Inhibitory and excitatory populations in parietal cortex are equally selective for decision outcome in both novices and experts

Farzaneh Najafi[1], Gamaleldin F Elsayed[2], Eftychios Pnevmatikakis[3], John P Cunningham[4], Anne K Churchland[1]

[1]Cold Spring Harbor Laboratory, Cold Spring Harbor, NY
[2]Columbia University, Neuroscience, New York, NY,
[3]Flatiron Institute, New York, NY,
[4]Columbia University, Statistics, New York, NY

Abstract

Decisions are driven by the coordinated activity of diverse neural populations. Inhibitory neurons play a critical role in decision-making models, but technical challenges have left untested their role in vivo, both in novice and expert decision-makers. To understand the contribution of excitatory and inhibitory neurons to decision-making, we simultaneously measured their activity in mice judging the repetition rate of multisensory pulses. Individual inhibitory neurons were slightly more selective and more strongly correlated than excitatory neurons. Further, inhibitory ensemble activity could be decoded to predict trial-by-trial choice with similar accuracy as excitatory ensembles. Finally, population activity in both cell types changed in parallel as mice transitioned from novice to expert decision-makers: population activity gradually became more choice-selective and prompt. The observations advocate for models in which excitatory and inhibitory connectivity (1) preserves choice selectivity in both populations and (2) is fine-tuned with experience to support expert decision-making.

Main

Theoretical models have been influential in interpreting behavioral and neural measurements during perceptual decision-making[1-4]. Models that incorporate inhibitory neurons are widely accepted[1,5], but key aspects of model architecture and connectivity remain untested. In some architectures, inhibitory neurons are part of a single pool that is broadly innervated by excitatory neurons (Fig. 1a top[1,4,6]). In alternative architectures, inhibitory neurons are selective, because of targeted connectivity with excitatory neurons (Fig. 1a bottom). Targeted connectivity is supported on theoretical grounds because of stability[2,7], and has experimental support in the goldfish oculomotor integrator[8]. However, such models have not been tested in decision-making due to the challenges in identifying inhibitory neurons reliably and in large numbers.

Outside of decision-making, the selectivity and connectivity of inhibitory neurons is well studied. Excitatory neurons in most cases are sharply tuned to visual stimulus features[9-13], reflecting their specific and non-random connectivity[12-17]. Inhibitory neurons are likewise tuned to stimulus features, though their tuning is often reported as broader[10-12,18-21] (but see [22]). A growing body of evidence suggests that the tuning in inhibitory neurons arises from strong connectivity with excitatory neurons tuned for the same stimulus[7,23,24].

The applicability of these findings for decision-making is unclear: the experiments were carried out mainly in V1, largely in passively viewing or anaesthetized animals. In decision-making, by contrast, areas beyond V1 are recruited, including the posterior parietal cortex (PPC)[25-27]. Further, decisions require the animal to learn, as novices, the abstract relationship between a sensory stimulus and a motor response, and then, as experts, report binary choices reliably. The computations needed to fulfill these requirements could rely on quite different circuits from those that are activated during passive viewing.

Here, we aimed to compare the responses of excitatory and inhibitory neurons during decision-making. We demonstrate that inhibitory neurons are selective for the animal's choice, both at the single-neuron and population level. These results argue that in decision structures, as in V1, there are subnetworks of neurons with similar response properties, conferring network stability and robustness.

Simultaneous imaging of excitatory and inhibitory neurons during decision-making

To test how excitatory and inhibitory neurons coordinate during decision-making, we measured neural activity in transgenic mice. First, we trained mice to report decisions about the repetition rate of a sequence of multisensory events by licking to a left or right waterspout (Fig. 1b; Extended Data Fig. 1a). Trials consisted of a series of auditory clicks and visual flashes, simultaneously presented at a rate that fluctuated stochastically over a 1000 ms period[28,29]. Mice reported whether event rates were high or low compared to an abstract category boundary (16 Hz) that they learned with experience. Decisions depended strongly on stimulus rate: performance was at chance when the stimulus rate was at the category boundary, and was higher at rates further from the category boundary (Fig. 1c). A logistic regression model demonstrated that choice depends on the current evidence strength, previous choice outcome[30], and the time passed since the previous trial (Extended Data Fig. 1b). We imaged excitatory and inhibitory neural activity by injecting a viral vector containing the calcium indicator GCaMP6f to layer 2/3 of mouse Posterior Parietal Cortex (PPC; 2mm posterior to Bregma, 1.7mm lateral to midline[25,26,30-33]). Mice expressed the red fluorescent protein tdTomato transgenically in all GABAergic inhibitory neurons. We used a two-channel two-photon microscope to record the activity of all neurons, a subset of which were identified as inhibitory neurons (Fig. 1d). This allowed us to measure the activity of excitatory and inhibitory populations in the same animal.

