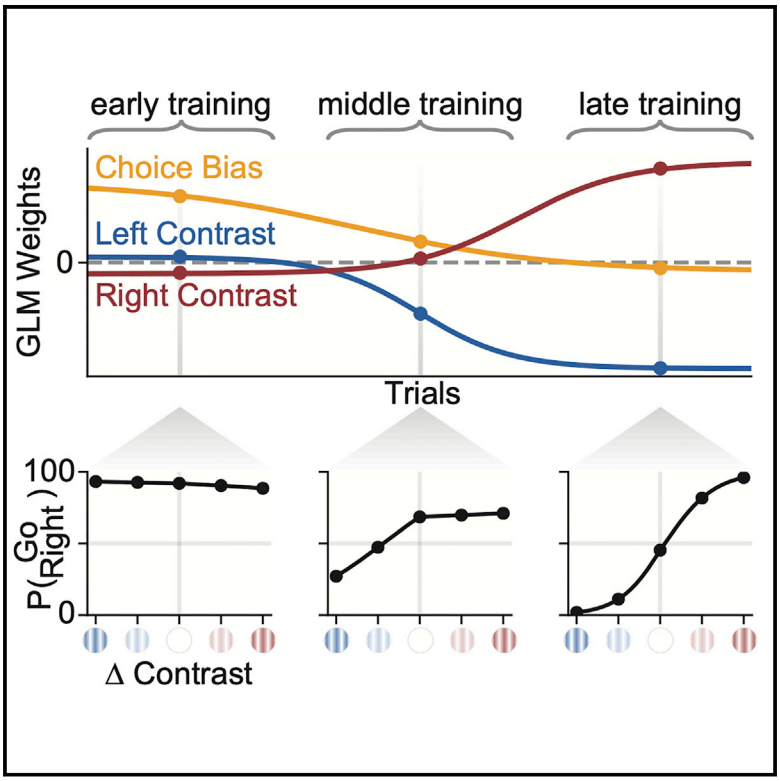
Neuron

NeuroResource

Extracting the dynamics of behavior in sensory decision-making experiments

Graphical Abstract



Authors

Nicholas A. Roy, Ji Hyun Bak, The International Brain Laboratory, Athena Akrami, Carlos D. Brody, Jonathan W. Pillow

Correspondence

nicholas.roy.42@gmail.com (N.A.R.), pillow@princeton.edu (J.W.P.)

In Brief

Roy et al. present a method for inferring the time course of behavioral strategies in sensory decision-making tasks, which they use to analyze how behavior evolves during training in rats, mice, and humans.

Highlights

- Dynamic model for time-varying sensory decision-making behavior
- Visualize changes in behavioral strategies of mice, rats, and humans across training
- Infer how quickly different parameters change between trials and between sessions
- Colab notebook reproduces all figures and analyses, facilitating application to new data

Roy et al., 2021, Neuron *109*, 1–14
February 17, 2021
https://doi.org/10.1016/j.neuron.2020.12.004

CellPress

Neuron

CellPress

NeuroResource

Extracting the dynamics of behavior in sensory decision-making experiments

Nicholas A. Roy,[1,7,9,*] Ji Hyun Bak,[2,3,8] The International Brain Laboratory, Athena Akrami,[1,4] Carlos D. Brody,[1,5] and Jonathan W. Pillow[1,6,*]

[1]Princeton Neuroscience Institute, Princeton University, Princeton, NJ 08544, USA
[2]Korea Institute for Advanced Study, Seoul 02455, South Korea
[3]Redwood Center for Theoretical Neuroscience, University of California, Berkeley, Berkeley, CA 94720, USA
[4]Sainsbury Wellcome Centre, University College London, London W1T 4JG, UK
[5]Howard Hughes Medical Institute, Princeton University, Princeton, NJ 08544, USA
[6]Department of Psychology, Princeton University, Princeton, NJ 08544, USA
[7]Present address: DeepMind, London N1C 4AG, UK
[8]Present address: University of California, San Francisco, San Francisco, CA 94158, USA
[9]Lead contact
*Correspondence: nicholas.roy.42@gmail.com (N.A.R.), pillow@princeton.edu (J.W.P.)
https://doi.org/10.1016/j.neuron.2020.12.004

Summary

Decision-making strategies evolve during training and can continue to vary even in well-trained animals. However, studies of sensory decision-making tend to characterize behavior in terms of a fixed psychometric function that is fit only after training is complete. Here, we present PsyTrack, a flexible method for inferring the trajectory of sensory decision-making strategies from choice data. We apply PsyTrack to training data from mice, rats, and human subjects learning to perform auditory and visual decision-making tasks. We show that it successfully captures trial-to-trial fluctuations in the weighting of sensory stimuli, bias, and task-irrelevant covariates such as choice and stimulus history. This analysis reveals dramatic differences in learning across mice and rapid adaptation to changes in task statistics. PsyTrack scales easily to large datasets and offers a powerful tool for quantifying time-varying behavior in a wide variety of animals and tasks.

Introduction

The behavior of well-trained animals in carefully designed tasks is a pillar of modern neuroscience research (Carandini, 2012; Krakauer et al., 2017; Niv, 2020). In sensory decision-making experiments, animals must learn to integrate relevant sensory signals while ignoring a large number of task-irrelevant covariates (Gold and Shadlen, 2007; Brunton et al., 2013; Hanks and Summerfield, 2017). However, the sensory decision-making literature has tended to focus on characterizing the decision-making behavior of fully trained animals in terms of fixed strategies, as in signal detection theory (Green and Swets, 1966) or the drift-diffusion model (Ratcliff and Rouder, 1998). This approach neglects the dynamics of decision-making behavior across trials, which may be essential for understanding learning, exploration, adaptation to task statistics, and other forms of non-stationary behavior (Usher et al., 2013; Pisupati et al., 2019; Brunton et al., 2013; Piet et al., 2018).

Characterizing the dynamics of sensory decision-making behavior is challenging due to the fact that decisions may depend on a large number of task covariates, including the sensory stimuli, an animal's choice bias, past stimuli, past choices, and past rewards. Detecting and disentangling the influence of these variables on a single choice is an ill-posed problem due to the fact that we have many unknowns (the weights on each variable) and a single observation (the animal's choice). As a result, it is common to assume that the decision-making rule, or strategy, of an animal is fixed over some reasonably large number of trials. However, this assumption is at odds with the fact that decision-making strategies may change on a trial-to-trial basis and may evolve rapidly during training, when animals are learning a new task (Carandini and Churchland, 2013).

Understanding what drives changes in decision-making behavior has long been the domain of reinforcement learning (RL) (Sutton and Barto, 2018; Sutton, 1988). In this paradigm, behavioral dynamics are examined through the lens of the rewards and punishments that may accompany each decision. RL-based approaches are generally normative, meaning that they describe changes in behavior as resulting from the optimization of some measure of future reward (Niv, 2009; Daw 2011; Niv et al., 2015; Samejima et al., 2004; Daw and Courville, 2008; Ashwood et al., 2020). By contrast, descriptive modeling approaches seek only to infer time-varying changes in strategy from the observed choices of an animal, without attributing such changes to any notion of optimality. Previous studies in this tradition include those of Smith et al. (2004) and Suzuki and Brown (2005), which focused on identifying the time at which an untrained animal began to learn. Other work from Kattner

Please cite this article in press as: Roy et al., Extracting the dynamics of behavior in sensory decision-making experiments, Neuron (2020), https://doi.org/10.1016/j.neuron.2020.12.004

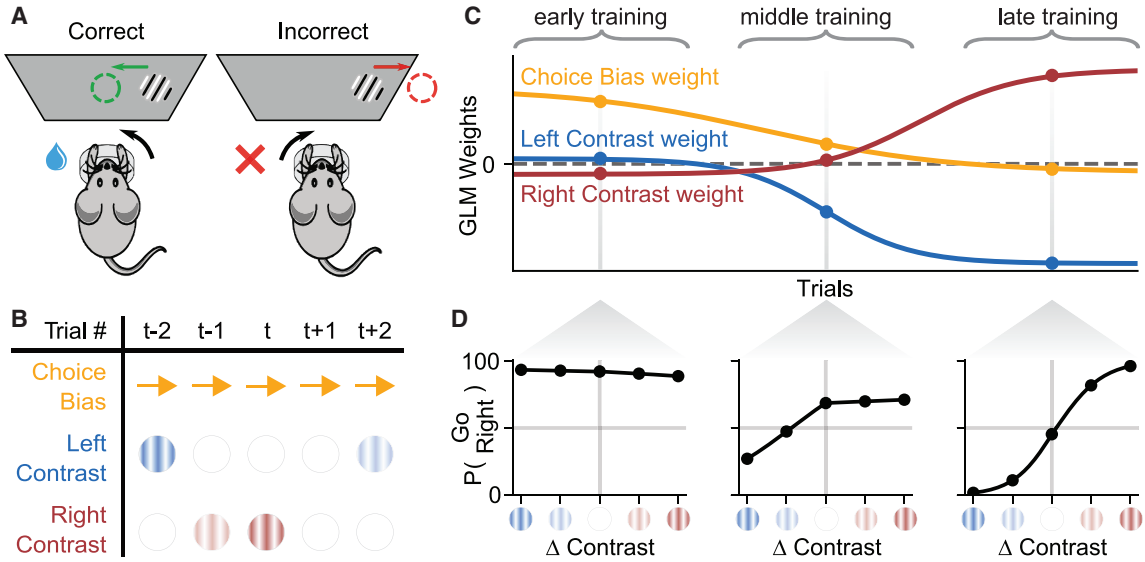


Figure 1. Schematic of binary decision-making task and dynamic psychophysical model

(A) A schematic of the IBL sensory decision-making task. On each trial, a sinusoidal grating (with contrast values between 0% and 100%) appears on either the left or right side of the screen. Mice must report the side of the grating by turning a wheel (left or right) to receive a water reward (see STAR methods for details) (International Brain Laboratory et al., 2020).

(B) An example table of the task variables $\mathbf{x}_t$ assumed to govern behavior for trials $t-2$ to $t+2$, consisting here of a choice bias (a constant rightward bias, encoded as "+1" on each trial), the contrast value of the left grating, and the contrast value of the right grating.

(C) Hypothetical time course of the psychophysical weights $W = [\mathbf{w}_1, \ldots, \mathbf{w}_T]$, which evolve smoothly over the course of training. Each weight corresponds to one of the $K = 3$ components of $\mathbf{x}_t$, such that the weight value at trial t indicates how the corresponding variable affects the animal's choice on that trial.

(D) Psychometric curves induced by the psychophysical weights $\mathbf{w}_t$ on particular trials in "early," "middle," and "late" training periods, as defined in (C). Early behavior is highly biased and insensitive to stimuli. Over the course of training, behavior evolves toward unbiased, high-accuracy performance consistent with a steep psychometric function.

et al. (2017) extended the standard psychometric curve to allow its parameters to vary continuously across trials.

Here, we present PsyTrack, a descriptive modeling approach for inferring the trajectory of an animal's decision-making strategy across trials, building on ideas developed by Bak et al. (2016) and Roy et al. (2018a). Our model describes decision-making behavior at the resolution of single trials, allowing for visualization and analysis of psychophysical weight trajectories both during and after training. It contains interpretable hyperparameters governing the rates of change of different weights, allowing us to quantify how rapidly different weights evolve between trials and between sessions. We apply PsyTrack to behavioral data collected during training in two different experiments (auditory and visual decision making) and three different species (mouse, rat, and human). After validating the method on simulated data, we use it to analyze an example mouse that learns to track block structure in a non-stationary visual decision-making task (International Brain Laboratory et al., 2020). We then examine how trial history influences rat (but not human) decisions during early training on an auditory parametric working memory task (Akrami et al., 2018). To facilitate application to new datasets, we provide a publicly available software implementation in Python, along with a Google Colab notebook that precisely reproduces all of the figures in this article directly from publicly available raw data (see STAR methods).

Results

Our primary contribution is a method for characterizing the evolution of animal decision-making behavior on a trial-to-trial basis. Our approach consists of a dynamic Bernoulli generalized linear model (GLM), defined by a set of smoothly evolving psychophysical weights. These weights characterize the decision-making strategy of the animal at each trial in terms of a linear combination of available task variables. The larger the magnitude of a particular weight, the more the decision of the animal relies on the corresponding task variable. Learning to perform a new task therefore involves driving the weights on "relevant" variables (e.g., sensory stimuli) to large values, while driving weights on irrelevant variables (e.g., bias, choice history) to zero. However, classical modeling approaches assume that weights remain constant over long blocks of trials, which precludes the tracking of trial-to-trial behavioral changes that arise during learning and in non-stationary environments. Below, we describe our modeling approach in more detail.

Dynamic psychophysical model for decision-making tasks

Although PsyTrack is applicable to any binary decision-making task, for concreteness, we introduce our method in the context of the task used by the International Brain Laboratory (IBL) (illustrated in Figure 1A; International Brain Laboratory et al., 2020). In

Please cite this article in press as: Roy et al., Extracting the dynamics of behavior in sensory decision-making experiments, Neuron (2020), https://doi.org/10.1016/j.neuron.2020.12.004

this visual detection task, a mouse is positioned in front of a screen and a wheel. On each trial, a sinusoidal grating (with contrast values between 0% and 100%) appears on either the left or right side of the screen. The mouse must report the side of the grating by turning the wheel (left or right) to receive a water reward (see STAR methods for more details).

Our modeling approach assumes that on each trial the animal receives an input $\mathbf{x}_t$ and makes a binary decision $y_t \in \{0,1\}$. Here, $\mathbf{x}_t$ is a K-element vector containing the task variables that may affect an animal's decision on trial $t \in \{1,\ldots,T\}$. For the IBL task, $\mathbf{x}_t$ could include the contrast values of left and right gratings, as well as stimulus history, a bias term, and other covariates available to the animal during the current trial (Figure 1B). We model the decision-making process of the animal with a Bernoulli GLM, also known as the logistic regression model. This model characterizes the strategy of the animal on each trial t with a set of K linear weights $\mathbf{w}_t$. The weight vector $\mathbf{w}_t$ describes how the different components of the input vector $\mathbf{x}_t$ affect the choice of the animal on trial t. The probability of a "rightward" decision ($y_t = 1$) is given by

$$p(y_t = 1|\mathbf{x}_t, \mathbf{w}_t) = \phi(\mathbf{x}_t \cdot \mathbf{w}_t), \qquad \text{(Equation 1)}$$

where $\phi(\cdot)$ denotes the logistic function, $\phi(z) = 1/(1+\exp(-z))$. Unlike standard psychophysical models, which assume that weights are constant across time, we assume that the weights evolve gradually over time (Figure 1C). Specifically, we model the weight change after each trial with a Gaussian distribution (Bak et al., 2016; Roy et al., 2018a):

$$\mathbf{w}_{t+1} = \mathbf{w}_t + \boldsymbol{\eta}_t, \quad \eta_{t,k} \sim \mathcal{N}(0, \sigma_k^2), \qquad \text{(Equation 2)}$$

where $\boldsymbol{\eta}_t$ is the vector of weight changes on trial t, and σ_k^2 denotes the variance of the changes in the k^{th} weight. The rate of change of the K different weights in $\mathbf{w}_t$ is thus governed by a vector of smoothness hyperparameters $\theta = \{\sigma_1,\ldots,\sigma_K\}$. A larger σ_k implies larger trial-to-trial changes in the k^{th} weight. Note that if $\sigma_k = 0$ for all k, then the weights are constant, and we obtain the classic psychophysical model with a fixed set of weights for the entire dataset.

Learning to perform a new task can be formalized under this model by a trajectory in weight space. Figures 1C and 1D shows a schematic example of such learning in the context of the IBL task. Here, the behavior of the hypothetical mouse is governed by three weights: a left contrast weight, a right contrast weight, and a (choice) bias weight. The first two weights capture how sensitive the choice of the animal is to left and right gratings, respectively, whereas the bias weight captures an independent, additional bias toward leftward or rightward choices.

In this hypothetical example, the weights evolve over the course of training as the animal learns the task. Initially, during "early training," the left and right contrast weights are close to zero and the bias weight is large and positive, indicating that the animal pays little attention to the left and right contrasts and exhibits a strong rightward choice bias. As training proceeds, the contrast weights diverge from zero and separate, indicating that the animal learns to compute a difference between right and left contrast. By the "late training" period, left and right contrast weights have grown to equal and opposite values, while the bias weight has shrunk to nearly zero, indicating unbiased, high-accurary performance of the task.

Although we have arbitrarily divided the data into three different periods—designated "early," "middle," and "late training"—the three weights change gradually after each trial, providing a fully dynamic description of the decision-making strategy of the animal as it evolves during learning. To better understand this approach, we can compute an "instantaneous psychometric curve" from the weight values at any particular trial (Figure 1D). These curves describe how the mouse converts the visual stimuli to a probability over choice on any trial. Together, the weights in this example illustrate the gradual evolution from a strongly right-biased strategy (Figure 1D, left) toward a high-accuracy strategy (Figure 1D, right). Of course, by incorporating weights on additional task covariates (e.g., choice and reward history), the model can characterize time-varying strategies that are more complex than those captured by a simple psychometric curve.

