampled from this distribution are shown in Fig. 3b. As can be seen from the latter figure, large values of the x-coordinate correspond to nearly ballistic wheel movements while small values correspond to slow, meandering trajectories, and large values of the y-coordinate correspond to right choices while small values correspond to left choices. Based on the idea that ballistic movements are a signature of an engaged state, we refer to the value of the x-coordinate as the (z-scored) engagement index. Because we are interested in engagement, in subsequent analysis we ignore the wheel direction and focus on the engagement index.

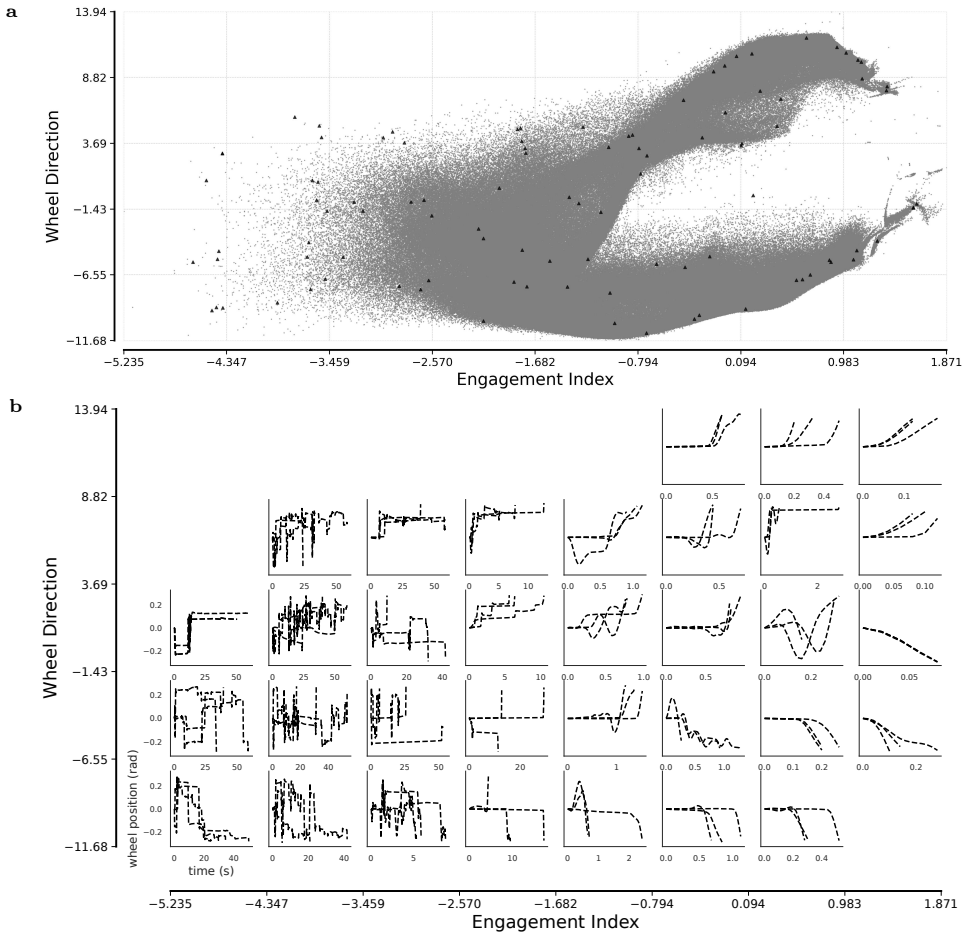

Fig. 3: **a**. Scatterplot of the z-scored engagement index versus wheel direction, both obtained by applying a linear transformation to the first two principal components of the VAE representation (see main text for details). The black dots correspond to the trajectories plotted in panel b. **b**. Trajectories corresponding to various values of the first two principal components. The engagement index is z-scored to facilitate comparison with other metrics.

2.2.2 The GLM-HMM model

Because the choices an animal makes depend on its internal state, it should be possible to infer internal state purely from choices. Conceptually, this is straightforward: build a model in which the current choice (in the case of the IBL task, right or left movement of the wheel) depends parametrically on current and past sensory input, past choices, and internal state; infer the parameters of the model; and then use those parameters to estimate the internal state on each trial. Practically, though, this is difficult because

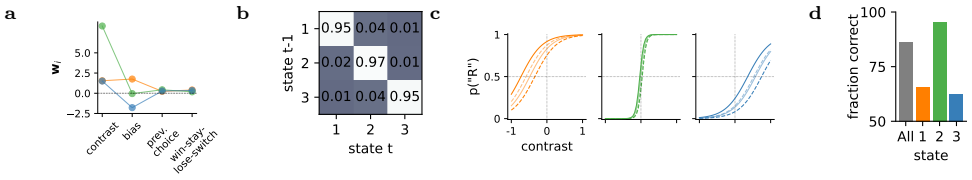

Fig. 4: The GLM-HMM model for an example mouse **a**. GLM weights corresponding a 3-state model. State 2 (green line) exhibits a strong stimulus weight, consistent with an "engaged" decision-making mode. States 1 and 3 (orange and blue lines) show weak stimulus dependence and strong rightward (state 1) and leftward (state 3) biases. **b**. Probability matrix, A_{ij} (see Eq. (2)) dictating state transition. Large diagonal values indicate strong state persistence. **c**. Psychometric curves generated using the fitted GLM-HMM. The four curves are the four combinations of choice and the outcome of the previous trial; more concisely, they are plots of $\phi(w_{i1}c + w_{i2} \pm w_{i3} \pm w_{i4})$ versus c for the four combinations of $\pm$; see Eq. (3). **d**. Overall accuracy (gray) and state-specific accuracies (using the same color code as in panels a and c).

the internal state predicts only the probability of a correct choice, and probability cannot be measured on a trial by trial basis.

To get around this practical difficulty, the typical assumption is that the internal state doesn't change too rapidly. This is the approach taken by the GLM-HMM model [2, 3], which assumes that there are discrete hidden states that change rarely, and the degree to which the mouse pays attention to the sensory input is state-dependent. Loosely, there are "engaged" states, where performance is near optimal given sensory input, and "disengaged" states, where performance is suboptimal, and often biased toward one side or another. For a complete description of the GLM-HMM, see [2, 3]; here we describe its basic elements, as they will be necessary to understand the analysis in Sec. 2.3 below.

Given its name, it's not surprising that the GLM-HMM combines a hidden Markov model (HMM) with a generalized linear model (GLM) [2–4]. The Hidden Markov Model is characterized by a state transition matrix, denoted A_{ij}, which determines transitions across trials,

$$A_{ij} = p(\text{state } i \text{ on current trial} \mid \text{state } j \text{ on previous trial}). \tag{2}$$

Each state defines a distinct decision strategy via a weight vector, $\mathbf{w}_i$: the probability of choosing right given state i is given by

$$p(\text{choose right}|\text{state } i) = \phi(w_{i1}c_i + w_{i2} + w_{i3}\text{ch}_{i-1} + w_{i4}\text{wsls}_{i-1}) \tag{3}$$

where $\phi(z) \equiv e^z/(1+e^z)$ is the standard sigmoidal function, c_i is the signed contrast, ch_i represents choice ($+1$ for right and -1 for left), and wsls_i encourages a win-stay-lose-switch strategy ($+1$ for a rewarded right choice or unrewarded left choice; -1 for a rewarded left choice or unrewarded right choice). The second weight, w_{i2}, for which x_{i1} is always 1, represents bias.