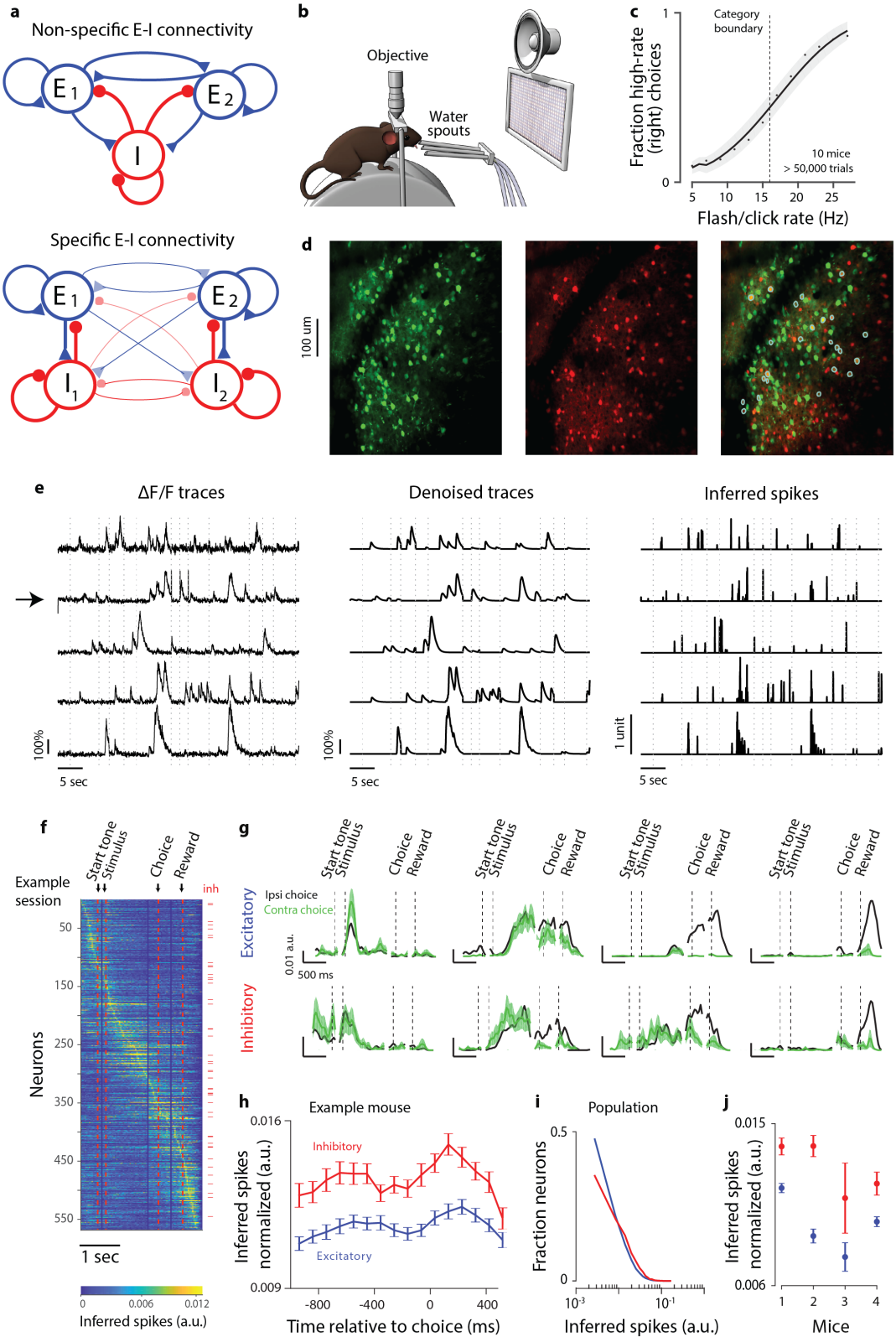


Figure 1. Simultaneous imaging of inhibitory and excitatory populations during decision-making constrains decision-making models.
a. Schematic of candidate decision-making model architectures. E1 and E2 represent pools of excitatory neurons, each favoring a different choice. **Top:** both pools excite a single pool of non-selective inhibitory neurons (I), which, in turn, provides inhibition to both excitatory pools. **Bottom:** excitatory pools (E_1 and E_2) target specific pools of inhibitory neurons (I_1 and I_2). Thin lines indicate weaker connections. In both cases, recurrent excitation drives persistent activity in the network, and inhibition allows for competition between the two choices[1,2]. **b,** Behavioral apparatus in which a head-fixed mouse is atop a cylindrical wheel. Visual display and speaker present the multisensory stimulus. To initiate a trial, mice licked the middle waterspout. To report the decision about the stimulus rate, mice licked left/right spouts. Objective belongs to the 2-photon microscope used to image neural activity through a window implanted in the skull. **c.** Psychometric function showing the fraction of trials in which the mouse judged the stimulus as high rate as a function of stimulus rate. Dots: data; Line: GLM fit; mean+/-standard deviation across 10 mice. Dashed vertical line: category boundary (16Hz). **d,** Average projection image of 10,000 frames. **Left:** green channel showing GCaMP6f expression. **Middle:** red channel showing tdTomato expression. **Right:** merge of middle and right. Cyan circles indicate GCaMP6f-expressing neurons that were identified as inhibitory. **e,** 5 example neurons identified by the CNMF algorithm (arrow: inhibitory neuron). **Left:** raw ΔF/F traces. **Middle:** de-noised traces. **Right:** inferred spiking activity; normalized so that 1 unit corresponds to 100% DF/F. Imaging was not performed during inter-trial intervals; traces from 13 consecutive trials are concatenated; dashed lines: trial onsets. **f,** Example session with 568 neurons. Each row shows the trial-averaged inferred spiking activity of a neuron (frame resolution: 32.4ms). Neurons are sorted according to the timing of their peak activity. Inhibitory neurons (n=45) are indicated by red ticks on the right. Red vertical lines mark trial events: initiation tone, stimulus onset, choice, and reward. Duration between events varied across trials, so traces were aligned for each trial event, and then concatenated. Vertical blue lines are due to concatenation. **g,** Trial-averaged traces of 4 excitatory (top) and 4 inhibitory (bottom) neurons, for ipsi (black) and contra-lateral (green) choices (mean +/- standard error; ~250 trials per session). **h,** Inferred spiking activity for excitatory (blue) and inhibitory (red) neurons during the course of a trial. Example mouse; mean +/- standard error across days (n=46). Inferred spiking activity was significantly higher for inhibitory neurons (t-test; $p<0.001$) at all times. **i,** Distribution of inferred spiking activity at the time bin before the choice for all mice and all sessions (bin size=97ms; 41,723 excitatory and 5,142 inhibitory neurons). **j,** Inferred spiking activity at the time bin before the choice for each individual mouse, for excitatory vs. inhibitory neurons (mean +/- standard error across days). Differences were significant for all subjects (t-test; $p<0.001$). In (**f-j**), inferred spiking activity of each neuron is normalized to its max spiking activity (Methods).

To detect neurons and extract calcium signals from imaging data, we leveraged an algorithm that simultaneously identifies neurons, de-noises the fluorescence signal and de-mixes signals from spatially overlapping components[34,35] (Fig. 1e middle). The algorithm also estimates spiking activity for each neuron (Fig. 1e right). We refer to this as "inferred spiking activity", for simplicity, acknowledging that estimating spikes from calcium signals is challenging[36]. Analyses were performed on inferred spiking activity. To identify inhibitory neurons, we first corrected for bleed-through from the green to the red channel. Next, we identified a subset of GCaMP6f-expressing neurons as inhibitory neurons based on the signal intensity on the red channel as well as the spatial correlation between red and green channels (Fig. 1d right, cyan circles). Inhibitory neurons constituted ~10% of the population, close to previous reports[37].

Confirming previous reports[31,32,38], we observed that the activity of individual neurons peaked at time points that spanned the trial (Fig. 1f,g). Diverse temporal dynamics were evident in both cell types (Fig. 1f,g). The inferred spiking activity was significantly higher for inhibitory

compared to excitatory neurons throughout the trial, as expected[10,39] (Fig. 1h; t-test, $p<0.001$). In the moments before the choice (97.1ms, average of 3 frames), this difference was clear (Fig. 1i) and significant for all mice (Fig. 1j). The probable differences in GCaMP expression levels and calcium buffering between excitatory and inhibitory neurons make a direct estimate of the underlying firing rates difficult[39]; however, the significant difference in the inferred spiking activity between excitatory and inhibitory neurons provides further evidence that we successfully identified two separate neural populations.