Inferring weight trajectories from data

The goal of PsyTrack is to infer the full time course of the decision-making strategy of an animal from the observed sequence of inputs $X = [\mathbf{x}_1, \ldots, \mathbf{x}_T]$, and choices $Y = [y_1, \ldots, y_T]$ over the course of an entire experiment. To do so, we estimate the time-varing weights $W = [\mathbf{w}_1,\ldots,\mathbf{w}_T]$ of the animal using the dynamic psychophysical model defined above (Equations 1 and 2), where T is the total number of trials in the dataset. Each of the K rows of W represents the trajectory of a single weight across trials, while each column provides the vector of weights governing decisions on a single trial. The method therefore involves inferring $K \times T$ weights from only T binary decision variables Y.

To estimate W from data, we use a two-step inference procedure called empirical Bayes (Bishop, 2006). First, we estimate θ, the hyperparameters governing the smoothness of the weight trajectories, by maximizing $p(Y|X,\theta)$, known as the evidence, which is the probability of choice data Y given the inputs X and hyperparameters "θ", with W integrated out. Second, we compute the maximum *a posteriori* (MAP) estimate for W given the choice data and the estimated hyperparameters $\hat{\theta}$. Although this optimization problem is computationally demanding, we have developed fast approximate methods that allow us to model datasets with tens of thousands of trials within minutes on a desktop computer (see STAR methods for details; see also Figure S1).

To validate the method, we generated an artificial dataset from a simulated observer with $K = 4$ weights that evolved according to a Gaussian random walk over $T = 5{,}000$ trials (Figure 2A). Each weight had a different standard deviation σ_k (Equation 2), producing weight trajectories with differing average rates of change. We sampled input vectors $\mathbf{x}_t$ for each trial from a standard normal distribution, then sampled the observer's choices y_t according to Equation 1. We then applied PsyTrack to this simulated dataset, which computed estimates of the four hyperparameters $\hat{\theta} = \{\hat{\sigma}_1, \ldots, \hat{\sigma}_4\}$ and weight trajectories $\widehat{W}$ (Figures 2A and 2B).

Please cite this article in press as: Roy et al., Extracting the dynamics of behavior in sensory decision-making experiments, Neuron (2020), https://doi.org/10.1016/j.neuron.2020.12.004

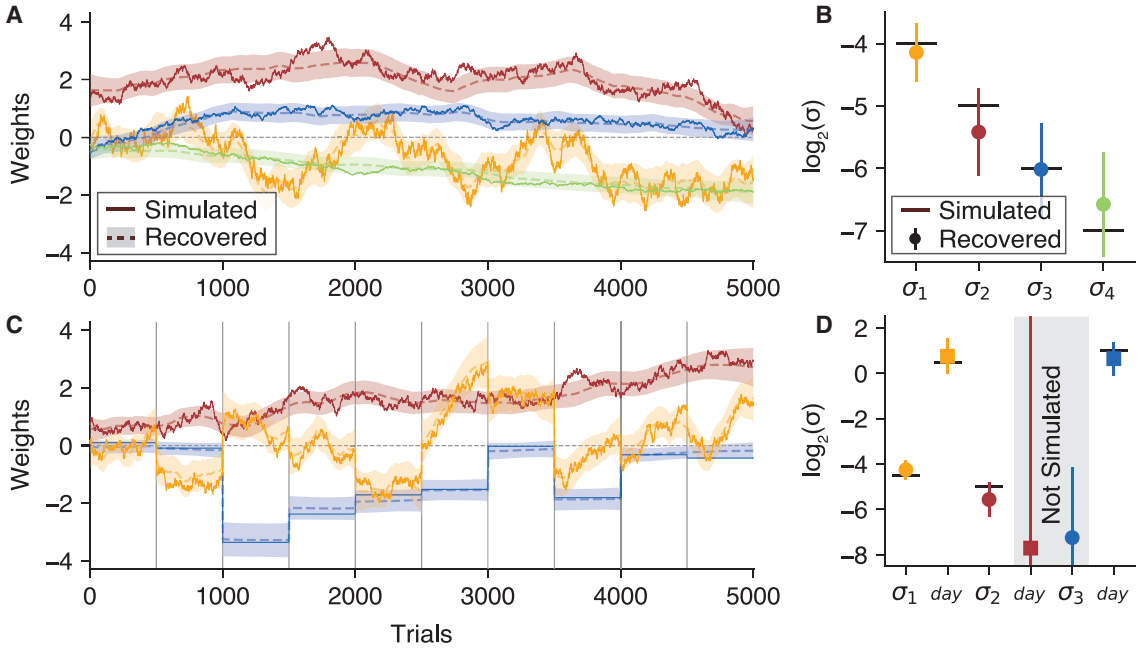


Figure 2. Recovering psychophysical weights from simulated data

(A) We simulated a set of $K = 4$ weights W that evolved for $T = 5{,}000$ trials (solid lines). We then used PsyTrack to recover these weights (dashed lines), with a shaded region indicating a 95% credible interval. The full optimization takes <1 min on a laptop; see Figure S1 for more information.

(B) In addition to recovering the weights, we recovered the smoothness hyperparameter σ_k for each weight, also plotted with a 95% credible interval. True values σ_k are plotted as solid black lines.

(C) We simulated a set of K = 3 weights with session boundaries every 500 trials (vertical black lines) and added a second set of hyperparameters "σ_{day}" allowing for larger weight changes between sessions. The yellow weight had non-zero σ and σ_{day} hyperparameters, allowing it to evolve trial-to-trial as well as "jump" at session boundaries. The blue weight, however, had $\sigma = 0$, so it was constant during each session and jumped only at session boundaries. The red weight, conversely, had $\sigma_{day} = 0$, and thus evolved like the weights in (A). See Figure S2 for weight trajectories recovered for this dataset without the use of any σ_{day} hyperparameters.

(D) We recovered the smoothness hyperparameters θ for the weights in (C). Although the simulation had only 4 non-zero hyperparameters, PsyTrack inferred both a σ and a σ_{day} hyperparameter for all 3 weights. The model appropriately assigned small values to the two zero-valued hyperparameters (gray shading).

Augmented model for capturing changes between sessions

One limitation of the model described above is that it does not account for the fact that experiments are typically organized into sessions, each containing tens to hundreds of consecutive trials, with large gaps of time between them. The basic PsyTrack model makes no allowance for the possibility that weights may change much more between sessions than between other pairs of consecutive trials. This assumption is unrealistic; if the animal either forgets or exhibits consolidation between sessions, the weights may exhibit much larger changes than between typical pairs of trials.

To overcome this limitation, we augmented the model to allow for larger weight changes between sessions. The augmented model has K additional hyperparameters, denoted $(\sigma_{day_1}, \ldots, \sigma_{day_K})$, which specify the prior standard deviation over weight changes between sessions or "days." A large value for σ_{day_k} means that the k^{th} weight can change by a large amount between sessions, regardless of how much it changes between other pairs of consecutive trials. The augmented model thus has $2K$ hyperparameters, with a pair of hyperparameters $(\sigma_k, \sigma_{day_k})$ for each of the K weights in $\mathbf{w}_t$.

We tested the performance of this augmented model using a second simulated dataset that included session boundaries every 500 trials (Figure 2C). We simulated $K = 3$ weights for $T = 5{,}000$ trials, with the input vector $\mathbf{x}_t$ and choices y_t on each trial sampled as in the first dataset. The red weight was simulated like the red weight in Figure 2A; that is, using only the standard σ and no σ_{day} hyperparameter. Conversely, the blue weight was simulated with a non-zero σ_{day} hyperparameter and $\sigma = 0$, making the weight constant within each session, but allowing "jumps" at session boundaries. The yellow weight was simulated with non-zero values for both hyperparameters, allowing it to smoothly evolve within a session and jump by larger amounts betwen sessions. Once again, we found that the recovered weights closely agree with the true weights (see Figure 2).

Note that we can also consider scenarios in which behavior changes suddenly at an arbitrary trial within a session, contradicting the assumptions of the PsyTrack model that weights evolve smoothly within a session. If we apply PsyTrack to datasets in which weights undergo such a step change, the method will infer a smoothed version of the true step, where the steepness is controlled by that weight's σ value (larger σ allows for a steeper slope). Figure S2 shows an empirical test of this phenomenon

Please cite this article in press as: Roy et al., Extracting the dynamics of behavior in sensory decision-making experiments, Neuron (2020), https://doi.org/10.1016/j.neuron.2020.12.004

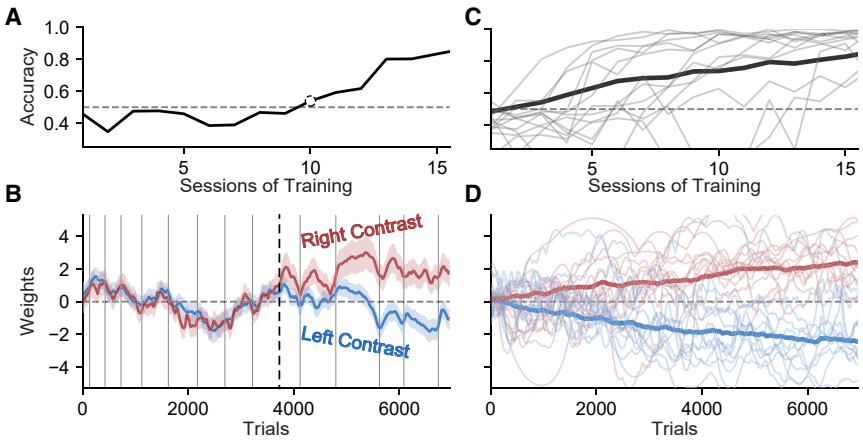


Figure 3. Visualization of early learning in IBL mice

(A) The accuracy of an example mouse over the first 16 sessions of training on the IBL task. We calculated accuracy only from "easy" high-contrast (50% and 100%) trials because lower-contrast stimuli were only introduced later in training. The first session with above-chance performance (50% accuracy) is marked with a dotted circle.

(B) Inferred weights for left (blue) and right (red) contrasts for the same example mouse and sessions shown in (A). Gray vertical lines indicate session boundaries. The black dotted line marks the start of the 10th session, when left and right weights first diverged, corresponding to the first session with above chance performance (shown in A). See Figure S3 for models using additional weights.

(C) Accuracy for a random subset of individual IBL mice (gray), along with average accuracy of the entire population (black).

(D) The psychophysical weights for left and right contrasts for the same subset of mice depicted in (C) (light red and blue), along with population averages (dark red and blue). (For visual clarity, we omitted the σ_{day} hyperparameters that would allow for jumps between sessions).

using the simulated dataset from Figure 2C but without the σ_{day} hyperparameters to capture the jumps at session boundaries.

Characterizing learning trajectories in the IBL task

We now turn to real data and show how PsyTrack can be used to characterize diverse trajectories of learning in a large cohort of animals. We examined a dataset from the IBL containing behavioral data from over 100 mice on a standardized sensory decision-making task (Figure 1A).

We began by analyzing choice data from the first 16 sessions of training. Figure 3A shows the learning curve (defined as the fraction of correct choices per session) for an example mouse over the first several weeks of training. Early training sessions used "easy" stimuli (100% and 50% contrasts) only, with harder stimuli (25%, 12.5%, 6.25%, and 0% contrasts) introduced only later in training as the animal's accuracy improved. To keep the metric consistent, we calculated accuracy only from easy-contrast trials on all of the sessions.

Although traditional analyses of learning rely on coarse performance metrics such as accuracy-per-session, PsyTrack offers a detailed characterization of the evolving strategies at the timescale of single trials. Figure 3B shows estimates of the time-varying weights on left-side contrast values (blue) and right-side contrast values (red) for an example mouse. During the first nine sessions, these two weights fluctuated together, indicating that the probability of making a rightward choice was independent of whether the stimulus was on the left or the right side of the screen. Positive (negative) fluctuations in these weights corresponded to a bias toward rightward (leftward) choices. These fluctuations indicated that the strategy was not constant across these sessions, even though accuracy remained at chance level.

At the start of the 10th session, the left and right stimulus weights began to diverge. Positive values of the right weight mean that right-side stimuli led to rightward choices, while negative values of the left weight mean that left-side stimuli led to leftward choices. The divergence of left and right weights therefore corresponds to an increase in accuracy. This divergence continued throughout the subsequent 6 sessions, gradually increasing accuracy to over 80% by the 16th session.

However, the learning trajectory of this example mouse was by no means characteristic of the entire cohort. Figures 3C and 3D shows the empirical learning curves (above) and inferred weight trajectories (below) from a dozen additional mice selected randomly from the IBL dataset. The light red and blue lines show the right and left weights for individual mice, whereas the dark red and blue lines show the average weights across the entire population. While we see a smooth and gradual divergence between the average left and right weights, there is great diversity in the dynamics of the weight trajectories of individual mice.

Adaptive bias modulation in a non-stationary task

Once training has progressed to include all contrast values, the IBL task undergoes a final modification. Instead of left and right stimuli appearing with an equal probability of 0.5 on each trial, the task statistics become nonstationary with alternating "left bias" and "right bias" blocks. Within a left bias block, the ratio of left contrasts to right contrasts is 80:20, whereas within right blocks the ratio is 20:80. These "bias blocks" are of variable duration, and some sessions begin with an "unbiased" 50:50 block for calibration purposes.

Figure 4 shows an analysis of behavior from the same example mouse from Figure 3 over its first 50 sessions of training, which includes the introduction of bias blocks. The pink box (Figure 4A) indicates a period of 3 sessions, which includes the last session without bias blocks and the first 2 sessions with bias blocks. The purple box indicates 2 sessions several weeks of training later, at the end of a period designated as "late bias blocks." In Figure S5, we show weight trajectories for an example session from each of these periods in training and validate that psychometric curves predicted from our model closely match curves computed directly from the behavioral data.

To examine how behavior changed during the onset of bias blocks, we applied PsyTrack to the three early bias block sessions (Figure 4B). The left and right stimulus weights are shown

Please cite this article in press as: Roy et al., Extracting the dynamics of behavior in sensory decision-making experiments, Neuron (2020), https://doi.org/10.1016/j.neuron.2020.12.004

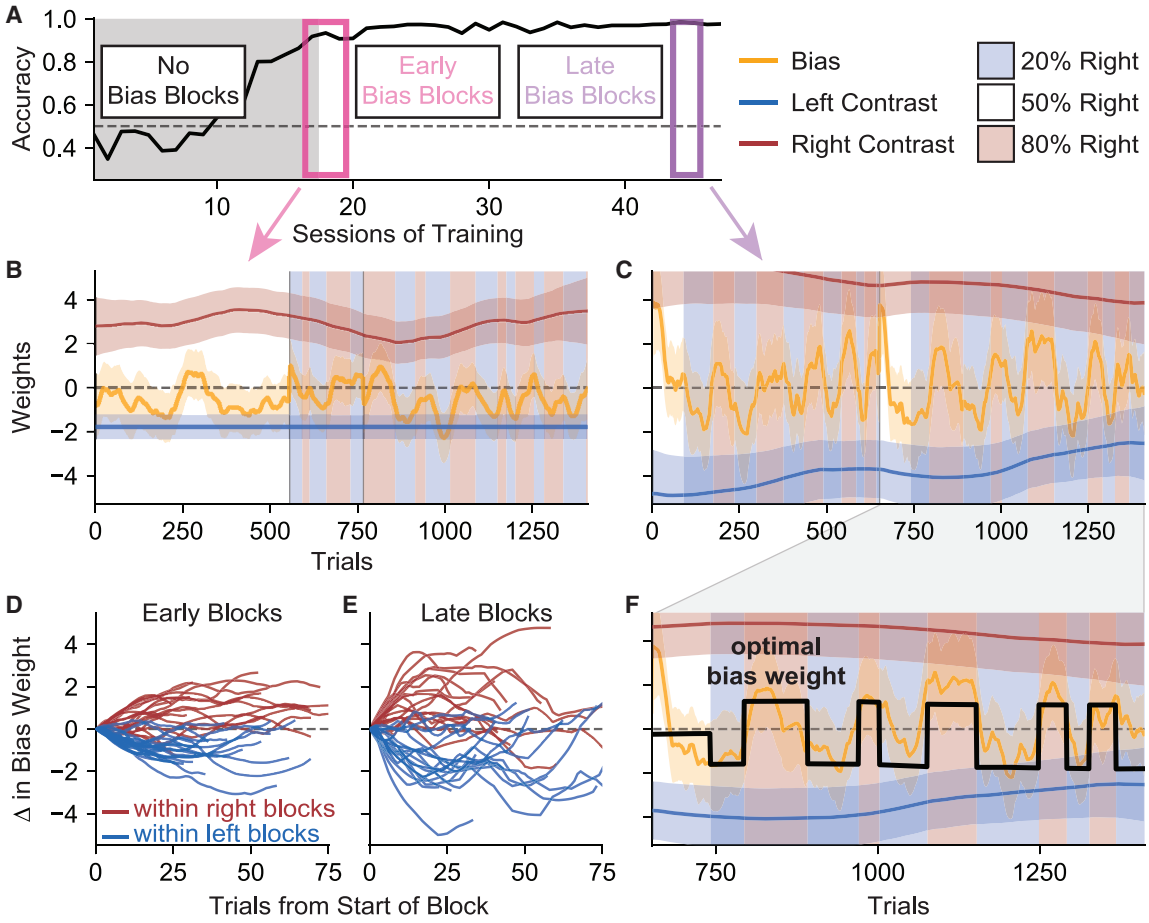


Figure 4. Adaptation to bias blocks in an example IBL mouse

(A) An extension of Figure 3A to include the first 50 sessions (several months) of training. Starting on session 17, our example mouse was introduced to alternating blocks of 80% right stimulus trials (right blocks) and 80% left stimulus trials (left blocks). The sessions in which these bias blocks were first introduced are outlined (pink), as are 2 sessions from later in training in which the mouse has adapted to the block structure (purple). Figure S5 validates model fits to sessions 10, 20, and 40 against psychometric curves generated directly from behavior.