We trained 3- and 4-state GLM-HMM models using the same sessions and mice we used for the VAE (see Methods, Sec. M.1 for details on data selection and inclusion criteria). In the main text we report results from the 3-state model, as they are representative; the 4-state model yielded highly similar results (see Supporting Information). The training yielded, for each mouse, a set of transition probabilities, A_{ij}, and weights, $\mathbf{w}_i$. Weights and transition probabilities from the 3-state model are plotted in Fig. 4a and 4b for a typical mouse; note that each color in Fig. 4a corresponds to a vector of weights for a different state. The first component of the weight (w_{i1} in state i), denoted "contrast" in Fig. 4a, is the GLM coefficient of the signed stimulus contrast; thus, high w_{i1} suggests high task engagement. This is reflected in the psychometric curves (Fig. 4c) and in state accuracy (Fig. 4d).

Once we know the parameters for a particular session, we can use the choices of the animal to determine, on each trial, the probability, p_i, that an animal is in each state, as described in [2, 3]. We then compute the degree to which the mouse is using the stimulus, denoted $s(n)$, via

$$s(n) = \sum_i p_i(n) w_{i1} \tag{4}$$

where n denotes trial and $p_i(n)$ is the posterior probability of being in state i on trial n, as inferred by the GLM-HMM model. We call $s(n)$ the "stimulus coefficient", and it's what we use in our analysis in the main text. An alternative measure of engagement is the probability of being in the engaged state (Methods, Sec. M.4). As we show in Supporting information, that gives nearly identical results as the stimulus coefficient.

2.2.3 Pupil diameter

Pupil diameter is known to be associated with arousal and attention [6–8, 11]. Although its relationship with these internal states might not be monotonic [12], it's still a potential indicator of internal state. To extract pupil diameter, we used the Lighting Pose algorithm [13]. We had access to pose on a subset of the trials used to train the GLM-HMM and VAE (145 sessions from 44 mice). To ensured that the recorded pupil data reflected meaningful physiological variability, we included only those sessions in which pupil diameter showed reliable sensitivity to stimulus contrast, as assessed by a stimulus-period pupillometry analysis. That resulted in 93 sessions from 35 mice (see Methods, Sec. M.2, for details).

Pupil diameter does not reflect only internal state; it also changes reflexively in response to visual stimulus features such as luminance and contrast [14–16]. To avoid these visual stimulus-driven responses, we analyzed pupil dilation at 600 ms preceding stimulus onset.

2.2.4 Comparison of the measures

Despite the fact that all three measures are intended to assess internal state, they are almost completely uncorrelated. Figures 5a-c show density plots of the three measures against each other, across all mice and sessions. Visually, no clear relationships are

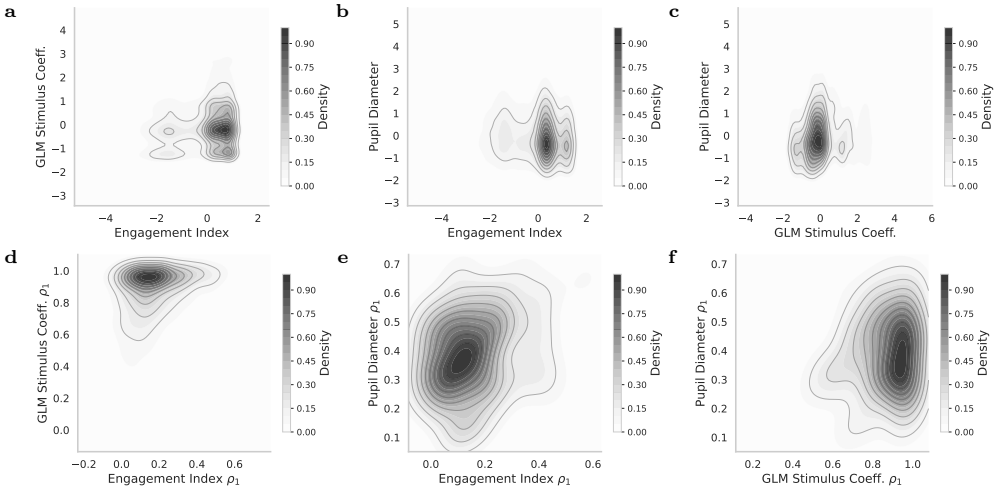

Fig. 5: Comparison of our three measures via pairwise density plots (see Methods, Sec. M.5 for details). For all panels the axes are z-scores. **a-c**: Direct comparison. **d-f**: Comparison of temporal autocorrelation (Eq. (15)).

apparent, and this is confirmed by the low correlation coefficients: 0.061, -0.127 and 0.091 in Figs. 5a-5c respectively.

We also assessed the temporal stability of our three measures using the the autocorrelation across sessions with a trial lag of 1 (see Methods, Sec. M.5; in particular Eq. (15)). As can be seen in Figs. 5d-5f, the three measures have very different temporal structure: for the GLM stimulus the autocorrelation coefficient is near 1, indicating that it is very persistent; for the engagement index it's centered only slightly above zero, indicating that it is short-lived, and for the pupil diameter the autocorrelation is somewhere in-between.

2.3 Task Behavior as a Readout of Internal State

Here we investigate how our potential measures of internal state – engagement index, GLM-HMM stimulus coefficient, and pupil diameter – are related to behavior. In Fig. 6 we plot psychometric curves versus these three quantities. In each panel there are three curves, corresponding, from top to bottom, to the right biased blocks, unbiased blocks, and left biased blocks. We extracted two features from these plots: the slope of the psychometric curves (which we average over the three curves), and the gap between the top and bottom curves at zero contrast. The latter is a measure of how much the mice pay attention to the prior; we refer to it as the prior-induced gap.

Plots of the slope, prior-induced gap, and performance versus z-score are shown in Fig. 7. Both the slope and the gap have different behavior as a function of our measures of internal state. The slope increases rapidly versus stimulus coefficient, but it depends much more weakly on both engagement index and pupil diameter (Fig. 7a). The gap increases rapidly versus the engagement index, generally decreases versus the

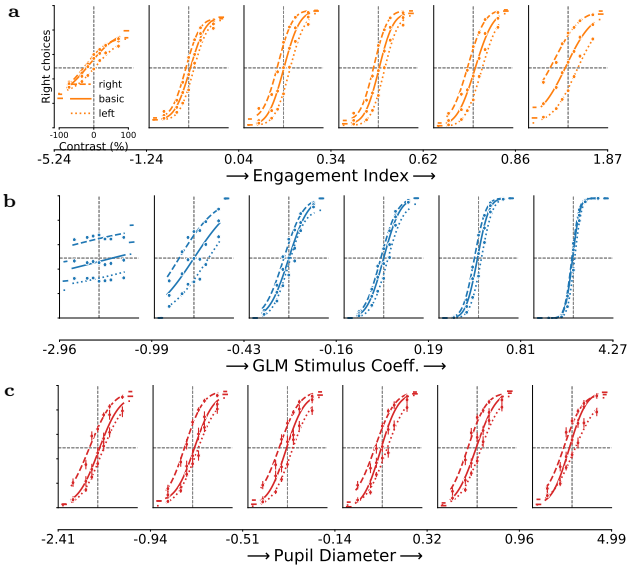

Fig. 6: Psychometric curves averaged over all sessions. The three curves in each panel correspond to the right, unbiased and left blocks, respectively (corresponding to rightward probability of 80%, 50% and 20%). Each panel corresponds to a different quantile of the z-score. **a**. GLM-HMM stimulus coefficient. **b**. Engagement index. **c**. Pupil diameter.

stimulus coefficient, and is more or less flat versus pupil diameter (Fig. 7b). Although the slope and bias behave very differently with respect to the stimulus coefficient and engagement index, the performance on the task is very similar for these two measures (compare the blue and orange lines in Fig. 7c). And consistent with the first two plots, performance is almost completely independent of pupil diameter.