(B) Three psychophysical weights evolving during the transition to bias blocks, with right (left) blocks indicated by red (blue) vertical stripes. Colored lines show left (blue) and right (red) stimulus weights, as well as a bias weight (yellow). See Figure S4 for analyses using alternate parameters.

(C) After several weeks of training on the bias blocks, the mouse learned to quickly adapt its behavior to the alternating block structure, as can be seen in the dramatic oscillations of the yellow bias weight in phase with the blocks of biased stimuli.

(D) Changes in bias weight relative to block transition for all blocks during the first 3 sessions with bias blocks, normalized to begin at zero. Even during these early bias block sessions, the red (blue) lines show that the bias tended to move up (down) during right (left) blocks, consistent with a strategy in which the bias weight tracks the stimulus probability.

(E) Same as (D) for 3 sessions during the late bias blocks period. Note that bias weights depart more rapidly from zero at the start of each block, indicating that the mice have adapted their behavior to the block structure of the task.

(F) For the second session from the late bias blocks shown in (C), we calculated an "optimal" bias weight (black) given the animal's stimulus weights and the ground truth block transition times (inaccessible to the mouse). This optimal bias closely matches the empirical bias weight recovered using PsyTrack (yellow), indicating that the strategy of the animal was approximately optimal for maximizing reward under the task structure. Although the bias weight may appear to "anticipate" the start of the next block, this is an artifact of the smoothing induced by the model (see Figure S6).

in blue and red, respectively, and a third psychophysical weight, in yellow, corresponds to choice bias. When this choice bias weight is positive (negative), the animal has an increased probability of choosing right (left), independent of other inputs. While task accuracy improves as the right weight grows more positive and the left weight grows more negative, the "optimal" value of the bias weight is naively 0 (no *a priori* preference for either side). However, this is only true when stimuli are presented with a 50:50 ratio and the two stimulus weights are of equal and opposite size.

For this mouse, we see that on the last session before the introduction of bias blocks, the stimulus weights were large and opposite, and bias was near zero. When the bias blocks began in the next session, the bias weight did not change in any obvious way to reflect the change in stimulus ratio. In

Please cite this article in press as: Roy et al., Extracting the dynamics of behavior in sensory decision-making experiments, Neuron (2020), https://doi.org/10.1016/j.neuron.2020.12.004

Figure 4C, however, we see that after several weeks of training with bias blocks, the bias weight exhibited large fluctuations synchronized to the block transitions.

We examined this phenomenon more fully in Figures 4D–4F. To better examine how the mouse's strategy changed within a bias block, we plotted the bias weight as a function of time since the start of each block, normalized to begin at zero. Figure 4D shows the bias weight changes for all bias blocks in the the first 3 sessions with bias blocks, with "left block" weight changes in red and "right block" changes in blue. Viewed in this way, we can see that there was some adaptation to the stimulus statistics within a bias block, even within the first few sessions. Within only a few dozen trials, the choice bias of the mouse tended to slowly drift rightward during right blocks and leftward during left blocks. We ran the same analysis on 3 sessions near the end of the "late bias blocks" period. This revealed substantially larger changes in the bias weight after a transition to a new block (Figure 4E). While it may appear that the animal proactively adjusted its bias toward the end of the longer blocks, as if in anticipation of the coming block, this is largely an artifact of the smoothing induced by the PsyTrack model. Figure S6 shows further analysis and presents a simple method to remove this artifact and test for true anticipation effects.

We can further analyze the choice bias of the animal in response to the bias blocks by returning to the notion of an "optimal" bias weight. As mentioned above, the optimal bias weight for the standard "unbiased" task is zero when the stimulus weights are equal and opposite. However, a non-zero bias weight can improve performance during bias blocks. Even if the stimulus weights were large enough to give nearly perfect performance on all trials with non-zero contrast, $1/9^{th}$ of the trials had 0% contrast stimuli on both sides, meaning the animal had to guess. If the bias weight adapted perfectly with the bias blocks, the mouse could get the 0% contrasts trials correct with 80% accuracy instead of 50%, increasing its total reward rate.

To assess the degree to which the mouse adjusted its bias in an optimal manner, we used PsyTrack to calculate the optimal bias weight. Here, we define an optimal bias weight on each trial as the value of the weight that maximizes expected accuracy assuming that (1) the left and right contrast weights recovered from the data are considered fixed, (2) the precise timings of the block transitions are known, and (3) the distribution of contrast values within each block is known. Figure 4F shows the resulting optimal bias (black line), overlaid with the empirical bias inferred from the animal's behavior (yellow line). Note that the optimal bias weight relied on knowledge of the block transition times, which was inaccessible to the mouse, but also changed subtly within blocks to account for changes in stimulus weights. We found that, in most blocks, the empirical bias matched the optimal bias weight closely. In fact, we can calculate that the mouse would only increase its expected accuracy from 86.1% to 89.3% by using the optimal bias instead of its empirical bias weight.

Trial-history effects dominate early behavior in Akrami rats

To further explore the capabilities of PsyTrack, we analyzed behavioral data from another binary decision-making task previously reported in Akrami et al. (2018), in which both rats and human subjects were trained on versions of the task (referred to hereafter as "Akrami rats" and "Akrami humans"). This auditory parametric working memory task requires an observer to listen to two white noise auditory stimuli, stimulus A, then stimulus B, which have different amplitudes and are separated by a delay (Figure 5A). If A is louder than B, then the rat must nose poke right to receive a reward and vice-versa.

Figure 5B shows an analysis of 12,500 trials of behavior from an example rat. Despite the new task and species, there are several similarities to the results from the IBL mice shown in Figures 3 and 4. The auditory stimuli A and B are the task-relevant variables (red and blue weights, respectively) and are similar to the left and right contrasts in the IBL task, and animals exhibit bias (yellow weight) in both tasks. However, while at least one stimulus was non-zero on each trial in the IBL task, both the auditory stimuli were present on every trial in the Akrami task. (Inputs were also parametrized differently, see STAR methods).

For this dataset, we added "history" weights that capture dependencies on the previous trial. We included a "previous stimuli" variable that represents the average of the stimulus A and B amplitudes on the previous trial (green), a "previous answer" variable, which indicates the rewarded (or correct) side on the previous trial (purple), and a "previous choice" variable, which indicates the choice on the previous trial (cyan). Note that these trial-history variables were always irrelevant in this task, meaning that the animal would maximize performance by setting the corresponding weights to zero. Despite this fact, previous work has shown that choice behavior often depends on trial history, especially early in training (for the impact of history regressors during early training in an IBL mouse, see Figure S3A) (Busse et al., 2011; Frund et al., 2014; Akrami et al., 2018).

For the example rat shown in Figure 5B, we found that the trial-history and bias weights dominated behavior early in training. In contrast, the weights on auditory stimuli A and B were initially close to zero, indicating that stimuli had a minimal effect on choice at the start of training. However, as training progressed, the history and bias weights shrank while the stimulus weights diverged from zero with opposite signs. Note that the stimulus B weight diverged from zero very early in training, while the stimulus A weight did not become positive until after several tens of sessions. In the context of the task, this makes intuitive sense: the association between a louder stimulus B and reward on the left is comparatively easy, since B occurs immediately before the choice. Making the association between a louder stimulus A and reward on the right is much more difficult to learn due to the delay period between A and B.

The positive value of all three history weights matched expectations from previous literature. The positive value of the weight on previous answer indicates that the animal preferred to go right (left) when the correct answer on the previous trial was also right (left). This is a commonly observed behavior known as a "win-stay/lose-switch" strategy, in which an animal will repeat its choice from the previous trial if it was rewarded or otherwise switch sides. The positive value of the previous-choice weight

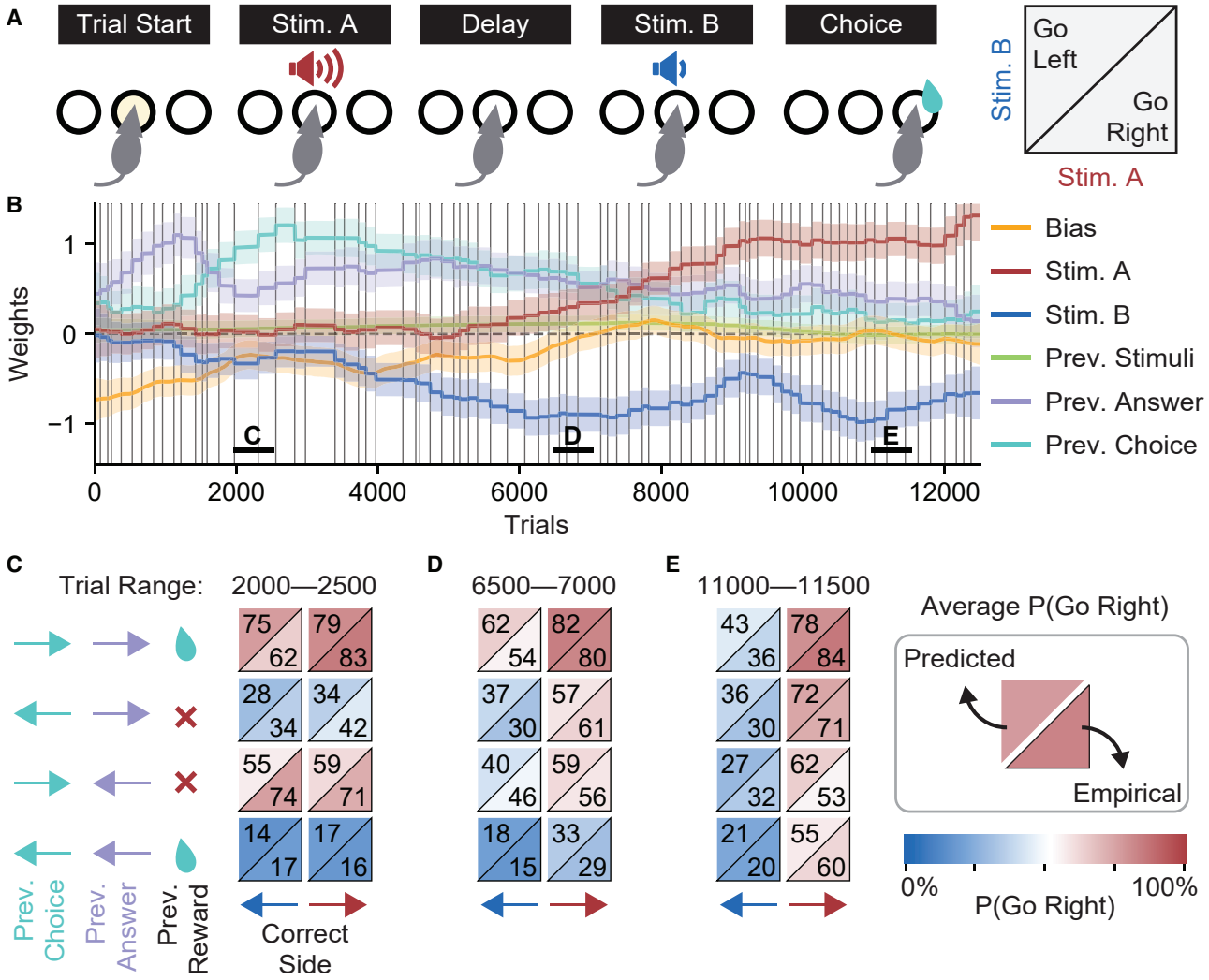


Figure 5. Visualization of learning in an example Akrami rat

(A) For these data from Akrami et al. (2018), a delayed response auditory discrimination task was used in which a rat experiences an auditory white noise stimulus of a particular amplitude (stimulus A), a delay period, a second stimulus of a different amplitude (stimulus B), and finally the choice to go either left or right. If stimulus A was louder than stimulus B, then a rightward choice triggers a reward, and vice-versa.

(B) The psychophysical weights recovered from the first 12,500 trials of an example rat. "Prev. stimuli" is the average amplitude of stimuli A and B presented on the previous trial; "prev. answer" is the rewarded (correct) side on the previous trial; "prev. choice" is the animal's choice on the previous trial. Black vertical lines are session boundaries. Figure S7 reproduces the analyses of this figure using a model with no history regressors; the remaining three weights look qualitatively similar.

(C–E) Within 500 trial windows starting at trials 2,000 (C), 6,500 (D), and 11,000 (E), trials are binned into 1 of 8 conditions according to 3 variables: the previous choice, the previous answer, and the correct side of the current trial. For example, the bottom left square is for trials in which the previous choice, the previous correct answer, and the current correct side are all left. The number in the bottom right of each square is the percent of rightward choices within that bin, calculated directly from the empirical behavior. The number in the top left is a prediction of that same percentage made using the cross-validated weights of the model. Close alignment of predicted and empirical values in each square indicate that the model is well validated, which is the case for each of the 3 training periods.

indicates that the animal preferred to go right (left) when it also went right (left) on the previous trial. This is known as a "perseverance" behavior: the animal prefers to simply repeat the same choice it made on the previous trial, independent of reward or task stimuli (Busse et al., 2011). Finally, the slight positive weight on the previous stimuli indicates that the animal was biased toward the right when the stimuli on the previous trial were louder than average, just as a louder stimulus A leads to more rightward choices. This corroborates an important finding from the original paper: choice biases are consistent with the rat's memory of stimulus A contracting toward the stimulus mean from the previous trial (Akrami et al., 2018), although we note that the analysis there was done on post-training behavior and used the 20–50 most recent trials to calculate an average

Please cite this article in press as: Roy et al., Extracting the dynamics of behavior in sensory decision-making experiments, Neuron (2020), https://doi.org/10.1016/j.neuron.2020.12.004

previous stimulus term (see also Papadimitriou et al., 2015 and Lu et al., 1992). To validate our use of these history regressors, we first analyzed the empirical choice behavior to verify that this strong dependence on the previous trial exists. Then, we examined whether the PsyTrack model with history regressors effectively captures this dependence.

Figures 5C–5E shows 3 different windows of 500 trials each, taken from different points during the training of our example rat. Within a particular window, the trials were binned according to 3 conditions: the choice on the previous trial ("prev. choice"), the correct answer of the previous trial ("prev. answer"), and the correct answer on the current trial ("correct side"). This gives $2^3 = 8$ conditions, represented by the 8 boxes. For example, the box in the lower left corresponds to trials in which the previous choice, previous answer, and current answer all are left. The 2 numbers within each box represent the percentage of trials within that condition (and within that 500-trial window) in which the rat chose to go right. The number in the bottom right of each square was calculated directly from the empirical behavior, whereas the number in the top left was calculated from the fitted model. Specifically, for each trial, we used the dot product of the model weights and the inputs on that trial to calculate the probability of a rightward choice (Equation 1); we then averaged these probabilities for all trials within a particular box. The color of each half-box maps directly to its average *P*(rightward) value, red for rightward and blue for leftward.

Focusing first on empirical data from early in training (the values in the bottom right of each square in Figure 5C), we can verify that behavior was strongly dependent on the previous trial. In a well-trained rat, the left column would all be blue, and the right column would all be red, since choices would align with the correct answer on the current trial. Instead, the empirical behavior was strongly history dependent, with the rat only going right on 16% of rightward trials if the previous choice and answer were both left (fourth row, right column). As the rat continued to train, we observed that the influence of the previous trial on choice behavior decreased, although it had not fully disappeared even by trial 11,000 (Figure 5E).

With the influence of the previous trial firmly established, we would like to compare these measurements against the model predictions. For almost all of the boxes in each of the trial windows, we observed that the predictions of the model align closely with the empirical choice behavior. In fact, only one predicted value existed outside the 95% confidence interval of our empirical measurement (third row, left column in Figure 5C). This is a strong validation that the PsyTrack model accurately captured the rat's decision-making behavior. We repeated this analyses in Figure S7 with a model without history regressors and show that this model was not able to capture any of the dependence on the previous trial evident in the empirical choice behavior. Note that all model predictions were evaluated via 10-fold cross-validation, so they represented true predictions on held-out data (see STAR methods).

Behavioral trends across the population of Akrami rats

We applied PsyTrack to our entire population of Akrami rats in order to identify commonalities and differences in learning behavior across animals during training. Figure 6 shows inferred weight trajectories for the first 20,000 trials of training for a population of 19 rats. Figure 6A shows the weights for auditory stimuli A and B, confirming the observation from our example rat that the stimulus B weight tended to diverge from zero earlier than the stimulus A weight. We observed large variability across animals in the bias weight trajectory (Figure 6B), although this variability was greatest early in training and averaged out to approximately 0 across animals. The slight positive weight on the previous stimuli was consistent across all rats (Figure 6C). Finally, the prevalence of both "win-stay/lose-switch" and "perseverance" behaviors across the population was evident in the positive weights on the previous answer and previous choice (Figures 6D and 6E), although there was substantial variation in the trajectories of these weights.

Figure 6F shows the average σ and σ_{day} values for each weight. The similar average σ values of the weights of stimuli A and B (blue and red circles) indicate a similar degree of smoothness in weight trajectories within a session, but the stimulus A weight exhibited much smaller changes (or "jumps") between sessions, as indicated by its relatively low σ_{day} average (red square). The individual bias weights in Figure 6B exhibited the largest within-session fluctuations, as reflected by the bias weight having the highest average σ value (yellow circle). Finally, the previous-stimuli weight had almost no variation across trials, as reflected in its markedly low average σ and σ_{day} values (green circle and square, Figure 6F).

In contrast to rats, human behavior is stable

Akrami et al. (2018) adapted the same auditory discrimination task to human subjects (Figure 7A). The weights inferred from an example human subject are shown in Figure 7B, and the weights from all of the human subjects are shown together in Figure 7C.

It is useful to contrast the weights recovered from the human subjects to the weights recovered from the rats. Since the rules of the task were explained to human subjects, one would intuitively expect that human weights would initialize at "correct" values corresponding to high performance and would remain constant throughout training. PsyTrack allows us to test these expectations explicitly. Figure 7C shows that the model does indeed confirm our intuitions: all four weights remained relatively stable throughout the experiment, although choice bias did fluctuate around zero for some subjects. Two of the history weights that dominated early behavior in rats, previous answer and previous choice, did not improve predictions in human data, so we removed those weights from the model (see Figure S8). The slight positive weight on previous stimuli remained for all subjects, however, and there was a slight asymmetry in the magnitudes of the A and B stimulus weights in many subjects.

Including history regressors boosts predictive power

To explore additional applications of PsyTrack, we can extend our analysis of the Akrami rats to quantify the importance of history regressors for characterizing decision-making behavior. Using the example rat from Figure 5, we sought to quantify the difference between a model that included the 3 history regressors and a model without them. To do this, we

Please cite this article in press as: Roy et al., Extracting the dynamics of behavior in sensory decision-making experiments, Neuron (2020), https://doi.org/10.1016/j.neuron.2020.12.004

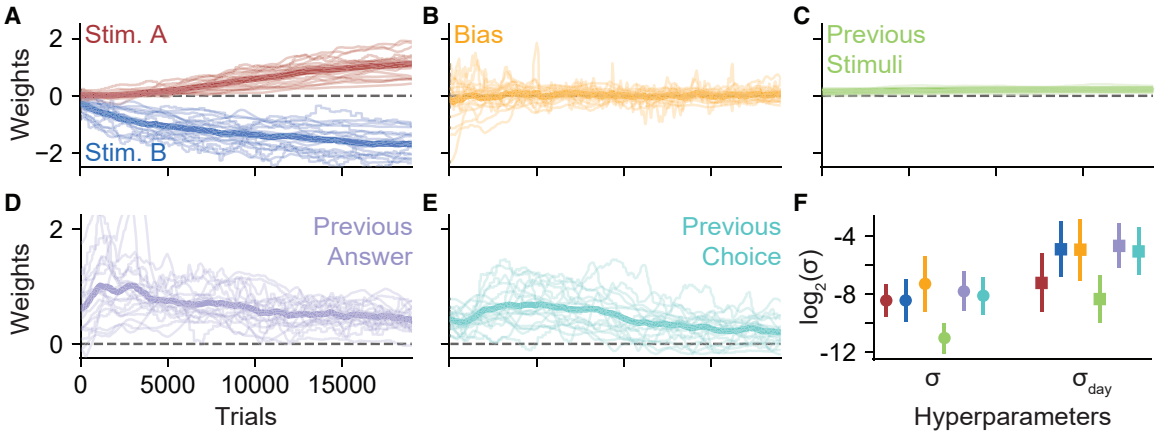


Figure 6. Population psychophysical weights from Akrami rats

The psychophysical weights during the first T = 20,000 trials of training, plotted for all rats in the population (light lines), plus the average weight (dark line): (A) auditory stimuli A and B, (B) bias, (C) previous stimuli, (D) previous answer, and (E) previous choice. (F) The average σ and σ_{day} for each weight (±1 SD), color coded to match the weight labels in (A)–(E).

refit a model to the data from Figure 5B using only bias and stimulus A and B weights (see Figure S7) and then calculated the predicted accuracy of both models at each trial via cross-validation (see STAR methods). Next, we binned the trials according to the predicted accuracy of the model. Finally, for the trials within each bin, we computed the empirical accuracy of the model (defined as the fraction of trials in a bin in which the prediction of the model matched the animal's choice). Figure 8A shows a comparison of model-predicted accuracy and empirical accuracy for the basic model without history weights. Points below the diagonal represent overconfident predictions, where the model is correct less often than expected from the probabilities it produces, whereas points above the line indicate underconfidence, where the model is correct more often than expected. The fact that the data (shown with 95% confidence intervals on their empirical accuracy) lie mostly along the diagonal shows that the model is well calibrated.

The histogram in Figure 8B shows the number of trials in each of the probability bins in Figure 8A. We can see that the model almost never predicted choices with >80% probability. The black star shows the average model-predicted accuracy (61.9%) and corresponding empirical accuracy (also 61.9%).

Figures 8C and 8D show an identical analysis for the full model with history dependence. That most data points still lie along the diagonal means that the model remains well calibrated. Note, however, that the data now extend much further up the diagonal, indicating that the model makes more confident predictions on some trials. Figure 8D shows that a meaningful fraction of trials now have a predicted accuracy greater than the most confident predictions made by the model without history regressors. In fact, the choices on some trials can be predicted with >95% probability. As the black star in Figure 8C indicates, the inclusion of history regressors improves the predicted accuracy of the model to 68.4%. This confirms the importance of history regressors when characterizing the decision-making behavior of the Akrami rats.

Discussion

Sensory decision-making strategies evolve continuously over the course of training, driven by learning signals as well as noise. Even after training is complete, these strategies can continue to fluctuate, both within and across sessions. Tasks with non-stationary stimuli or rewards may require continual changes in strategy to maximize reward (Piet et al., 2018). However, standard methods for quantifying sensory decision-making behavior, such as learning curves (Figure 3A) and psychometric functions (Figures 1D and S5), are not able to describe the evolution of complex decision-making strategies over time. To address this shortcoming, we developed PsyTrack, which parametrizes time-varying behavior using a dynamic generalized linear model with time-varying weights. The rates of change of these weights across trials and across sessions are governed by weight-specific hyperparameters, which we infer directly from data using evidence optimization. We have applied PsyTrack to data from two tasks and three species and shown that it can characterize trial-to-trial fluctuations in decision-making behavior and provide a quantitative foundation for more targeted analyses.

PsyTrack contributes to a growing literature on the quantitative analysis of time-varying behavior. An influential paper by Smith et al. (2004) introduced a change-point-detection approach for identifying the trial at which learning produced a significant departure from chance behavior in a decision-making task. We have extended this model by moving beyond change points to track detailed changes in behavior over the entire training period and beyond, and by adding regression weights for a wide variety of task covariates that may also influence choices. Previous work has shown that animals frequently adopt strategies that depend on previous choices and previous stimuli (Busse et al., 2011; Frund et al., 2014; Akrami et al., 2018), even when such strategies are suboptimal. Our approach builds upon a state-space approach for the dynamic tracking of behavior (Bak et al., 2016) using the decoupled Laplace method (Wu et al., 2017; Roy et al., 2018a) to scale up analysis and make it practical for modern

Please cite this article in press as: Roy et al., Extracting the dynamics of behavior in sensory decision-making experiments, Neuron (2020), https://doi.org/10.1016/j.neuron.2020.12.004

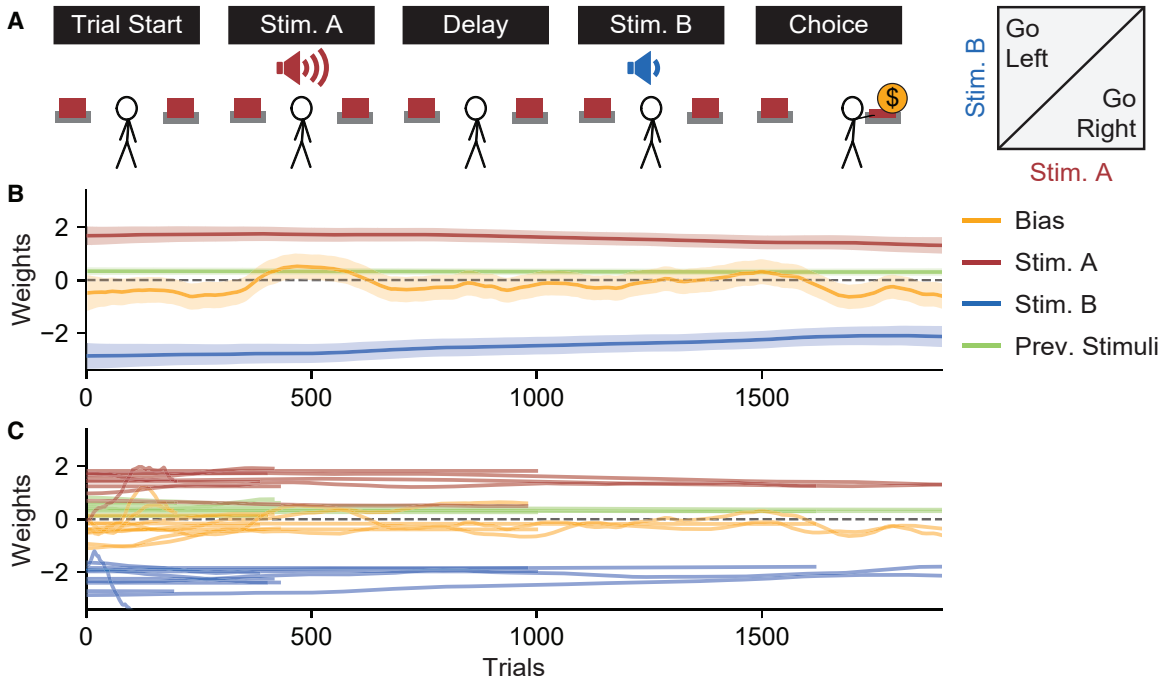


Figure 7. Population psychophysical weights from Akrami human subjects

(A) The same task used by the Akrami rats in Figure 5A, adapted for human subjects.

(B) The weights for an example human subject. Human behavior is not sensitive to the previous correct answer or previous choice, so the corresponding weights are not included in the model (see Figure S8 for a model that includes these weights).

(C) The weights for the entire population of human subjects. Human behavior was evaluated in a single session of variable length.

behavioral datasets (STAR methods). In particular, the efficiency of our algorithm allows for routine analysis of large behavior datasets, with tens of thousands of trials, within minutes on a laptop (see Figure S1).

We anticipate a variety of use cases for PsyTrack. First, experimenters training animals on binary decision-making tasks can use PsyTrack to better understand the diverse range of behavioral strategies seen in early training (Cohen and Schneidman, 2013). This will facilitate the design and validation of new training strategies, which could ultimately open the door to more complex tasks and faster training times. Second, studies of learning will benefit from the ability to analyze behavioral data collected during training, which is often discarded and left unanalyzed. Third, the time-varying weights inferred under the PsyTrack model will lend themselves to downstream analyses. Specifically, our general method can enable more targeted investigations, acting as one step in a larger analysis pipeline (as in Figures 4 and 8). To facilitate these uses, we have released PsyTrack as a publicly available Python package (Roy et al., 2018b). The Google Colab notebook accompanying this work provides many flexible examples, and we have included a guide in the STAR methods devoted entirely to practical considerations.

The two assumptions of PsyTrack, that (1) decision-making behavior can be described by a set of GLM weights and (2) that these weights evolve smoothly over training, are well validated in the datasets explored here. However, these assumptions may not be true for all datasets. Behaviors that change suddenly may not be well described by the smoothly evolving weights in our model (see Figure S2), although allowing for weights to evolve more dramatically between sessions can mitigate this model mismatch (as in Figure 2C). Determining which input variables to include can also be challenging. For example, task-irrelevant covariates of decision making typically include the history of previous trials (Akrami et al., 2018; Frund et al., 2014; Corrado et al., 2005), which is not always clearly defined; depending on the task, the task-relevant feature may also be a pattern of multiple stimulus units (Murphy et al., 2008). We do not currently consider high-dimensional inputs to our model (e.g., images, complex natural signals), but incorporating an automatic relevance determination prior is an exciting future direction that would allow for model weights to be automatically pruned during inference (Tipping, 2001).

Furthermore, deciding how input variables ought to be parameterized can dramatically affect the accuracy of the model fit. For example, a transformation of the contrast levels used in our analysis of data from the IBL task allows the model to discount the impact of incorrect choices under extreme levels of perceptual evidence (e.g., 100% contrasts; see Figure S4) (Nassar and Frank, 2016). While many models discount the influence of such trials using a lapse rate, PsyTrack relies on a careful parametrization of the perceptual input in lieu of any explicit lapse (STAR methods). This flexibility gives the model the ability to account for a wide variety of behavioral strategies.

In the face of these potential pitfalls, it is important to validate our results. Thus, we have provided comparisons to more conventional measures of behavior to help assess the accuracy of our

Please cite this article in press as: Roy et al., Extracting the dynamics of behavior in sensory decision-making experiments, Neuron (2020), https://doi.org/10.1016/j.neuron.2020.12.004

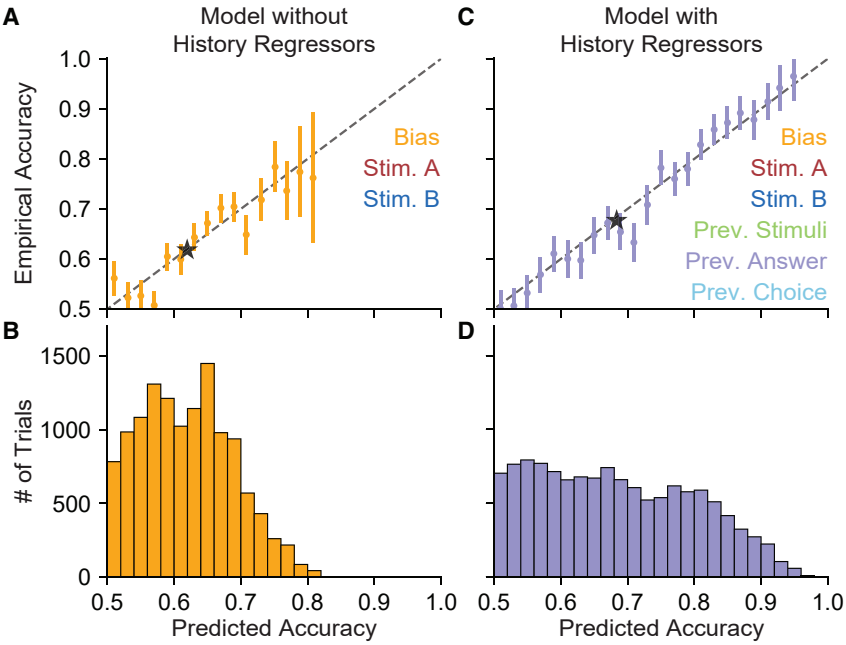


Figure 8. History regressors improve model accuracy for an example Akrami rat

(A) Using a model of our example rat that omits history regressors, we plot the empirical accuracy of the choice predictions of the model against the cross-validated predicted accuracy of the model. The black dashed line is the identity, in which the predicted accuracy of the model exactly matches the empirical accuracy (i.e., points below the line are overconfident predictions). The choice of the animal is predicted with 61.9% confidence on the average trial, precisely matching the empirical accuracy of the model of 61.9% (black star). Each point represents data from the corresponding bin of trials seen in (B). Empirical accuracy is plotted with a 95% confidence interval. See STAR methods for more information on the cross-validation procedure.

(B) A histogram of trials binned according to the predicted accuracy of the model.

(C) Same as (A), but for a model that also includes 3 additional weights on history regressors: previous stimuli, previous answer, and previous choice. We see that data for this model extends into regions of higher predicted and empirical accuracy, as the inclusion of history regressors allows the model to make stronger predictions. The choice of the animal is predicted with 68.4% confidence on the average trial, slightly overshooting the empirical accuracy of 67.6% (black star) of the model.

(D) Same as (B), but for the model including history regressors.

model (see Figures 5, 8, and S5). Furthermore, the Bayesian setting of our modeling approach provides approximate posterior credible intervals for both weights and hyperparameters, allowing for an evaluation of the uncertainty of our inferences about behavior.

The ability to quantify complex and dynamic behavior at a trial-by-trial resolution enables exciting future opportunities for animal training. The descriptive model underlying PsyTrack could be extended to incorporate an explicit model of learning that makes predictions about how strategy will change in response to different stimuli and rewards (Ashwood et al., 2020). Ultimately, this could guide the creation of automated optimal training paradigms in which the stimuli predicted to maximize learning on each trial is presented (Bak et al., 2016). There are also opportunities to extend the model beyond binary decision-making tasks, so that multi-valued choices (or non-choices—e.g., interrupted or "violation" trials) could also be included in the model (Churchland et al., 2008; Bak and Pillow, 2018). Our work opens up the path toward a more rigorous understanding of the behavioral dynamics at play as animals learn. As researchers continue to ask challenging questions, new animal training tasks will grow in number and complexity. We expect that PsyTrack will help guide those looking to better understand the dynamic behavior of their experimental subjects.