Given that the pupil diameter has very little effect on performance, in what follows we'll mainly ignore it, and focus on the GLM-HMM stimulus coefficient and the engagement index. To better illustrate the relationship between the prior-induced gap and the slope shown in Figs. 7a and 7b, in Fig. 8a we plot these quantities against each other. For the GLM stimulus coefficient (blue), the gap initially increases with slope, followed by a relatively slow decrease, such that the overall range of the prior-induced gap is not very large. The engagement index (orange) behaves very differently. As with the stimulus coefficient, the prior-induced gap initially increases with slope. Here, though, the increase is much more rapid. And after a slope of about 3.5 the gap isn't even single-valued. Moreover, the slope has a much smaller range than the stimulus coefficient, and the prior-induced gap has a much larger range.

The plots in Fig. 8a are somewhat abstract; to make sense of them we build a Bayesian model of the task parameterized by two quantities: observation noise, and the assumed prior probability that the grating appears on the right. The model starts with the assumption that mice observe a variable, ξ, that is probabilistically related

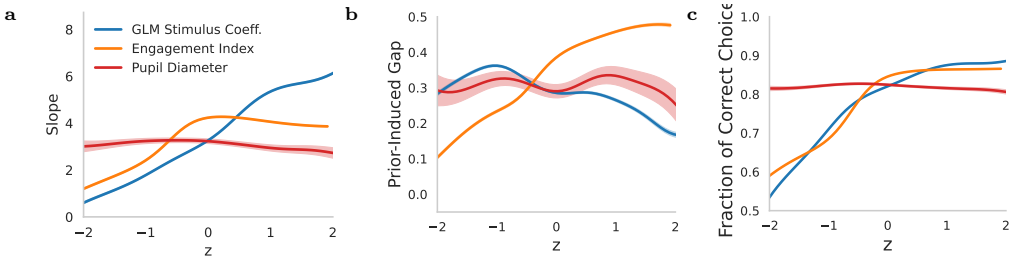

Fig. 7: Behavioral metrics versus our measures of internal state, all versus z-score. Blue: stimulus coefficient; orange: engagement index; red: pupil diameter. Solid lines are mean; shaded regions are standard deviation; see Methods, Sec. M.6 for details. **a**. Slope of psychometric curves at $c = 0$. **b**. Gap between the psychometric curves for the left and the right block at $c = 0$. **c**. Fraction of correct choices.

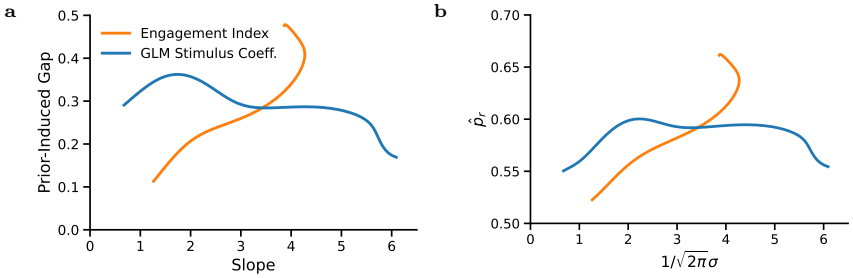

Fig. 8: **a**. Prior-induced gap versus slope. **b**. Inferred prior used by the mice, $\hat{p}_r$, versus $1/\sqrt{2\pi}\sigma$.

to the true contrast, c, via

$$P(\xi|c) = \frac{e^{-(\xi-c)^2/2\sigma^2}}{\sqrt{2\pi\sigma^2}} \tag{5}$$

where σ is the standard deviation of the observation noise. Given this model, it's straightforward to compute psychometric curves of our "Bayesian" mouse (see Methods, Sec. M.7). These curves – and thus the slope and prior-induced gap – depend on the noise, σ, and on the assumed prior that the grating appears on the right hand side, denoted $\hat{p}_r$. We can, therefore, map each point on Fig. 8a (which are parameterized by slope and prior-induced gap) to its corresponding value of σ and $\hat{p}_r$ The mapping from slope to σ is straightforward; it's just $1/\sqrt{2\pi}\sigma$ – exactly what is expected for Gaussian noise. The mapping from gap to assumed prior is more complicated, but it's mainly linear in $(\hat{p}_r - 1/2)$, although with a slope that depends on σ (see Methods, Fig. M.2, and Eq. (36) for an analytic expression).

Figure 8b shows the same curves as in Fig. 8a, but plotted versus $1/\sqrt{2\pi}\sigma$ and $\hat{p}_r$. These curves are nearly identical to the ones in Fig. 8a. This in itself is not especially

surprising. But it's important, as it allows us to interpret the slope and prior-induced gap in terms of easily understood parameters: observation noise and assumed prior that the grating appears on the right side.

The different dependencies for the stimulus coefficient and engagement index illustrated in Figs. 8a and 8b suggest that these quantities measure different things. In particular, a change in the stimulus coefficient mainly causes the slope of the psychometric curve to change, without much affect on the prior used by the animals, whereas a change in the engagement index mainly causes the prior to change. A natural interpretation is that the stimulus coefficient measures how much the mice pay attention to the current stimulus, without much reference to the prior, while the engagement index measure how much the mice pay attention to the prior, with relatively less attention placed on the current stimulus.

Despite their difference, performance versus stimulus coefficient and engagement index is very similar (Fig. 7c). This seems hard to reconcile with the very different slopes of the psychometric curves. To understand why it happens, in Fig. 9 we plot the performance of our Bayesian model, measured as fraction correct, versus the assumed prior, $\hat{p}_r$, for various levels of the noise, σ. This figure shows that increasing their assumed prior can have a large effect on performance, especially in the range used by the mice (0.5-0.7; see Fig. 8b). It's this large increase in prior-induced gap with engagement index (which, via Fig. 8b, corresponds to a large increase in $\hat{p}_r$) that leads to the performance gains for large engagement index.