STAR★Methods

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Mouse subjects
  - Rat subjects
  - Human subjects
- METHOD DETAILS
  - Optimization: psychophysical weights
  - Optimization: smoothness hyperparameters
  - Selection of input variables
  - Parameterization of input variables
  - IBL task
  - Akrami task
  - A practical guide
- QUANTIFICATION AND STATISTICAL ANALYSIS
  - Cross-validation procedure
  - Calculation of posterior credible intervals

Supplemental Information

Supplemental Information can be found online at https://doi.org/10.1016/j.neuron.2020.12.004.

Acknowledgments

The authors thank A. Churchland, A. Pouget, M. Carandini, A. Urai, and Y. Niv for helpful feedback on the manuscript. We also thank K. Osorio and J. Teran for animal and laboratory support in collecting the rat high-throughput data. This work was supported by grants from the Wellcome Trust (209558 and 216324) and the Simons Foundation, to the IBL and the Simons Collaboration on the Global Brain (SCGB AWD543027), the NIH BRAIN Initiative (NS104899, R01EB026946), and a U19 NIH-NINDS BRAIN Initiative Award (5U19NS104648) (N.A.R. and J.W.P.).

Author contributions

Conceptualization, N.A.R., J.H.B., and J.W.P.; Methodology, N.A.R., J.H.B., and J.W.P.; Software, N.A.R.; Formal Analysis, N.A.R.; Investigation, N.A.R., the IBL, and A.A.; Resources, the IBL, A.A., C.D.B., and J.W.P.; Data Curation, N.A.R., the IBL, and A.A.; Writing – Original Draft, N.A.R.; Writing – Review & Editing, N.A.R., J.H.B., the IBL, A.A., C.D.B., and J.W.P.; Visualization, N.A.R.; Supervision, J.W.P.; Project Administration, N.A.R. and J.W.P.; Funding Acquisition, the IBL and J.W.P.

Declaration of interests

The authors declare no competing interests.

Received: May 25, 2020
Revised: October 23, 2020
Accepted: December 3, 2020
Published: December 30, 2020

References

Akrami, A., Kopec, C.D., Diamond, M.E., and Brody, C.D. (2018). Posterior parietal cortex represents sensory history and mediates its effects on behaviour. Nature *554*, 368–372.

Ashwood, Z., Roy, N.A., Bak, J.H., and Pillow, J.W. (2020). Inferring learning rules from animal decision-making. Adv. Neural Inf. Process. Syst. *34*, 33.

Bak, J.H., and Pillow, J.W. (2018). Adaptive stimulus selection for multi-alternative psychometric functions with lapses. J. Vision *18*, 4.

Bak, J.H., Choi, J.Y., Akrami, A., Witten, I., and Pillow, J.W. (2016). Adaptive optimal training of animal behavior. Adv. Neural Inf. Process. Syst *30*, 1947–1955.

Bishop, C.M. (2006). Pattern Recognition and Machine Learning (Springer).

Brunton, B.W., Botvinick, M.M., and Brody, C.D. (2013). Rats and humans can optimally accumulate evidence for decision-making. Science *340*, 95–98.

Burgess, C.P., Lak, A., Steinmetz, N.A., Zatka-Haas, P., Bai Reddy, C., Jacobs, E.A.K., Linden, J.F., Paton, J.J., Ranson, A., Schröder, S., et al. (2017). High-yield methods for accurate two-alternative visual psychophysics in head-fixed mice. Cell Rep. *20*, 2513–2524.

Busse, L., Ayaz, A., Dhruv, N.T., Katzner, S., Saleem, A.B., Schölvinck, M.L., Zaharia, A.D., and Carandini, M. (2011). The detection of visual contrast in the behaving mouse. J. Neurosci. *31*, 11351–11361.

Carandini, M. (2012). From circuits to behavior: a bridge too far? Nat. Neurosci. *15*, 507–509.

Carandini, M., and Churchland, A.K. (2013). Probing perceptual decisions in rodents. Nat. Neurosci. *16*, 824–831.

Churchland, A.K., Kiani, R., and Shadlen, M.N. (2008). Decision-making with multiple alternatives. Nat. Neurosci. *11*, 693–702.

Cohen, Y., and Schneidman, E. (2013). High-order feature-based mixture models of classification learning predict individual learning curves and enable personalized teaching. Proc. Natl. Acad. Sci. USA *110*, 684–689.

Corrado, G.S., Sugrue, L.P., Seung, H.S., and Newsome, W.T. (2005). Linear-Nonlinear-Poisson models of primate choice dynamics. J. Exp. Anal. Behav. *84*, 581–617.

Daw, N., and Courville, A. (2008). The pigeon as particle filter. Adv. Neural Inf. Process. Syst. *20*, 369–376.

Daw, N.D. (2011). Trial-by-trial data analysis using computational models. In Decision Making, Affect, and Learning: Attention and Performance XXIII, M.R. Delgado, E.A. Phelps, and T.W. Robbins, eds. (Oxford University Press) https://oxford.universitypressscholarship.com/view/10.1093/acprof:oso/9780199600434.001.0001/acprof-9780199600434-chapter-001.

Fassihi, A., Akrami, A., Esmaeili, V., and Diamond, M.E. (2014). Tactile perception and working memory in rats and humans. Proc. Natl. Acad. Sci. USA *111*, 2331–2336.

Frund, I., Wichmann, F.A., and Macke, J.H. (2014). Quantifying the effect of intertrial dependence on perceptual decisions. J. Vision *14*, 9.

Gold, J.I., and Shadlen, M.N. (2007). The neural basis of decision making. Annu. Rev. Neurosci. *30*, 535–574.

Green, D.M., and Swets, J.A. (1966). Signal Detection Theory and Psychophysics (Wiley).

Guo, Z.V., Hires, S.A., Li, N., O'Connor, D.H., Komiyama, T., Ophir, E., Huber, D., Bonardi, C., Morandell, K., Gutnisky, D., et al. (2014). Procedures for behavioral experiments in head-fixed mice. PLOS ONE *9*, e88678.

Hanks, T.D., and Summerfield, C. (2017). Perceptual decision making in rodents, monkeys, and humans. Neuron *93*, 15–31.

Hunter, J.D. (2007). Matplotlib: A 2d graphics environment. Comput. Sci. Eng. *9*, 90–95.

International Brain Laboratory, Bonacchi, N., Chapuis, G., Churchland, A., Harris, K.D., Rossant, C., Sasaki, M., Shen, S., Steinmetz, N.A., Walker, E.Y., et al. (2019). Data architecture and visualization for a large-scale neuroscience collaboration. BioRxiv, 827873.

International Brain Laboratory, Aguillon-Rodriguez, V., Angelaki, D.E., Bayer, H.M., Bonacchi, N., Carandini, M., Cazettes, F., Chapuis, G.A., Churchland, A.K., Dan, Y., et al. (2020). A standardized and reproducible method to measure decision-making in mice. bioRxiv. https://doi.org/10.1101/2020.01.17.909838.

Jones, E., Oliphant, T., and Peterson, P. (2001). In SciPy: open source scientific tools for Python https://www.scipy.org/.

Kattner, F., Cochrane, A., and Green, C.S. (2017). Trial-dependent psychometric functions accounting for perceptual learning in 2-AFC discrimination tasks. J. Vis. *17*, 3.

Krakauer, J.W., Ghazanfar, A.A., Gomez-Marin, A., MacIver, M.A., and Poeppel, D. (2017). Neuroscience needs behavior: correcting a reductionist bias. Neuron *93*, 480–490.

Lu, Z.L., Williamson, S.J., and Kaufman, L. (1992). Behavioral lifetime of human auditory sensory memory predicted by physiological measures. Science *258*, 1668–1670.

Murphy, R.A., Mondragon, E., and Murphy, V.A. (2008). Rule learning by rats. Science *319*, 1849–1851.

Nassar, M.R., and Frank, M.J. (2016). Taming the beast: extracting generalizable knowledge from computational models of cognition. Curr. Opin. Behav. Sci. *11*, 49–54.

Niv, Y. (2009). Reinforcement learning in the brain. J. Math. Psychol. *53*, 139–154.

Niv, Y. (2020). The primacy of behavioral research for understanding the brain. PsyArXiv. https://doi.org/10.31234/osf.io/y8mxe.

Niv, Y., Daniel, R., Geana, A., Gershman, S.J., Leong, Y.C., Radulescu, A., and Wilson, R.C. (2015). Reinforcement learning in multidimensional environments relies on attention mechanisms. J. Neurosci. *35*, 8145–8157.

Nocedal, J., and Wright, S.J. (2006). Quasi-Newton methods. In Numerical Optimization (Springer), pp. 135–163.

Papadimitriou, C., Ferdoash, A., and Snyder, L.H. (2015). Ghosts in the machine: memory interference from the previous trial. J. Neurophysiol. *113*, 567–577.

Piet, A.T., Hady, A.E., and Brody, C.D. (2018). Rats adopt the optimal timescale for evidence integration in a dynamic environment. Nat. Commun. *9*, 4265.

Pisupati, S., Chartarifsky-Lynn, L., Khanal, A., and Churchland, A.K. (2019). Lapses in perceptual decisions reflect exploration. bioRxiv, 613828.

Ratcliff, R., and Rouder, J.N. (1998). Modeling response times for two-choice decisions. Psychol. Sci. *9*, 347–356.

Roy, N.A., Bak, J.H., Akrami, A., Brody, C., and Pillow, J.W. (2018a). Efficient inference for time-varying behavior during learning. Adv. Neural Inf. Process Syst. *31*, 5695–5705.

Roy, N.A., Bak, J.H., and Pillow, J.W. (2018b). PsyTrack: open source dynamic behavioral fitting tool for Python. https://github.com/nicholas-roy/psytrack.

Rybicki, G.B., and Hummer, D.G. (1991). An accelerated lambda iteration method for multilevel radiative transfer. I-Non-overlapping lines with background continuum; Appendix B. Astron. Astrophys. *245*, 171–181.

Sahani, M., and Linden, J.F. (2003). Evidence optimization techniques for estimating stimulus-response functions. Adv. Neural Inf. Process. Syst. *15*, 317–324.

Samejima, K., Doya, K., Ueda, Y., and Kimura, M. (2004). Estimating internal variables and parameters of a learning agent by a particle filter. Adv. Neural Inf. Process.Syst. *16*, 1335–1342.

Smith, A.C., Frank, L.M., Wirth, S., Yanike, M., Hu, D., Kubota, Y., Graybiel, A.M., Suzuki, W.A., and Brown, E.N. (2004). Dynamic analysis of learning in behavioral experiments. J. Neurosci. *24*, 447–461.

Sutton, R.S. (1988). Learning to predict by the methods of temporal differences. Mach. Learn. *3*, 9–44.

Sutton, R.S., and Barto, A.G. (2018). Reinforcement Learning: An Introduction (MIT Press).

Suzuki, W.A., and Brown, E.N. (2005). Behavioral and neurophysiological analyses of dynamic learning processes. Behav. Cogn. Neurosci. Rev. *4*, 67–95.

Tipping, M.E. (2001). Sparse bayesian learning and the relevance vector machine. J. Mach. Learn. Res. *1*, 211–244.

Usher, M., Tsetsos, K., Yu, E.C., and Lagnado, D.A. (2013). Dynamics of decision-making: from evidence accumulation to preference and belief. Front. Psychol. *18*, https://doi.org/10.3389/fpsyg.2013.00758.

Wu, A., Roy, N.A., Keeley, S., and Pillow, J.W. (2017). Gaussian process based nonlinear latent structure discovery in multivariate spike train data. Adv. Neural Inf. Process. Syst. *30*, 3499–3508.

Please cite this article in press as: Roy et al., Extracting the dynamics of behavior in sensory decision-making experiments, Neuron (2020), https://doi.org/10.1016/j.neuron.2020.12.004

STAR★Methods

Key Resources Table

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Deposited data		
Human subject behavior dataset	Akrami et al., 2018	https://doi.org/10.6084/m9.figshare.12213671.v1
Mouse behavior dataset	International Brain Laboratory et al., 2020	https://doi.org/10.6084/m9.figshare.11636748.v7
Rat behavior dataset	Akrami et al., 2018	https://doi.org/10.6084/m9.figshare.12213671.v1
Software and algorithms		
IBL Python Library / ONE Light	International Brain Laboratory et al., 2019	https://github.com/int-brain-lab/ibllib/tree/master/oneibl
PsyTrack	Roy et al., 2018b	https://github.com/nicholas-roy/psytrack
SciPy ecosystem	Jones et al., 2001; Hunter, 2007	https://www.scipy.org/

Resource availability

Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the Lead Contact, Nicholas A. Roy (nicholas.roy.42@gmail.com).

Materials availability

This study did not generate any new materials.

Data and code availability

Each of the three datasets analyzed are publicly available. The mouse training data, from International Brain Laboratory et al. (2020): (https://doi.org/10.6084/m9.figshare.11636748.v7); the rat training data, from Akrami et al. (2018): (https://doi.org/10.6084/m9.figshare.12213671.v1); the humans subject training data, also from Akrami et al. (2018): (https://doi.org/10.6084/m9.figshare.12213671.v1).

Our code for fitting psychophysical weights to behavioral data is distributed as a GitHub repository (under a MIT license): https://github.com/nicholas-roy/PsyTrack. This code is also made easily accessible as a Python package, PsyTrack (installed via *pip install psytrack*). Our Python package relies on the standard SciPy scientific computing libraries as well as the Open Neurophysiology Environment produced by the IBL (Jones et al., 2001; Hunter, 2007; International Brain Laboratory et al., 2019).

We have assembled a Google Colab notebook (https://tinyurl.com/PsyTrack-colab) that will automatically download the raw data and precisely reproduce all figures from the paper. Our analyses can be easily extended to additional experimental subjects and act as a template for application of PsyTrack to new datasets.

Experimental model and subject details

Mouse subjects

101 experimental subjects were all female and male C57BL/6J mice aged 3-7 months, obtained from either Jackson Laboratory or Charles River. All procedures and experiments were carried out in accordance with the local laws and following approval by the relevant institutions: the Animal Welfare Ethical Review Body of University College London; the Institutional Animal Care and Use Committees of Cold Spring Harbor Laboratory, Princeton University, and University of California at Berkeley; the University Animal Welfare Committee of New York University; and the Portuguese Veterinary General Board. This data was first reported in International Brain Laboratory et al. (2020).

Rat subjects

19 experimental subjects were male Long-Evans rats (*Rattus norvegicus*) between the ages of 2 and 24 months. Animal use procedures were approved by the Princeton University Institutional Animal Care and Use Committee and carried out in accordance with National Institutes of Health standards. This data was first reported in Akrami et al. (2018).

Please cite this article in press as: Roy et al., Extracting the dynamics of behavior in sensory decision-making experiments, Neuron (2020), https://doi.org/10.1016/j.neuron.2020.12.004

Human subjects

11 human subjects (8 males and 3 females, aged 22–40) were tested and all gave their informed consent. Participants were paid to be part of the study. The consent procedure and the rest of the protocol were approved by the Princeton University Institutional Review Board. This data was first reported in Akrami et al. (2018).

Method details

Optimization: psychophysical weights

Our method requires that weight trajectories be inferred from the response data collected over the course of an experiment. This amounts to a very high-dimensional optimization problem when we consider models with several weights and datasets with tens of thousands of trials. Moreover, we wish to learn the smoothness hyperparameters $\theta = \{\sigma_1, \ldots, \sigma_K\}$ in order to determine how quickly each weight evolves across trials. The theoretical framework of our approach was first introduced in Bak et al. (2016). The statistical innovations facilitating application to large datasets, as well as the initial release of our Python implementation PsyTrack (Roy et al., 2018b), were first presented in Roy et al. (2018a).

We describe our full inference procedure in two steps. The first step is optimizing for weight trajectories W given a fixed set of hyperparameters, while the second step optimizes for the hyperparameters θ given a parametrized Gaussian approximation to the posterior. The full procedure involves alternating between the two steps until the hyperparameters converge.

For now, let $\mathbf{w}$ denote the massive weight vector formed by concatenating all of the K individual length-T trajectory vectors, where T is the total number of trials. We then define $\boldsymbol{\eta} = D\mathbf{w}$, where D is a block-diagonal matrix of K identical $T \times T$ difference matrices (i.e., 1 on the diagonal and -1 on the lower off-diagonal), such that $\boldsymbol{\eta}_t = \mathbf{w}_t - \mathbf{w}_{t-1}$ for each trial t. Because the prior on $\boldsymbol{\eta}$ is simply $\mathcal{N}(0, \Sigma)$, where Σ has each of the σ_k^2 stacked T times along the diagonal, the prior for $\mathbf{w}$ is $\mathcal{N}(0, C)$ with $C^{-1} = D^\top \Sigma^{-1} D$. The log-likelihood is simply a sum of the log probability of the animal's choice on each trial, $L = \sum_{t=1}^{T} \log p(y_t | \mathbf{x}_t, \mathbf{w}_t)$.

The log-posterior is then given by

$$\log p(\mathbf{w}|\mathcal{D}) = \frac{1}{2}\left(\log|C^{-1}| - \mathbf{w}^\top C^{-1}\mathbf{w}\right) + \sum_{t=1}^{T} \log p(y_t|\mathbf{x}_t, \mathbf{w}_t) + const, \qquad \text{(Equation 3)}$$

where $\mathcal{D} = \{(\mathbf{x}_t, y_t)\}_{t=1}^{T}$ is the set of user-defined input features (including the stimuli) and the animal's choices, and *const* is independent of $\mathbf{w}$.

Our goal is to find the $\mathbf{w}$ that maximizes this log-posterior; we refer to this maximum *a posteriori* (MAP) vector as $\mathbf{w}_{\text{MAP}}$.

We observe that the Hessian of our log-posterior is sparse:

$$H = \frac{\partial^2}{\partial \mathbf{w}^2} \log p(\mathbf{w}|\mathcal{D}) = C^{-1} + \frac{\partial^2 L}{\partial \mathbf{w}^2} \qquad \text{(Equation 4)}$$

where C^{-1} is a sparse (banded) matrix, and $\partial^2 L / \partial \mathbf{w}^2$ is a block-diagonal matrix. The block diagonal structure arises because the log-likelihood is additive over trials, and weights at one trial t do not affect the log-likelihood component from another trial t'.

We take advantage of this sparsity, using a variant of conjugate gradient optimization that only requires a function for computing the product of the Hessian matrix with an arbitrary vector (Nocedal and Wright, 2006). Since we can compute such a product using only sparse terms and sparse operations, we can utilize quasi-Newton optimization methods in SciPy to find a global optimum for $\mathbf{w}_{\text{MAP}}$, even for very large T (Jones et al., 2001).

Optimization: smoothness hyperparameters

So far we have addressed the problem of finding a global optimum $\mathbf{w}_{\text{MAP}}$, given a specific hyperparameter setting θ; now we must also find the optimal hyperparameters. A common approach for selecting hyperparameters would be to optimize for cross-validated log-likelihood. Given the potential number of different smoothness hyperparameters and the computational expense of calculating $\mathbf{w}_{\text{MAP}}$, this is not feasible. We turn instead to an optimization of the (approximate) marginal likelihood, or model evidence, called empirical Bayes (Bishop, 2006).

To select between models optimized with different θ, we use a Laplace approximation to the posterior, $p(\mathbf{w}|\mathcal{D}, \theta) \approx \mathcal{N}(\mathbf{w}|\mathbf{w}_{\text{MAP}}, -H^{-1})$, to approximate the marginal likelihood as in Sahani and Linden (2003):

$$p(\mathbf{y}|X, \theta) = \frac{p(\mathbf{y}|X, \mathbf{w})\, p(\mathbf{w}|\theta)}{p(\mathbf{w}|\mathcal{D}, \theta)} \approx \frac{\exp(L) \cdot \mathcal{N}(\mathbf{w}|0, C)}{\mathcal{N}(\mathbf{w}|\mathbf{w}_{\text{MAP}}, -H^{-1})}. \qquad \text{(Equation 5)}$$

Naive optimization of θ requires a re-optimization of $\mathbf{w}$ for every change in θ, strongly restricting the dimensionality of tractable θ. Under such a constraint, the simplest approach is to reduce all σ_k to a single σ, thus assuming that all weights have the same smoothness (as done in Bak et al., 2016).

Here we use the decoupled Laplace method (Wu et al., 2017; Roy et al., 2018a) to avoid the need to re-optimize for our weight parameters after every update to our hyperparameters by making a Gaussian approximation to the likelihood of our model. This optimization is given in Algorithm 1. By circumventing the nested optimizations of θ and $\mathbf{w}$, we can consider larger sets of hyperparameters and more complex priors over our weights (e.g., σ_{day}) within minutes on a laptop (see Figure S1). In practice, we

Please cite this article in press as: Roy et al., Extracting the dynamics of behavior in sensory decision-making experiments, Neuron (2020), https://doi.org/10.1016/j.neuron.2020.12.004

also parametrize θ by fixing $\sigma_{k,t=0} = 16$, an arbitrary large value that allows the likelihood to determine the starting value of the weights $\mathbf{w}_0$ rather than forcing the weights to initialize near some predetermined value via the prior.

Algorithm 1 Optimizing hyperparameters with the decoupled Laplace approximation

Require: inputs X, choices $\mathbf{y}$

Require: initial hyperparameters θ_0, subset of hyperparameters to be optimized θ_{OPT}

1: repeat
2: Optimize for $\mathbf{w}$ given current $\theta \rightarrow \mathbf{w}_{MAP}$, Hessian of log-posterior H_θ, log-evidence E
3: Determine Gaussian prior $\mathcal{N}(0, C_\theta)$ and Laplace appx. posterior $\mathcal{N}(\mathbf{w}_{MAP}, -H_\theta^{-1})$
4: Calculate Gaussian approximation to likelihood $\mathcal{N}(\mathbf{w}_L, \Gamma)$ using product identity, where $\Gamma^{-1} = -(H_\theta + C_\theta^{-1})$ and $\mathbf{w}_L = -\Gamma H_\theta \mathbf{w}_{MAP}$
5: Optimize E w.r.t. θ_{OPT} using closed form update (with sparse operations) $\mathbf{w}_{MAP} = -H_\theta^{-1}\Gamma^{-1}\mathbf{w}_L$
6: Update best θ and corresponding best E
7: until θ converges
8: return $\mathbf{w}_{MAP}$ and θ with best E

Selection of input variables

The variables that make up the model input X are entirely user-defined. The decision as to what variables to include when modeling a particular dataset can be determined using the approximate log-evidence (log of Equation 5). The model with the highest approximate log-evidence would be considered best, though this comparison could also be swapped for a more expensive comparison of cross-validated log-likelihood (using the cross-validation procedure discussed below).

Non-identifiability is another issue that should be taken into account when selecting the variables in the model. A non-identifiability in the model occurs if one variable in X is a linear combination of some subset of other variables, in which case there are infinite weight values that all correspond to a single identical model. Fortunately, the posterior credible intervals on the weights will help indicate that a model is in a non-identifiable regime — since the weights can take a wide range of values to represent the same model, the credible intervals will be extremely large on the weights contributing to the non-identifiability. See Figure S3B for an example and further explanation.

Parameterization of input variables

It is important that the variables used in X are standardized such that the magnitudes of different weights are comparable. For categorical variables, we constrain values to be $\{-1, 0, +1\}$. For example, the previous-choice variable is coded as a -1 if the choice on the previous trial was left, $+1$ if right, and 0 if there was no choice on the previous trial (e.g., on the first trial of a session). Additionally, variables depending on the previous trial can be set to 0 if the previous trial was a mistrial. Mistrials (instances where the animal did not complete a trial, e.g., by violating a "go" cue) are otherwise omitted from the analysis. The choice bias is fixed to be a constant $+1$.

Continuous variables can be more difficult to parameterize appropriately. In the Akrami task, each of the variables for stimulus A, stimulus B, and previous stimuli are standardized such that the mean is 0 and the standard deviation is 1. The left and right contrast values used in the IBL task present a more difficult normalization problem. Suppose the contrast values were used directly. This would imply that a mouse should be twice as sensitive to a 100% contrast than to a 50% contrast. Empirically, however, this is not the case: mice tend to have little difficulty distinguishing contrasts of either value and perform comparably on these "easy" contrast values. Nonetheless, we used both contrast values as input to the same contrast weight, so the model always predicts a significant difference in behavior between 50% and 100% contrasts.

To improve accuracy of the model, we transformed the stimulus contrast values to better match the *perceived* contrasts of different stimuli, motivated by the fact that responses in the early visual system are not a linear function of physical contrast (Busse et al., 2011). For all mice, we used a fixed transformation of the contrast values used in the experiment (though this transformation could in principle be tuned separately for each mouse). The following tanh transformation of the contrasts x has a free parameter p which we set to $p = 5$ throughout the paper: $\hat{x}_p = \tanh(px)/\tanh(p)$. Specifically, this maps the contrast values from [0, 0.0625, 0.125, 0.25, 0.5, 1] to [0, 0.302, 0.555, 0.848, 0.987, 1]. See Figure S4 for a worked example of why this parametrization is useful.

More generally, this particular transformation of contrasts allows us to avoid the pressure exerted on our likelihood by the most extreme levels of perceptual evidence (e.g., 100% contrast values). In many behavior fitting procedures, this issue is avoided via the inclusion of an explicit lapse rate which discounts the influence of incorrect choices despite strong perceptual evidence. Since our model does not include explicit lapse rates, transformations to limit the influence of the strongest stimuli may be necessary for robust fitting under many psychophysical conditions (Nassar and Frank, 2016).

IBL task

Here we review the relevant features of the task and mouse training protocol from the International Brain Laboratory task (IBL task). Please refer to International Brain Laboratory et al. (2020) for further details.

Please cite this article in press as: Roy et al., Extracting the dynamics of behavior in sensory decision-making experiments, Neuron (2020), https://doi.org/10.1016/j.neuron.2020.12.004

Mice are trained to detect of a static visual grating of varying contrast (a Gabor patch) in either the left or right visual field (Figure 1A). The visual stimulus is coupled with the movements of a response wheel, and animals indicate their choices by turning the wheel left or right to bring the grating to the center of the screen (Burgess et al., 2017). The visual stimulus appears on the screen after an auditory "go cue" indicates the start of the trial and only if the animal holds the wheel for 0.2-0.5 s. Correct decisions are rewarded with sweetened water (10% sucrose solution; Guo et al., 2014), while incorrect decisions are indicated by a noise burst and are followed by a longer inter-trial interval (2 s).

Mice begin training on a "basic" version of the task, where the probability of a stimulus appearing on the left or the right is 50:50. Training begins with a set of "easy" contrasts (100% and 50%), and harder contrasts (25%, 12.5%, 6.25%, and 0%) are introduced progressively according to predefined performance criteria. After a mouse achieves a predefined performance criteria, a "biased" version of the task is introduced where the distribution of stimuli switches in blocks of trials between 20:80 favoring the right and 80:20 favoring the left. The length of each block is sampled from an exponential distribution of mean 50 trials, with a minimum block length of 20 trials and a maximum block length of 100 trials.

Akrami task

Here we review the relevant features of the task, as well as the rat and human subject training protocols. Please refer to Akrami et al. (2018) for further details.

Rats were trained on an auditory delayed comparison task, adapted from a tactile version (Fassihi et al., 2014). Training occurred within three-port operant conditioning chambers, in which ports are arranged side-by-side along one wall, with two speakers placed above the left and right nose ports. Figure 5A shows the task structure. Rat subjects initiate a trial by inserting their nose into the center port, and must keep their nose there (fixation period), until an auditory "go" cue plays. The subject can then withdraw and orient to one of the side ports in order to receive a reward of water. During the fixation period, two auditory stimuli, A and B, separated by a variable delay, are played for 400 ms, with short delay periods of 250 ms inserted before A and after B. The stimuli consist of broadband noise (2,000–20,000 Hz), generated as a series of sound pressure level (SPL) values sampled from a zero-mean normal distribution. The overall mean intensity of sounds varies from 60–92 dB. Rats should judge which out of the two stimuli, A and B, had the greater SPL standard deviation. If A > B, the correct action is to poke the nose into the right-hand nose port in order to collect the reward, and if A < B, rats should orient to the left-hand nose port.

Trial durations are independently varied on a trial-by-trial basis, by varying the delay interval between the two stimuli, which can be as short as 2 s or as long as 6 s. Rats progressed through a series of shaping stages before the final version of the delayed comparison task, in which they learned to: associate light in the center poke with the availability of trials; associate special sounds from the side pokes with reward; maintain their nose in the center poke until they hear an auditory "go" signal; and compare the A and B stimuli. Training began when rats were two months old, and typically required three to four months of training to display stable performance on the complete version of the task.

In the human version of the task, similar auditory stimuli to those used for rats were used (see Figure 7A). Subjects received, in each trial, a pair of sounds played from ear-surrounding noise-cancelling headphones. The subject self-initiated each trial by pressing the space bar on the keyboard. Stimulus A was then presented together with a green square on the left side of a computer monitor in front of the subject. This was followed by a delay period, indicated by "WAIT!" on the screen, then B was presented together with a red square on the right side of the screen. At the end of the second stimulus and after the go cue, subjects were required to compare the two sounds and decide which one was louder, then indicate their choice by pressing the "k" key with their right hand (B was louder) or the "s" key with their left hand (A was louder). Written feedback about the correctness of their response was provided on the screen, for individual trials as well as the average performance updated every ten trials.

A practical guide

In order to facilitate easy application of our model to new datasets, we have assembled a list of practical considerations. Many of these have already been addressed in the main text and in the STAR methods, but we will provide a comprehensive list here for easy access. We divide these considerations into three sections. First, *model specification* for considerations that arise before attempting to apply PsyTrack to new behavioral data. Second, *fitting and model-selection* for considerations that occur when trying to obtain the best possible characterization of behavior with the model. And finally, *post-modeling analysis* which is concerned with the interpretation and validation of the model results.

Model specification

Is the behavior appropriately described by a binary choice variable?

- The current method cannot handle more than two choices, though the extension to the multinomial setting is an exciting future direction that has already received some attention (Bak and Pillow, 2018).
- Early training in many tasks often results in a high proportion of "mistrials," i.e., trials where the animal responded in an invalid (rather than incorrect) way. These mistrials are omitted in our model (though they can indirectly influence valid trials via history regressors).

Please cite this article in press as: Roy et al., Extracting the dynamics of behavior in sensory decision-making experiments, Neuron (2020), https://doi.org/10.1016/j.neuron.2020.12.004

What sort of variability is expected from the behavior?

- If there is reason to believe that behavior is static (e.g., human behavior or behavior from the end of training), then PsyTrack may not offer anything beyond standard logistic regression. However, one major benefit is that PsyTrack will *infer* static behavior rather than *assume* it (see Figure 7).
- If behavior is expected to "jump" at known points in training (e.g., in-between sessions), consider using σ_{day} to accommodate that behavior (see Figures 2C and S6 for creative usage of σ_{day}).
- If behavior is expected to "jump," but it's uncertain when those sudden changes might occur, be aware of how the model makes a smooth approximation to a sudden change (Figure S2).

What input variables should be tried?

- Relevant task stimuli should certainly be included. Choice bias is also a good bet.
- The irrelevant input variables can be harder to settle upon. It is often worth a preliminary analyses of the empirical data to see if certain dependencies are clearly present (e.g., Figures 5C–5E verifies a behavioral dependence on the choice and correct answer of the previous trial).

Fitting and model-selection

Which candidate input variables should ultimately be chosen?

- For models with the same number of hyperparameters, we use the model evidence (Equation 5) to perform model selection.
- For models with different numbers of hyperparameters, we can use the Bayesian Information Criterion (BIC), given by $-2E + K\log T$, where E is the log-evidence, K is the number of hyperparameters, and T is the number of trials. Select the model with lowest BIC.
- Users may also perform model selection via cross-validation. This involves dividing the data into a training and test set, then selecting the model with highest cross-validated log-likelihood on the test set. This approach is more principled than the BIC but computationally more expensive.

How should input variables be parametrized?

- The default option is consolidated to $\{-1, 0, +1\}$ for discrete variables and to standardize continuous variables (subtract mean and divide by standard deviation).
- If weights ever grow to be greater than 5 (or less than -5), this is generally a red flag. This often means that the corresponding variable should be reparametrized to discount the effect of extreme evidence (see Figure S4 and STAR methods).
- Example alternate parametrizations in the current work include: (i) combining the left and right contrast weights in the IBL task into a single contrast weight (where left contrasts are encoded as negative values); (ii) increasing the number of history regressors in the Akrami task to capture more specific effects (e.g., a weight to capture the specific impact of previously choosing rightward on a leftward trial).
- Again, model evidence (Equation 5) can be used to decide between different parametrizations.

Over what time frame should the model be applied?

- While the weights of the model vary smoothly, the hyperparameters do not. For example, if you apply the model to two consecutive sessions where a particular weight is static for the first session then highly variable for the second session, the model will settle on a compromise value for that weight's σ. It may be preferred in such a situation to apply the model to each session separately.
- If you are looking for fine-scale fluctuations (e.g., bias weight fluctuations in response to bias blocks in Figure 4), it is often better to look at shorter time frames, down to a single session.

Post-modeling analysis

How should the weights of a model be interpreted?

- The credible intervals around the weights can help gauge the model's uncertainty about that weight's value at a particular point in training (and also alert you to non-identifiability in your input variables, see Figure S3).
- Keep in mind that the weights trajectories are conditioned on a point estimate (the MAP estimate) of the smoothness hyperparameters. Uncertainty in the σ values can also be calculated (Figures 2B and 2D).
- Remember that the model is not causal: information from future trials is free to affect the weight values at earlier points in training (see Figure S6).
- PsyTrack is a descriptive, not normative, model of behavior. There is no notion of what an animal *ought* to be doing; in fact, task rewards do not play a part in the model inference at all (though previous rewards could be included as an optional input variable).

How should the results of a model be validated?