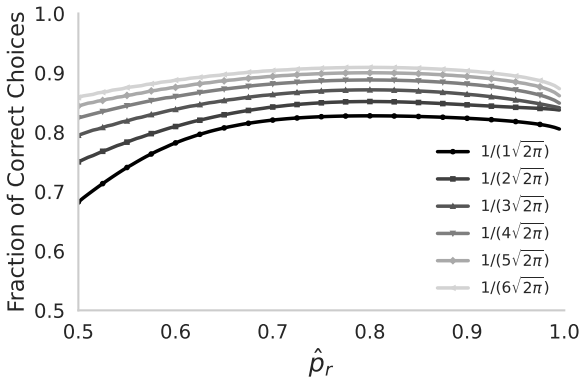

Fig. 9: Fraction correct versus assumed prior, $\hat{p}_r$, given that the true prior is 0.8. The different curves correspond, from bottom to top, of slopes ranging from 1 to 6; alternatively, from noise ranging from to $1/(\sqrt{2\pi}\sigma)$ to $1/(6\sqrt{2\pi}\sigma)$.

3 Discussion

Decision-making in animals is shaped not only by the immediate stimulus, but also by internal state – motivation, for instance, has a huge effect on performance. But how complex is the internal state? Is it primarily one-dimensional, say degree of motivation?

Or is it a multi-dimensional variable, with different dimensions affecting behavior in different ways?

To address this question, we considered three potential measures of internal state, all within the context of the International Brain Lab [9]: engagement index, which is a measure of the complexity of wheel movements; stimulus coefficient, which is the degree to which mice pay attention to the stimulus as assessed by the GLM-HMM model; and pupil diameter, an autonomic response. For each of them we asked how they affected behavior, as assessed by the slope of the psychometric curve, the prior-induced gap at zero contrast, and performance (Fig. 7).

As can be seen in Fig. 7, pupil diameter (red curves) is a poor predictor of behavior: it has virtually no effect on slope, prior-induced gap, or performance. This suggests that it is only weakly correlated with internal state – at odds with previous work showing that pupil diameter is associated with arousal and attention [6–8, 11]. However, there are at least two technical reasons why pupil diameter doesn't reflect internal state in this task. First, the inter-trial interval is very short ($< 2s$), so the visual stimulus and reward from the previous trial may influence pupil diameter. Second, pupil diameter was estimated from a camera capturing not just the eye, but also the head and hands, making extraction challenging. Thus, our results don't in general rule out pupil diameter as a measure of internal state.

The stimulus coefficient and engagement index, on the other hand, are strongly correlated with behavior. But they are correlated in very different ways, both for the slope and prior induced-gap. In particular, the slope of the psychometric curve increases approximately linearly with stimulus coefficient over the whole range of z-scores, whereas the slope increases with engagement index only for negative z-scores, while for positive z-scores (about half the data), it decreases (Fig. 7a). And the prior-induced gap exhibits first a small increase with stimulus coefficient, followed by a slightly larger decrease, but with overall very little change, whereas it increases monotonically, and relatively rapidly, with engagement index (Fig. 7b).

Given Fig. 8, which shows that the prior-induced gap is a reasonable proxy for the prior assumed by the mice, $\hat{p}_r$, and the slope is proportional to the inverse of the observation observation noise, σ, the following picture emerges. When the stimulus coefficient is high, the mice are paying close attention to the stimulus, in a way that reduces their internal noise. They are, though, a bit overconfident in their ability: their performance would have been better if they had relied more on past trials (by increasing the prior probability that the grating is on the right) relative to their sensory evidence. When the engagement index is high, on the other hand, the animals pay less attention to the stimulus, in the sense that their observation noise is higher, but to make up for it they rely more on the prior. Somewhat surprisingly, the different strategies yield nearly identical performance (Fig. 7c).

Taken together, this suggests that the internal state is at least two-dimensional. One dimension determines how much the animals rely on the immediate stimulus, and the other determines how much they take into account the prior. There is, though, a large difference in the stability of these dimensions: the degree to which the animals rely on the immediate stimulus changes slowly from one trial to the next; the amount they take into account the prior changes rapidly (Fig. 5d). Because the prior must be

computed from past trials, this makes the prediction that information about the prior is always present, but the animals sometimes pay attention to it and sometimes don't. This gives animals the flexibility to "space out" – to pay less attention to the current stimulus and use their knowledge of the world to make decisions. This is a very useful ability, as it allows animals to do more than one thing at a time without much loss in performance.

Methods

M.1 Wheel Data Preprocessing

The IBL dataset consists of sessions with variability in the set of contrasts used for the grating and in the block probabilities. To allow for easy comparison, we restricted our analysis to sessions in which the contrasts were drawn from the set $\pm\{0, 0.0625, 0.125, 0.25, 1.0\}$, and the sessions began with unbiased blocks and then alternated between rightward probabilities of 0.2 and 0.8. We also excluded sessions with:

- fewer than 400 trials;
- missing wheel movement recordings (typically due to hardware malfunction) on more than 10% of the trials;
- error rate on 100% contrast trials exceeding 50%;
- no decision within 60 seconds on more than 10% of the trials.

Finally, only animals contributing at least 13 valid sessions were retained. Given these criteria, the final dataset consisted of 2,415 sessions from 81 mice.

Wheel trajectories in the IBL experiment were recorded at non-uniform time intervals [9]. This irregular sampling poses challenges for modeling with sequential methods such as variational autoencoders. To address this, we linearly interpolated wheel position between measurements and resampled the trajectories at 50 Hz.

M.2 Pupil Data Selection

We obtained pupil data, via the Lightning Pose estimation algorithm [13], from a subset of the sessions and mice used for the VAE and GLM-HMM: 145 sessions from 44 mice. Pupil diameter measurements derived from the Lightning Pose estimation algorithm can be compromised by poor illumination, partial occlusion, and periods when the animal closes its eyes or turns away from the camera. To ensure that only high-quality, interpretable pupil data was included, we applied a session-level selection criterion based on stimulus-driven modulation. On each trial we took the pupil diameter to be the average in a window ranging from 1-2 s after stimulus onset. We then fit a linear regression model relating trial-by-trial fluctuations in mean pupil size to the absolute value of stimulus contrast, following the approach described in [17]. Only sessions exhibiting a significant correlation coefficient ($p < 0.05$, ordinary least squares) were retained for subsequent analyses. That left us with 93 sessions from 35 mice.

M.3 VAE

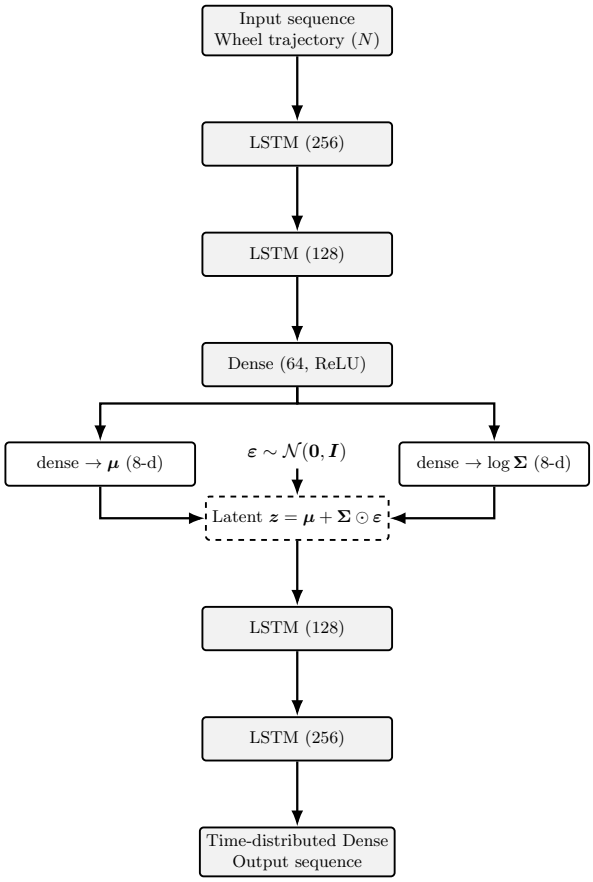

Fig. M.1: VAE architecture.

M.3.1 Architecture

The architecture of the VAE is shown in Fig. M.1. The encoder (top) consists of two stacked Long Short-Term Memory (LSTM) layers to capture the temporal structure of each trajectory. The first LSTM, with 256 units, takes the wheel position as input. Because wheel trajectories in the IBL dataset are not all the same length, missing items are padded with -2 – a value chosen to be outside the range of actual wheel position, which are constrained to lie between -0.3 and $+0.3$ radians. The output of the first LSTM is a sequence of hidden states, preserving information at each time point. The second LSTM, with 128 units, compresses the sequence into a single fixed-size vector summarizing the trajectory. This vector is further processed by a dense

layer with 64 units and ReLU activation. Two parallel dense layers then compute the mean and diagonal covariance of an 8-dimensional latent Gaussian distribution, which serves as the representation of the trajectory.

The decoder mirrors the encoder. A sample, $\mathbf{z}$, drawn from the latent space is turned into a constant 8-dimensional time series (i.e., $\mathbf{z}(t) = \mathbf{z}$ for all time points, t). This is passed through two LSTM layers (128 units followed by 256 units). A time-distributed dense layer generates the predicted wheel position at each time step, producing the reconstructed trajectory.

Training minimizes the loss given in Eq. (1). To allow gradients to flow through the stochastic sampling process, a sample from the corresponding Gaussian distribution is passed to the decoder using the reparameterization trick [18], which is then used to reconstruct the original input.

M.3.2 Evidence Lower Bound

Here we show that the loss given in Eq. (1) corresponds to the Evidence Lower Bound (ELBO) [10, 19]. For the observed wheel trajectory θ_n on trial n, the VAE defines a generative model

$$p_\phi(\theta_n, z_n) = p_\phi(\theta_n \mid z_n), p(z_n), \tag{6}$$

where $p(z_n) = \mathcal{N}(0, I)$ is the prior over latents, and $p_\phi(\theta_n \mid z_n)$ is the distribution over trajectories generated by the decoder. The log-likelihood of the data is

$$\log p_\phi(\theta_n) = \log \int p_\phi(\theta_n \mid z_n)\, p(z_n)\, dz_n\,. \tag{7}$$

This is generally intractable, so we introduce an approximate posterior,

$$q_\phi(z_n \mid \theta_n) = \mathcal{N}(\boldsymbol{\mu}_n, \boldsymbol{\Sigma}_n)\,. \tag{8}$$

This leads to the Evidence Lower Bound (ELBO) [10, 19],

$$\log p_\phi(\theta_n) \geq \mathbb{E}_{z_n \sim q_\phi(z_n|\theta_n)}\big[\log p_\phi(\theta_n \mid z_n)\big] - D_{\mathrm{KL}}\big(q_\phi(z_n \mid \theta_n) \,\|\, p(z_n)\big)\,. \tag{9}$$

Choosing

$$\log p_\phi(\theta_n \mid z_n) = -\frac{1}{\lambda}\int_0^{T_n} dt\, \big(\theta_n(t) - \hat{\theta}_n(t)\big)^2 + \text{ constant}\,, \tag{10}$$

we see that the first term in Eq. (9) is proportional to the negative of the reconstruction error given in Eq. (1). Taking into account the second term, Eq. (9) is proportional to the negative of our VAE loss, Eq. (1). Thus, maximizing the ELBO is equivalent to minimizing our VAE loss. As mentioned in the main text, for variable-length trajectories, the reconstruction term is computed only over the actual (non-padded) time indices, ignoring any padding.

M.4 Measures of engagement for the GLM-HMM

The GLM-HMM algorithm produces, on trial n, $p_i(n)$ – the probability that the animal is in state i [2, 3]. Two measures can be derived from these probabilities. The one we use in the main text is the stimulus coefficient, $s(n)$, which is a weighted average of w_{i1} (Eq. (4)). An alternative measure is the probability of being in the engaged state. Qualitatively, an engaged state is one for which w_{i1} is large. For the 3-state model there is typically only one large weight, so we set the probability of being engaged, denoted p_{engaged}, to the probability associated with the largest weight,

$$\begin{aligned} p^{\text{3-state}}_{\text{engaged}}(n) &= p_{i^*}(n) \\ i^* &= \underset{i}{\text{argmax}}\; w_{i1}(n)\,. \end{aligned} \tag{11}$$

For the 4-state model, however, there are often two large weights, corresponding to two engaged states. To take this into account, we instead use

$$p^{\text{4-state}}_{\text{engaged}}(n) = \sum_{w_{i1} \geq \alpha w_{\max}} p_i(n) \tag{12}$$

where the sum is over state, i, and

$$w_{\max} \equiv \max_i w_{i1}\,. \tag{13}$$

In our analysis we used $\alpha = 0.75$.

M.5 Association between engagement measures

In Fig. 5, we show the pairwise relationships between engagement index, GLM-HMM stimulus coefficient and pupil diameter using kernel density estimation. These were constructed as follows. Using z_n to denote z-score on trial n, the density is given by

$$f(x,y) = \frac{1}{Z}\sum_n e^{-\frac{1}{2h^2}((x-z^k_n)^2+(y-z^l_n)^2)} \tag{14}$$

where k and l refer to engagement index, GLM-HMM stimulus coefficient, or pupil diameter; h is the kernel bandwidth (determined automatically by the Python library); and Z is chosen so that the maximum value of $f(x,y)$ is 1.

To probe temporal dependencies, we analyzed session-level lag-1 autocorrelations for each measure in Figs. 5d-5f. For a variable X in session i with N trials, we define the lag-1 autocorrelation of variable X in session i as

$$\rho^k_{1i} = \frac{\text{Cov}(X^k_n, X^k_{n-1})}{\text{Var}(X^k_n)}\,. \tag{15}$$

where again k refers to engagement index, GLM-HMM stimulus coefficient, or pupil diameter. The density maps were constructed as in Eq. (14), with z_i^k and z_i^l replaced by ρ_{1i}^k and ρ_{1i}^l. Note that the density maps for the lag-1 autocorrelation is taken over sessions, not trials.