- The best way to validate a model is to compare model predictions (optimally using cross-validated weights) against empirical measure direct from the choice behavior.
- For example, Figures 5C–5E validates that the added history regressors are indeed capturing a dependence on the previous trial (especially when compared to Figures S7C–S7E). Figure S5 validates the predictions of the model against to the standard psychometric curve.

Please cite this article in press as: Roy et al., Extracting the dynamics of behavior in sensory decision-making experiments, Neuron (2020), https://doi.org/10.1016/j.neuron.2020.12.004

Quantification and statistical analysis

Cross-validation procedure

When making predictions about specific trials, the model should not be trained using those trials. We implemented a 10-fold cross-validation procedure where we fit the model using a training set composed of a random 90% of trials at a time, and used the remaining 10% for predicting and testing. We modified the prior Σ such that the gaps created by removing the 10% of test set trials were taken into account. For example, if trial t is in the test set and trials $t-1$ and $t+1$ are in the training set, then we modify the value on the diagonal of Σ corresponding to trial $t-1$ from σ^2 to $2\sigma^2$ to account for the missing entry in Σ created by omitting trial t from the training set.

To predict the animal's choice at a test trial t, we first inferred the weights for a training set of trials that excludes t, as described above. Then we approximated $\mathbf{w}_t$ by linearly interpolating from the nearest adjacent trials in the training set. We repeated this to obtain a set of *predicted weights* for each trial. Using the predicted weights $\mathbf{w}_t$ and the input vector $\mathbf{x}_t$ for that trial, we calculated a predicted choice probability $P(\text{Go Right})$ (as in Figure 5), and compared this to the actual choice y_t to calculate the predicted accuracy (as in Figure 8). The cross-validated log-likelihood L, calculated using predicted weights, can be used to choose between models in lieu of approximate log-evidence.

Calculation of posterior credible intervals

In order to estimate the extent to which our recovered weights $\mathbf{w}$ are constrained by the data, we calculated a posterior credible interval over the time-varying weight trajectories (e.g., the shaded regions shown in Figure 5B). Specifically, we approximated the 95% posterior credible interval by using 1.96 standard deviations. The standard deviation is calculated by taking the square-root of the diagonal of the covariance matrix at $\mathbf{w}_{\text{MAP}}$.

The covariance matrix can be approximated by the inverse Hessian, but inversion of a large matrix can be challenging. Here, we adapted a fast method for inverting block-tridiagonal matrices (Rybicki and Hummer, 1991), taking advantage of the fact that our Hessian (while extremely large) has a block-tridiagonal structure, and that we only need the diagonal of the inverse Hessian. If H is the Hessian matrix of our weights at the posterior peak $\mathbf{w}_{\text{MAP}}$ (also with the highest log-evidence, according to the optimization procedure in Algorithm 1), this method calculates a diagonal matrix,

$$A = \text{diag}\left(H^{-1}\right), \qquad \text{(Equation 6)}$$

such that we can take the square root of the diagonal elements of A to estimate one standard deviation for the corresponding weight on each trial. The algorithm requires order TK^3 scalar operations for calculating the central blocks of our inverse Hessian (for T trials and K weights).

To calculate the posterior credible intervals for our hyperparameters θ, we took the same approach of inverting the $K \times K$ hyperparameter Hessian matrix (or $2K \times 2K$ if σ_{day} is used). The difficulty here is not in inverting this Hessian matrix, since it is much smaller, but in calculating the Hessian in the first place. We calculated each entry in our Hessian numerically, determining each of the $K(K+1)/2$ unique entries with finite differencing. Once we have determined the Hessian matrix for the hyperparameters, it is straightforward to calculate the 95% posterior credible intervals for each hyperparameter using the square root of the diagonal elements of the inverse Hessian, the same procedure as in the case for the weights.

Neuron, Volume *109*

Supplemental Information

Extracting the dynamics of behavior in sensory decision-making experiments

Nicholas A. Roy, Ji Hyun Bak, The International Brain Laboratory, Athena Akrami, Carlos D. Brody, and Jonathan W. Pillow

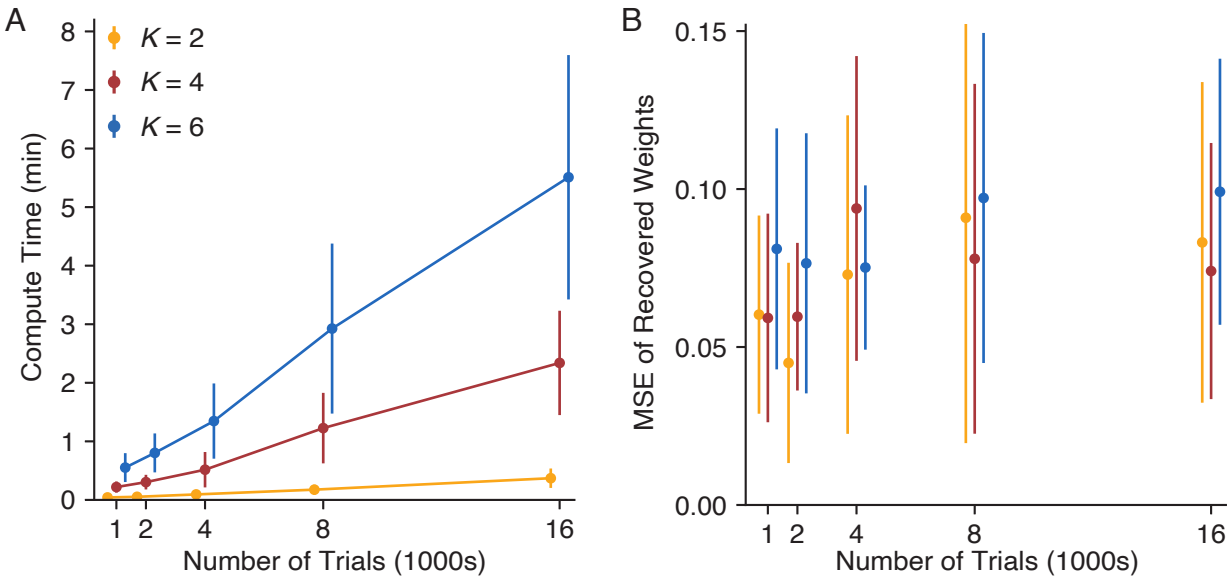


Figure S1. Related to Figure 2, Compute time and model accuracy

(A) The model fitting time as a function of the number of weights $K = \{2, 4, 6\}$ and the number of trials $T = \{1000, 200, 4000, 8000, 16000\}$. Weights are simulated as in Figure 2A. For each pair of K weights and T trials, 20 sets of weights are randomly simulated (with each $\log_2(\sigma_k) \sim \mathbb{U}(-7.5, -3.5)$; no σ_{day} was used), and recovered by the model (the calculation of credible intervals on the weights was omitted). We plot the average recovery time across the 20 iterations, ± 1 standard deviation. We can see that even a reasonably complex model (6 weights and 16000 trials) only takes around 5 minutes to fit on average. All models were fit on a 2012 MacBook Pro with a 2.3 GHz Quad-Core Intel i7 processor.

(B) The mean-squared error (MSE) calculated across all weights across all trials (between the true weights and recovered weights), as a function of the total number of weights and total number of trials in the model. Calculations used the same models run in (A). The average MSE across the 20 iterations of each model is plotted, ± 1 standard deviation. We can see that the recovery of the weights is relatively independent of the number of weights (color coded as in (A)) and number of trials.

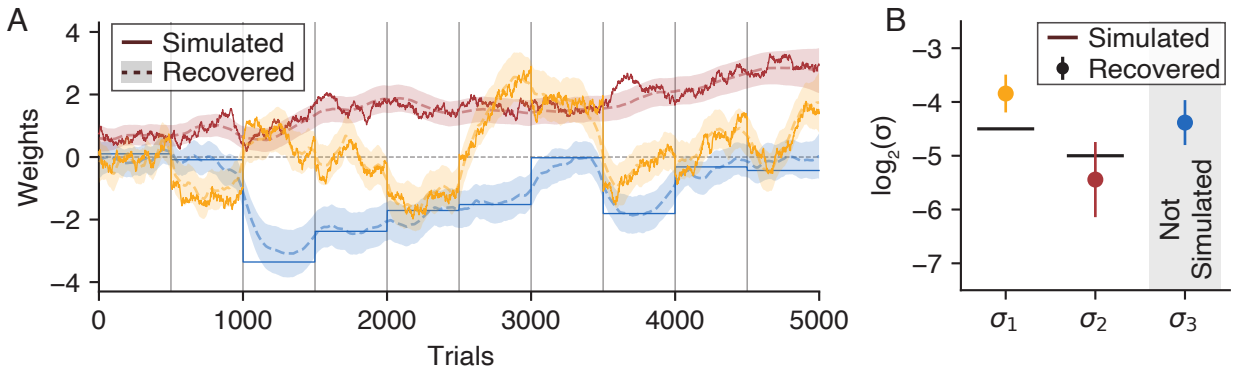


Figure S2. Related to Figure 2, Recovering sudden changes in behavior with smooth weight trajectories
(A) To examine how the model handles "jumps" in behavior (that do not occur at session boundaries), we attempt to recover the same simulated weights from Figure 2C without using any σ_{day} hyperparameters. While our model is no longer able to capture the explicit "jumps" in the true weights (solid lines, yellow and blue simulated with a σ_{day} hyperparameter), the recovered weights (dashed lines) smoothly track these sudden shifts as a sigmoid (e.g., the blue weight at trial 1000).
(B) Since the recovered σ hyperparameters also have to account for variability in the weight due to the simulated σ_{day} hyperparameters, the recovered σ values tend to be larger than the simulated σ values. The σ for the yellow weight (simulated with both σ and σ_{day}) does indeed overshoot it's true value, while the σ for the red weight (simulated with only σ) is accurately recovered. The blue weight was not simulated with σ at all, so the recovered σ is capturing variability exclusively due to the simulated σ_{day}.

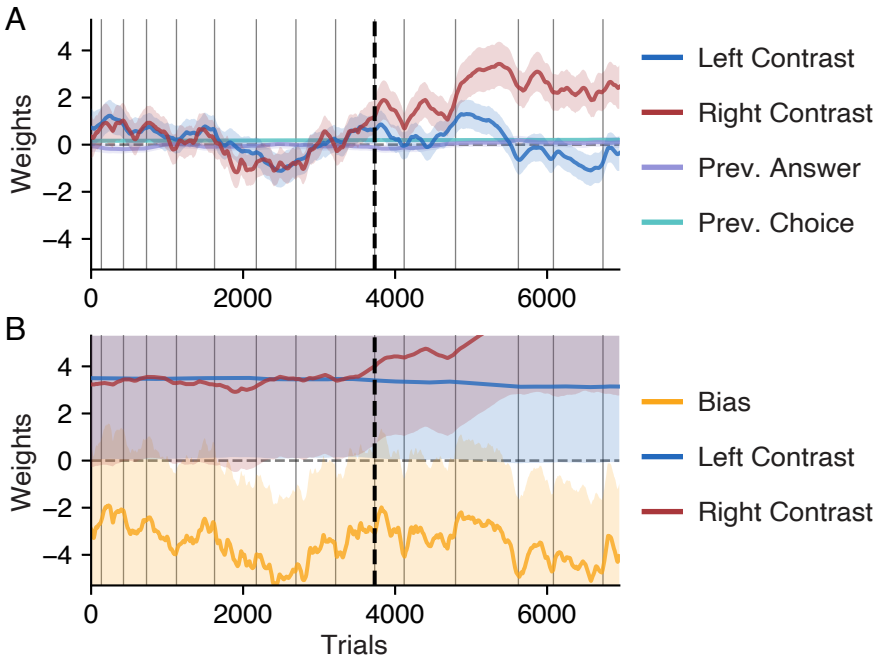


Figure S3. Related to Figure 3, Adding weights to early training sessions in IBL mice

(A) Here we refit the data first presented in Figure 3B, but now adding two additional weights: a Previous (Correct) Answer and Previous Choice weight. While these history weights massively impact choice behavior during early training in all the Akrami Rats (see Figure 6D-E), they have negligible impact on the behavior of our example IBL mouse.

(B) Here we refit the data first presented in Figure 3B, but now adding a bias weight. While a bias weight is very useful for describing the behavior of IBL mice later in training (see Figure 4), we omit the bias weight during early training due to an issue of non-identifiability. During the initial sessions of training, IBL mice are only presented with "easy" contrast values of 50% and 100%. These contrasts are perceptually very similar (i.e. a 100% contrast is not twice as difficult as a 50% contrast), which we account for with a tanh transformation of the contrasts (see Figure S4). Thus, the task in the earliest stages of training has effectively only two types of trials: $\sim$100% left contrast trials and $\sim$100% right contrast trials. With a task this simple, behavior is over-parameterized by having a bias weight and two contrast weights, introducing the non-identifiability. Explicitly, a model with hypothetical weight values of [bias, left, right] = [0, -1, +1] is nearly identical to a model with values [-1, 0, 2]; in fact, there are an infinite number of weight values for these three weights that would all describe behavior on this simplified task in approximately the same way. Fortunately, the posterior credible intervals on the weights can indicate that a model is in a non-identifiable regime. Because there are so many settings of possible weight value, the intervals become abnormally large and overlapping, as shown here. See the STAR Methods for more details about model non-identifiability.

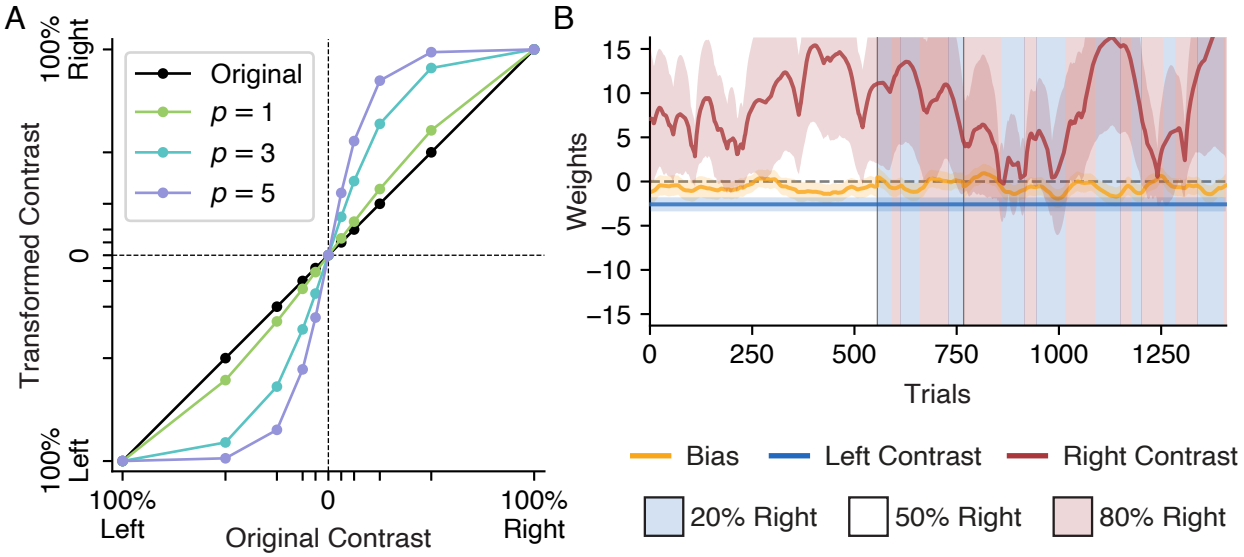


Figure S4. Related to Figure 4, The impact of the tanh transformation of IBL contrasts on model weights
(A) The effect of the tanh transformation on the IBL contrast values for several settings of the free parameter p. A tanh transformation is applied to the contrast values, c, in the IBL task such that the relative values of the transformed contrasts, $\hat{c}_p = \tanh(pc)/\tanh(p)$, better aligns with their relative perceptual difficulty. In this work, we use $p = 5$ to transform the contrasts (purple line), such that $c = \{0, 0.0625, 0.125, 0.25, 0.5, 1\}$ are transformed to $\hat{c}_5 = \{0, 0.303, 0.555, 0.848, 0.987, 1\}$ (left contrasts are coded as having negative value here). This value was anecdotally observed to work well across a large variety of mice and sessions, but it could be optimized for each model.
(B) Here we refit the data first presented in Figure 4B, forgoing the tanh transformation and using the original contrast values c instead. We see that the right contrast weight grows to massive values and fluctuates wildly. Using Equation 1, we can calculate that with a weight of 15, the model predicts that a 100% right contrast would result in a P(Go Right) of over 99.9999% (disregarding the impact of the much smaller bias weight, for simplicity). This is an absurdly confident prediction, even for the best trained mouse, but it represents a compromise the model was forced to make. We can calculate that the predicted P(Go Right) on the most difficult right contrast value, a 6.25% contrast, is a much more reasonable 71.9%. A well-trained mouse could certainly be performing better than this on the hardest contrast. However, in order to reflect a higher P(Go Right) on this hard contrast, the right contrast weight would need to become even greater, resulting in even more absurdly confident predictions on the 100% contrast trials. All right contrast values share the same weight, forcing a single compromise weight value between them. By using the raw contrast value, the model assumes that a 100% right contrast is 16 times more salient to the mouse than a 6.25% contrast, which is empirically untrue (Busse et al., 2011). Applying the tanh transformation brings the relative value of the two contrasts into a much more reasonable regime. With $p = 5$, a 6.25% contrast is encoded as 0.303 while the 100% contrast remains at 1.0, making the 100% contrast only 3 time more salient than the 6.25% contrast. See the STAR Methods for more details about the parametrization of input variables.