M.6 Prior-Induced Gap, Slope and Performance

In Fig. 7 we plot the slope of the psychometric curve, prior-induced gap and performance versus z-score. The first two quantities depend on the probability of a rightward choice given contrast and z; the last depends on the probability that the animal made a correct choice and z. For the first two,

$$\text{Slope}(z) = \frac{\hat{p}_r(z, c_2, 1) - \hat{p}_r(z, -c_2, 0)}{2c_2} \tag{16a}$$

$$\text{Gap}(z) = \hat{p}_r(z, 0, 1) - \hat{p}_r(z, 0, 0) \tag{16b}$$

where $\hat{p}_r(z, c, b)$ is the estimate of the probability that the animal chooses right given z, c (contrast), and b (prior block; 0 for a left and 1 for a right). Recall that $c_2 = 0.0625$, chosen because it's the smallest nonzero contrast. Note that we are slightly abusing notation, since we use $\hat{p}_r$ for the assumed prior, but $\hat{p}_r$ is distinguishable from $\hat{p}_r(z, c, b)$ because the latter has arguments.

To compute $\hat{p}_r(z, c, b)$, and the probability of making a correct choice given z (see Eq. (21) below), from data, we make use of the fact that on every trial our data consists ot the z-score (of engagement index, stimulus coefficient, and pupil diameter), contrast, block, the direction the animal chose, and whether or not it was correct. We'll denote this as a 5-tuple, $(z_n, c_n, b_n, d_n, r_n)$, where n denotes trial, z_n is the z-score of the variable of interest, d_n is the direction the animal chose (1 for right, 0 for left), and r_n is the reward the animal receives (1 for correct, 0 for incorrect). To get a smooth estimates using this 5-tuple, we'll use kernel density estimation. For the probability of a rightward choice given contrast and block, the kernel density estimator is

$$\hat{p}_r(z, c, b) = \frac{1}{Z(z, c, b)} \sum_{n\,:\,c_n = c, b_n = b} f(z - z_n) d_n \tag{17}$$

where $Z(z, c, b)$ is the contrast-dependent normalization,

$$Z(z, c, b) = \sum_{n\,:\,c_n = c, b_n = b} f(z - z_n)\,, \tag{18}$$

and we take $f(\cdot)$ to be Gaussian,

$$f(z) = \frac{e^{-\frac{z^2}{2h^2}}}{\sqrt{2\pi h^2}}\,. \tag{19}$$

To determine the error bars, we treat d_n as Bernoulli random variable with mean $\hat{p}_r(z_n, c_n, b_n)$. Thus,

$$\mathrm{Var}[\hat{p}_r(z,c,b)] = \sum_{n\,:\,c_n=c, b_n=b} \left(\frac{f(z-z_n)}{Z(z,c,b)}\right)^2 \hat{p}_r(z_n,c_n,b_n)\big(1-\hat{p}_r(z_n,c_n,b_n)\big)\,. \quad (20)$$

To compute the mean slope, Eq. (16a), and prior-induced gap, Eq. (16b), we simply insert Eq. (17) into these equations. Equation (20) can then be used to compute error bars. In practice, we estimate the variance using the kernel-weighted empirical variance of the Bernoulli samples (i.e., a plug-in estimator), which is equivalent to Eq. (20) up to finite-sample corrections.

The kernel density estimator for the probability of making a correct choice given z is similar to Eq. (17), except with d_n replaced by r_n and the summation is now over trials,

$$p_{\mathrm{correct}}(z) = \frac{1}{Z(z)} \sum_n f(z-z_n) r_n \quad (21)$$

where $Z(z)$ (distinguishable from $Z(z,c,b)$ by its arguments) is the normalization,

$$Z(z) = \sum_n f(z-z_n) \quad (22)$$

and again $f(z)$ is given by Eq. (19). As above we treat r_n as a Bernoulli random variable with mean $p_{\mathrm{correct}}(z_n)$, so the variance of $p_{\mathrm{correct}}(z_n)$ is given by

$$\mathrm{Var}[p_{\mathrm{correct}}(z)] = \sum_n \left(\frac{f(z-z_n)}{Z(z)}\right)^2 p_{\mathrm{correct}}(z_n)\big(1-p_{\mathrm{correct}}(z_n)\big)\,. \quad (23)$$

In our analysis we use $h = 0.4$.

M.7 Bayesian model

To gain further insight into the internal variables, we build a parametric Bayesian model of behavior, and then use behavioral data to infer the parameters. As discussed in Sec. 2.3, we assume that the mice observe a variable ξ that is probabilistically related to the true contrast, c, via Eq. (5). The probability that the right side was rewarded given ξ is

$$P(\text{right-rewarded}|\xi) = \int_0^\infty dc'\, \frac{P(\xi|c')\widehat{P}_0(c')}{P(\xi)} \quad (24)$$

where $\widehat{P}_0(c)$ is the prior assumed by the mice when making decisions. Note that we have equated reward with positive contrast. Because contrast can be zero this is not strictly true; we take care of this shortly with a small adjustment to the prior.

The likelihood is determined solely by the internal noise, σ, so we just need to parameterize the prior, $\widehat{P}_0(c)$. For that we start with the true prior, $P_0(c)$, which is given by

$$P_0(c) = \frac{p_r}{N_c}\sum_{i=1}^{N_c}\delta(c-c_i) + \frac{p_l}{N_c}\sum_{i=1}^{N_c}\delta(c+c_i) \tag{25}$$

where p_r and p_l are the true probabilities that the grating appears on the right and left side, respectively, and N_c is the number of contrasts. The set of contrasts, c_i, correspond to the positive contrasts used in the experiments: $\{c_1, ..., c_{N_c}\} = \{0, 0.0625, 0.125, 0.25, 1.0\}$. Since the mice see a large number of trials, we'll assume they know the c_i with high accuracy (approximated by delta-functions). However, the prior, p_r, has to be inferred on a trial-by-trial basis, and so carries a great deal of uncertainty. Thus, for the prior assumed by the mice, we'll replace p_r and p_l in Eq. (25) with $\hat{p}_r$ and $\hat{p}_l(\equiv 1-\hat{p}_r)$, the prior assumed by the mice, giving us

$$\widehat{P}_0(c) = \frac{\hat{p}_r}{N_c}\sum_{i=1}^{N_c}\delta(c-c_i) + \frac{\hat{p}_l}{N_c}\sum_{i=1}^{N_c}\delta(c+c_i)\,. \tag{26}$$

Our model thus consists of two parameters: internal noise, σ, and the rightward prior assumed by the mice, $\hat{p}_r$. We'll infer these from the average slope of the psychometric curve and the gap between the curves associated with the rightward and leftward block, both evaluated at zero contrast (see Eqs. (33) and (34) below). As just mentioned, there is a small, but important, technical issue: the probability of reward at zero contrast is not well defined. Here we follow the experiments, where the reward probability matches the prior probability on zero contrast trials: it's 0.8 for choosing right and 0.2 for choosing left when in rightward biased blocks, and the opposite, 0.2 and 0.8, when in leftward biased blocks. To implement this in our analysis, we let c_1 take on an infinitesimally small positive value. Thus, when we integrate contrast with a lower bound of 0 in Eq. (24), we pick up only the rightward probabilities. Consequently, inserting Eq. (26) into Eq. (24), we arrive at

$$P(\text{right-rewarded}|\xi) = \frac{\hat{p}_r f(\xi,\sigma)}{\hat{p}_r f(\xi,\sigma) + \hat{p}_l f(-\xi,\sigma)} \tag{27}$$

where

$$f(\xi,\sigma) \equiv \frac{1}{N_c}\sum_{i=1}^{N_c} e^{-(\xi-c_i)^2/2\sigma^2}\,. \tag{28}$$