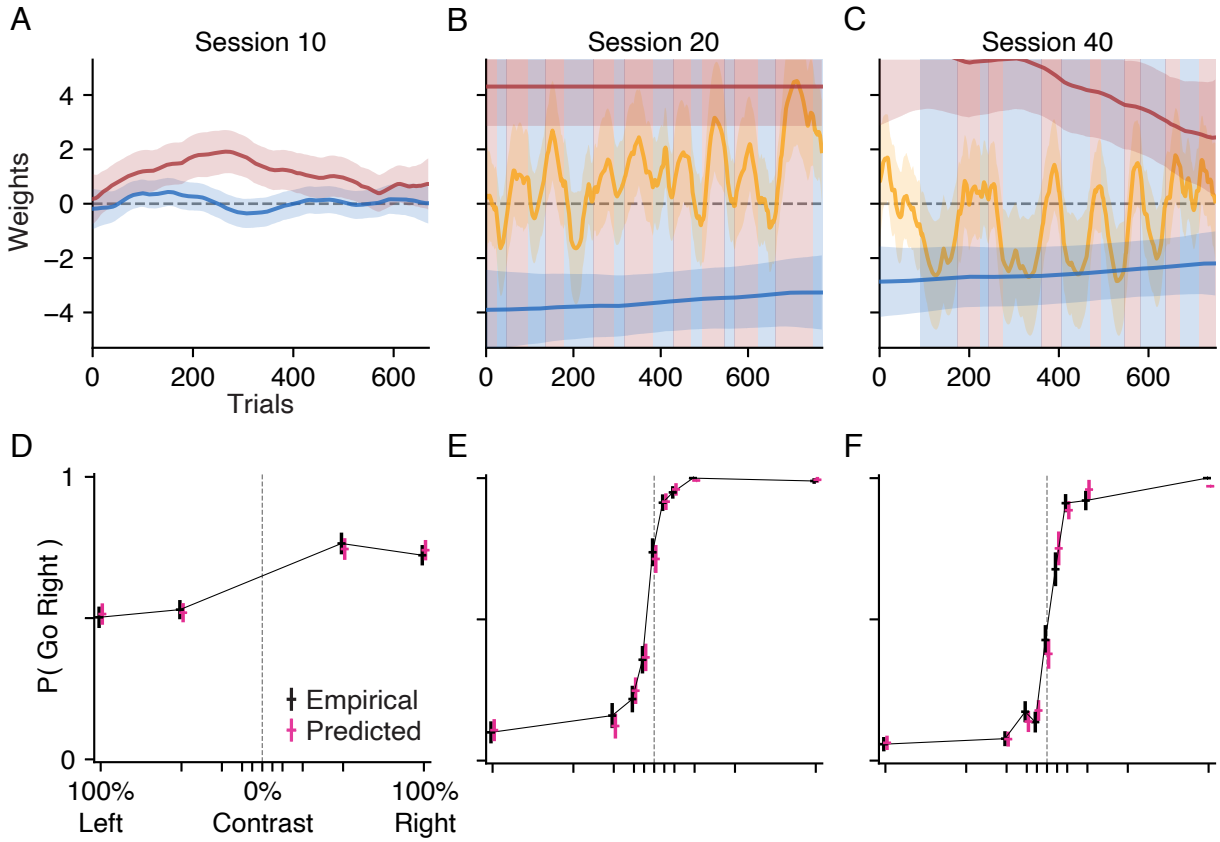


Figure S5. Related to Figure 4, Validating the model with a comparison to empirical psychometric curves
(A-C) In order to validate that our model is accurately characterizing choice behavior, we can compare a psychometric curve generated directly from the primary behavioral data against a curve generated from the (cross-validated) weights of our model. Here, we show a sample of three sessions from our example IBL mouse, taken from distinct periods in training (see Figure 4A). We use our cross-validation procedure to infer weights for each session.
(D-F) For each of the three example sessions, we show both the empirical (black) and model predicted (pink) psychometric curve (each with ±1SE). That is, for each unique stimulus used during the session (sessions early in training operate with a restricted "easy" stimulus set) we calculate the percent of trials where the animal went right in response to that stimulus. Similarly, for each trial we can derive from our model a cross-validated P(Go Right) and average those values according to the stimulus. We can see that for each of the example sessions there is agreement between the empirical data and the model's predictions.

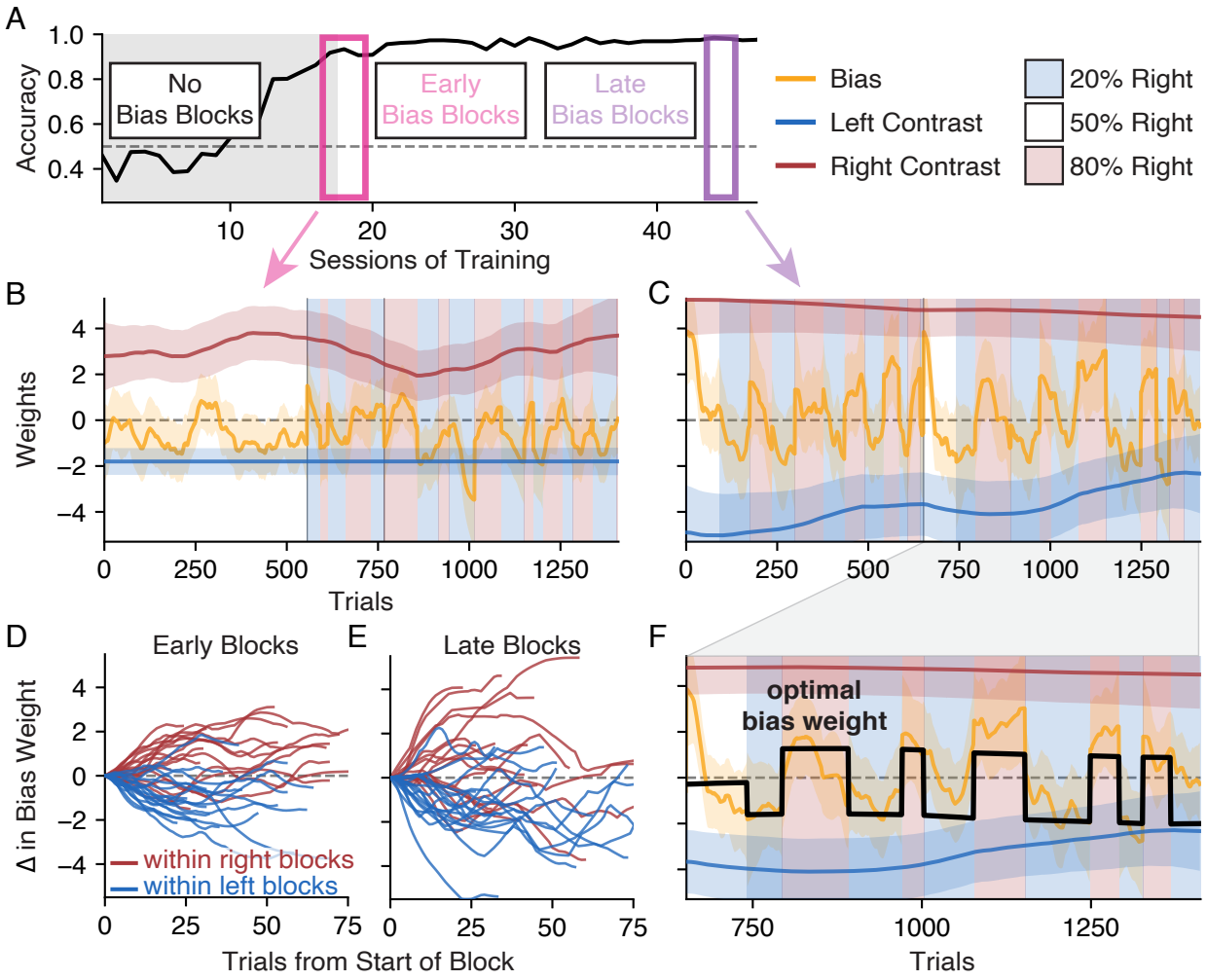


Figure S6. Related to Figure 4, Allowing the bias weight to reset between bias blocks with σ_{day}

When inferring weights with our model, it is important to remember that the value of a weight at a particular trial t is a function of not just previous trials, but also future trials. Our analyses in Figure 4 could be misinterpreted as showing that the mouse *anticipates* the start of a new block since the bias weight will often reverse direction before the end of the current block. This apparent anticipation is confounded by the smoothing of our model, which is especially dramatic around "jumps" in a weight (see Figure S2). Fortunately, we can adapt our model to remove this confound with two steps. First, we treat the boundary between each bias block as a session boundary such that the size of the jump at each boundary is parametrized not by σ, but by σ_{day} (for the bias weight only). Second, instead of inferring the value of our σ_{day} hyperparameter as we have done throughout the paper, we fix the value of σ_{day} to be very high ($\sigma_{\text{day}} = 2^5$). Effectively, this means that the bias weight is free to "reset" itself to any value at the start of each new bias block. This prevents blocks ahead of (or behind) the current block from influencing the values of the bias weight within that block. Figure 4 has been recreated here with this new adaptation to the model: the sharp transitions between blocks are now apparent in (B) and (C), but the qualitative results shown in (D-F) remain largely unaffected.

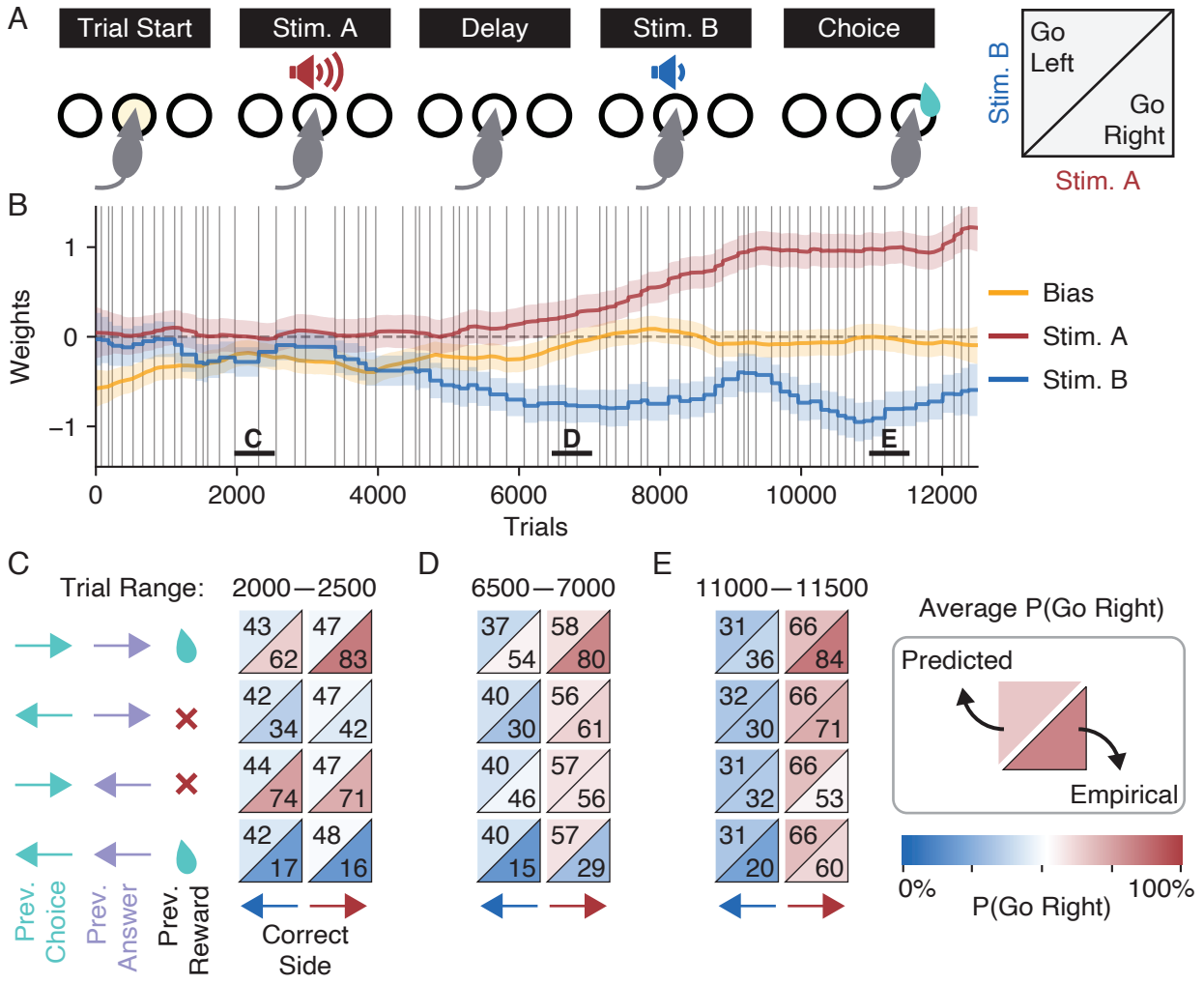


Figure S7. Related to Figure 5, Example Akrami rat without history regressors
In order to better understand the contribution of the three history regressors to our model, we reproduce the analyses of Figure 5 with a model that omits the history regressors. In (B), we see that the trajectories of the Stim. A, Stim. B, and Bias weights remain qualitatively the same. However, (C-E) shows that this model does not capture the dependence on the previous trial that is clearly reflected in the empirical choice behavior, especially early in training.

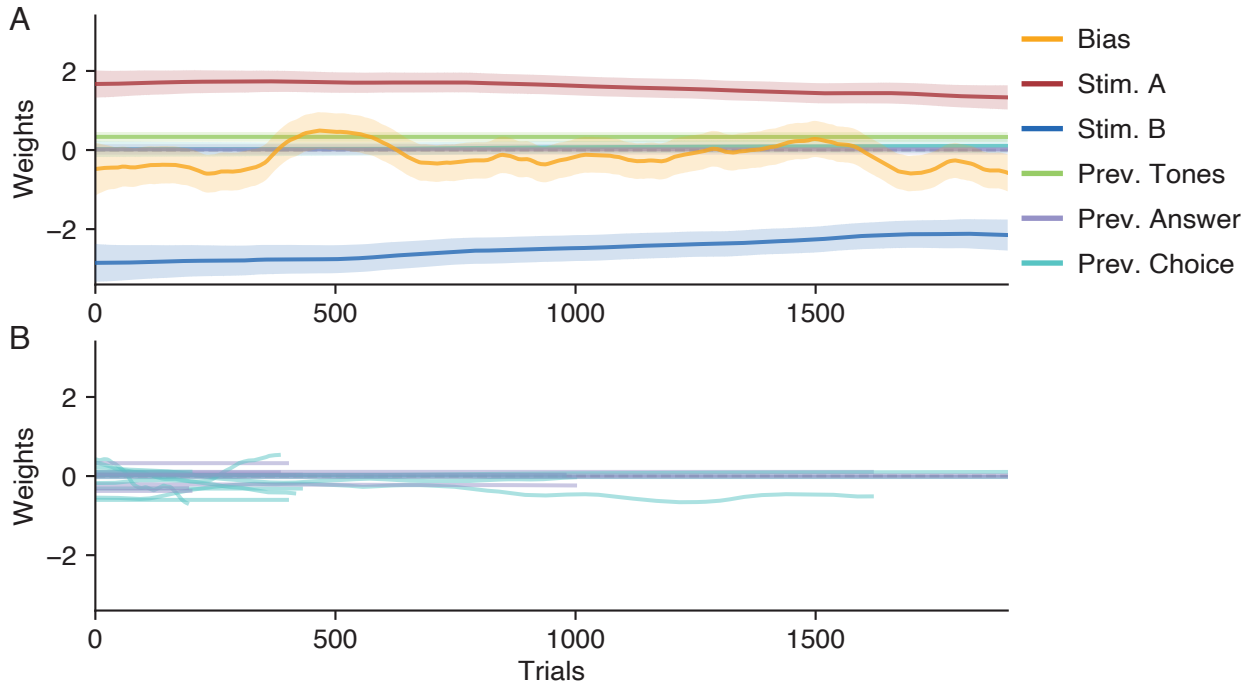


Figure S8. Related to Figure 7, Modeling the Akrami human subjects with the Previous Choice and Previous Answer weights

(A) Here we refit the data presented in Figure 7B, but now we add both a Previous Choice and a Previous (Correct) Answer weight. Since human subjects are given task instructions before starting, we would not expect their behavior to be affected by either their choice or the correct answer on the previous trial. Indeed, we see that both new history weights are always approximately 0 for our example subject.

(B) Here we refit the data presented in Figure 7C, but now we add both a Previous Choice and a Previous (Correct) Answer weight for all the human subjects. We plot only the two new weights from each refit model, for clarity. We see that, in general, both the choice and the correct answer on the previous trial have a relatively small impact on human choice behavior.