To construct psychometric curves, we need to know the probability of choosing right as a function of contrast. We'll assume that the mice choose the side that's most likely to be rewarded; that is, they choose if $P(\text{right-rewarded}|\xi) > 1/2$ and left

otherwise. Thus, using Eq. (27), we have

$$P(\text{choose-right}|c) = \int_{-\infty}^{\infty} d\xi\, P(\xi|c)\Theta\big(\hat{p}_r\, f(\xi,\sigma) - \hat{p}_l f(-\xi,\sigma)\big) \tag{29}$$

where Θ is the Heaviside step function. Letting ξ^* satisfy

$$\hat{p}_r\, f(\xi^*,\sigma) - \hat{p}_l f(-\xi^*,\sigma) = 0\,, \tag{30}$$

we have

$$P(\text{choose-right}|c) = \Phi\left(\frac{c - \xi^*(\hat{p}_r,\sigma)}{\sigma}\right) \tag{31}$$

where $\Phi(\cdot)$ is the cumulative normal function,

$$\Phi(z) \equiv \int_{-\infty}^{z} dy\, \frac{e^{-y^2/2}}{\sqrt{2\pi}}\,. \tag{32}$$

With this model, the slope of the psychometric curve evaluated at $c = 0$ and $\hat{p}_r = 1/2$ (the latter implying, by symmetry, that $\xi^* = 0$) is

$$\text{slope} = \left.\frac{dP(\text{choose-right}|c)}{dc}\right|_{c=\xi^*=0} = \frac{1}{\sqrt{2\pi}\sigma}\,. \tag{33}$$

The gap at zero contrast is slightly more complicated; it's given by

$$\begin{aligned}\text{gap} &= \Phi\left(\frac{-\xi^*(\hat{p}_r,\sigma)}{\sigma}\right) - \Phi\left(\frac{-\xi^*(\hat{p}_l,\sigma)}{\sigma}\right) \\ &= 2\Phi\left(\frac{-\xi^*(\hat{p}_r,\sigma)}{\sigma}\right) - 1\end{aligned} \tag{34}$$

where the second line follows from symmetry. A plot of the gap versus $\hat{p}_r$ is shown in Fig. M.2.

Note that this expression is easily inverted. First using the second line of Eq. (34), we have

$$\xi^* = -\sigma\Phi^{-1}\left(\frac{\text{gap}+1}{2}\right)\,. \tag{35}$$

Second, we can use Eq. (30) to express $\hat{p}_r$ in terms of ξ^*,

$$\hat{p}_r = \frac{f(-\xi^*,\sigma)}{f(\xi^*,\sigma) + f(-\xi^*,\sigma)}\,. \tag{36}$$

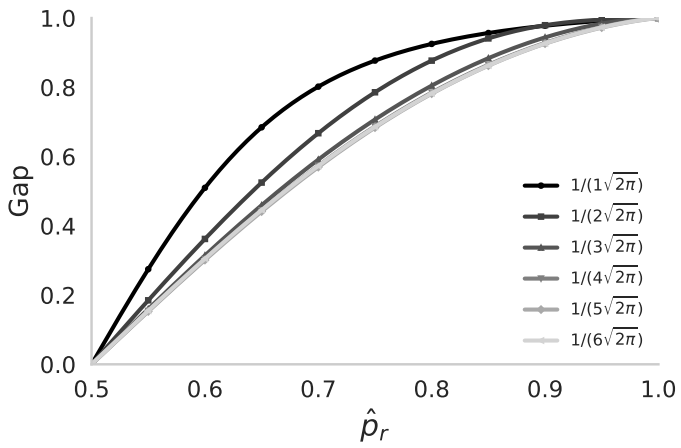

Fig. M.2: Gap (right hand side of Eq. (34)) versus $\hat{p}_r$, for $\sigma = 1/(n\sqrt{2\pi})$, $n = 1, ..., 6$ (top to bottom), which corresponds to slopes $1, ..., 6$. Only four lines are visible because the curves for $n = 4, 5$ and 6 lie on top of each other.

Finally, performance, measured as the probability of making the correct choice, is equal to the probability of choosing right given that the right side was rewarded plus the probability of choosing left given the left side was rewarded. That's given by

$$P_{\text{correct}} = \frac{p_r}{N_c} \sum_{i=1}^{N_c} \Phi\left(\frac{c_i - \xi^*(\hat{p}_r, \sigma)}{\sigma}\right) + \frac{p_l}{N_c} \sum_{i=1}^{N_c} \Phi\left(\frac{-c_i - \xi^*(\hat{p}_l, \sigma)}{\sigma}\right). \tag{37}$$

This expression is plotted in Fig. 9.

References

[1] Prins, N.: The psychometric function: The lapse rate revisited. Journal of Vision **12**(6), 25–25 (2012)

[2] Ashwood, Z.C., Roy, N.A., Stone, I.R., Laboratory, I.B., Urai, A.E., Churchland, A.K., Pouget, A., Pillow, J.W.: Mice alternate between discrete strategies during perceptual decision-making. Nature Neuroscience **25**(2), 201–212 (2022)

[3] Mohammadi, Z., Ashwood, Z.C., Laboratory, T.I.B., Pillow, J.W.: Identifying the factors governing internal state switches during nonstationary sensory decision-making. bioRxiv (2024)

[4] Cuturela, L.I., Laboratory, I.B., Pillow, J.W.: Internal states emerge early during learning of a perceptual decision-making task. bioRxiv, 2024–11 (2024)

[5] Niyogi, R.K., Shizgal, P., Dayan, P.: Some work and some play: microscopic and macroscopic approaches to labor and leisure. PLoS computational biology **10**(12), 1003894 (2014)

[6] Suzuki, T.W., Kunimatsu, J., Tanaka, M.: Correlation between pupil size and subjective passage of time in non-human primates. Journal of Neuroscience **36**(44), 11331–11337 (2016)

[7] Gilzenrat, M.S., Nieuwenhuis, S., Jepma, M., Cohen, J.D.: Pupil diameter tracks changes in control state predicted by the adaptive gain theory of locus coeruleus function. Cognitive, Affective, & Behavioral Neuroscience **10**(2), 252–269 (2010)

[8] Aston-Jones, G., Cohen, J.D.: Adaptive gain and the role of the locus coeruleus–norepinephrine system in optimal performance. Journal of Comparative Neurology **493**(1), 99–110 (2005)

[9] International Brain Laboratory, Aguillon-Rodriguez, V., Angelaki, D., Bayer, H., Bonacchi, N., Carandini, M., Cazettes, F., Chapuis, G., Churchland, A.K., Dan, Y., *et al.*: Standardized and reproducible measurement of decision-making in mice. Elife **10**, 63711 (2021)

[10] Kingma, D.P.: Auto-encoding variational bayes. arXiv preprint arXiv:1312.6114 (2013)

[11] Schwartz, Z.P., Buran, B.N., David, S.V.: Pupil-associated states modulate excitability but not stimulus selectivity in primary auditory cortex. Journal of neurophysiology **123**(1), 191–208 (2020)

[12] Kempen, J., Loughnane, G.M., Newman, D.P., Kelly, S.P., Thiele, A., O'Connell, R.G., Bellgrove, M.A.: Behavioural and neural signatures of perceptual decision-making are modulated by pupil-linked arousal. Elife **8**, 42541 (2019)

[13] Biderman, D., Whiteway, M.R., Hurwitz, C., Greenspan, N., Lee, R.S., Vishnubhotla, A., Warren, R., Pedraja, F., Noone, D., Schartner, M.M., *et al.*: Lightning pose: improved animal pose estimation via semi-supervised learning, bayesian ensembling and cloud-native open-source tools. Nature Methods **21**(7), 1316–1328 (2024)

[14] McGinley, M.J., Vinck, M., Reimer, J., Batista-Brito, R., Zagha, E., Cadwell, C.R., Tolias, A.S., Cardin, J.A., McCormick, D.A.: Waking state: rapid variations modulate neural and behavioral responses. Neuron **87**(6), 1143–1161 (2015)

[15] Reimer, J., Froudarakis, E., Cadwell, C.R., Yatsenko, D., Denfield, G.H., Tolias, A.S.: Pupil fluctuations track fast switching of cortical states during quiet wakefulness. neuron **84**(2), 355–362 (2014)

[16] 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**(3), 740–754 (2015)

[17] Johnson, P.A., Nieuwenhuis, S., Urai, A.: Pupil dynamics preceding switches in task engagement. In: Proceedings of the Annual Meeting of the Cognitive Science Society, vol. 46 (2024)

[18] Doersch, C.: Tutorial on variational autoencoders. arXiv preprint arXiv:1606.05908 (2016)

[19] Rezende, D.J., Mohamed, S., Wierstra, D.: Stochastic backpropagation and approximate inference in deep generative models. In: Proceedings of the 31st International Conference on Machine Learning (ICML), pp. 1278–1286 (2014). https://arxiv.org/abs/1401.4082

Supporting Information

In the main text we focused on the 3-state GLM-HMM model, and our measure of internal state was the stimulus coefficient (see Eq. (4)). Here we provide results for the 4-state model, and for the probability of being in the engaged state.

Figure S.1 shows the psychometric curves versus stimulus coefficient for the 3- and 4-state models (panels a and c) and probability engaged for the 3- and 4-state models (panels b and d). Panel a, which is included for comparison, is equivalent to Fig. 6b of the main text. The psychometric curves are nearly identical in all four panels, suggesting that our results are robust to the number of states and choice of stimulus coefficient versus probability engaged.

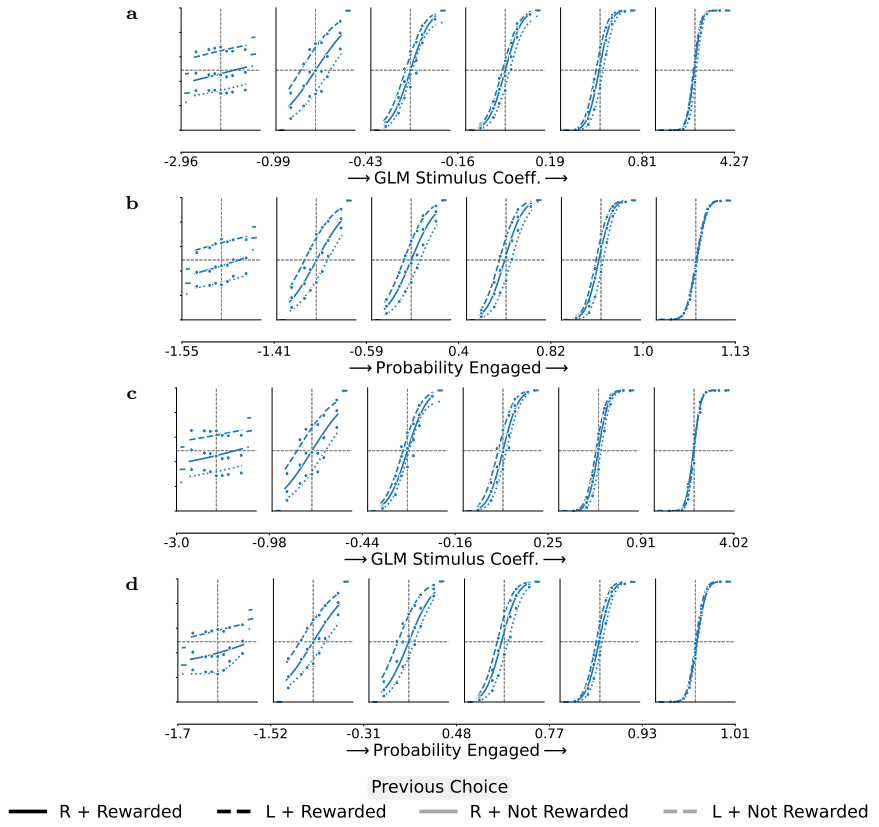

Fig. S.1: Alternative psychometric curves for the GLM-HMM model. As in the main text, the x-axis is z-score. **a**. 3-state model versus stimulus coefficient, for comparison (same as Fig. 6b). **b**. 3-state model versus probability engaged. **c**. 4-state model versus stimulus coefficient. **d**. 4-state model versus probability engaged.

In Fig. S.2 we plot the slope, prior-induced gap and fraction correct versus z-scored stimulus coefficient and probability engaged. The solid blue lines, included for comparison, are equivalent to the solid blue lines in Fig. 7; the other lines are either for the 4-state model or plots versus z-scored probability engaged. All curves are similar; the main difference being a smaller range of z for probability engaged, which is likely due to the fact that probability must be between 0 and 1. This further confirms that our results are robust.

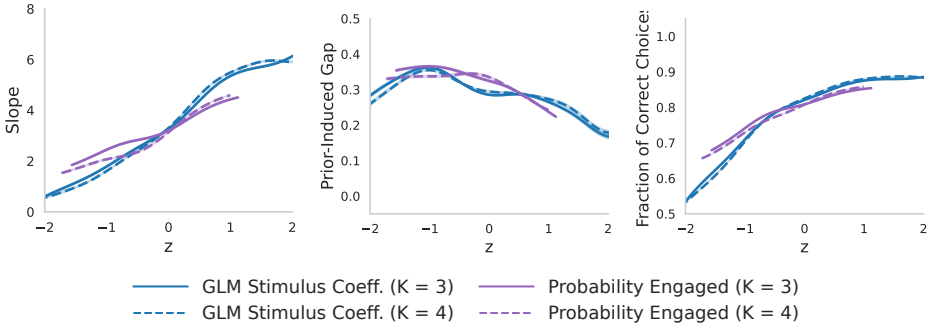

Fig. S.2: Behavioral metrics versus z-scored measures of internal state in the GLM-HMM algorithm. The solid line in each graph is the 3-state GLM-HMM coefficient we plotted in Fig. 7. K is the number of states used in the GLM-HMM.