Individual inhibitory neurons are selective for the animal's choice

To assess the selectivity of individual excitatory and inhibitory neurons for the decision outcome, we performed receiver operating characteristic (ROC) analysis[40] on single-neuron responses. For each neuron, at each time point, we calculated the area under ROC curve (AUC) as a measure of the amount of overlap between the response distributions for ipsilateral vs. contralateral choices. A neuron was identified as "choice-selective" if its AUC value was significantly different from a constructed shuffled distribution (Extended Data Fig 2a; Methods), indicating that the neural activity was significantly different for ipsi- vs. contralateral choices (Fig. 2a, shaded areas mark choice-selective neurons).

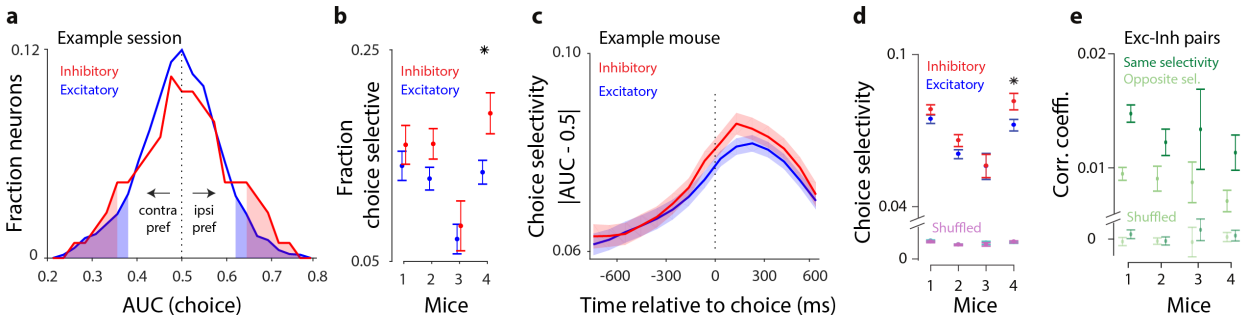


Figure 2. Single-cell and pairwise analyses argue for non-random connections between excitatory and inhibitory neurons.

a-d, Ideal observer analysis reveals the ability of individual neurons to distinguish left vs. right choices. In all panels, blue and red indicate excitatory and inhibitory neurons, respectively. **a,** Distribution of AUC values (area under the curve) of an ROC analysis for distinguishing choice from the activity of single neurons in an example session. Data correspond to the 97 ms window preceding the choice for 285 excitatory and 29 inhibitory neurons. Values larger (smaller) than 0.5 indicate neurons preferring the ipsi- (contra-) lateral choice. Dashed line at 0.5 indicates that neural activity was not distinguishable for left vs. right choices. Shaded areas mark significant AUC values (compared to a shuffle distribution). **b,** Fraction of excitatory and inhibitory neurons that are significantly choice-selective for each mouse: mean+/-standard error across days (n = 45, 48, 7, 35, per mouse). Star (*) indicates significant difference between excitatory and inhibitory neurons (t-test; $p<0.05$). **c,** ROC analysis performed on 97ms non-overlapping time windows. Time course of choice selectivity (absolute deviation of AUC from chance) shown for excitatory and inhibitory neurons in an example mouse; mean+/-standard error across days, n=45. **d,** Average choice selectivity of excitatory and inhibitory neurons during 97-0 ms before the choice is summarized for each mouse; mean+/-standard error across days. **e,** Pearson's correlation coefficients shown for pairs of excitatory-inhibitory neurons with the same choice selectivity (dark green) or opposite choice selectivity (light green, i.e. one neuron prefers ipsi, and the other neuron prefers contralateral choice). "Shuffled" in (d,e) denotes quantities were computed using shuffled trial labels. Mean+/-standard error across days; 97-0 ms before the choice.

Our analysis of choice selectivity in both cell types revealed that 13% of individual neurons were significantly choice selective for either ipsilateral or contralateral choices (Fig. 2b). Choice selectivity values computed on correct vs. error trials were positively correlated (Extended Data Fig 2c,d). Positive correlations indicate that the majority of neurons reflect the impending choice, as opposed to the sensory stimulus that informed it (Methods). Variability across mice in the strength of this correlation may indicate that the balance of sensory vs. choice signals within individual neurons varied across subjects (perhaps due to imaged subregions within the window, Extended Data Fig. 2e). Finally, we found that choice selectivity gradually increased during the trial, peaking just after the animal reported its choice (Fig. 2c).

A careful comparison of choice selectivity revealed that inhibitory neurons reflected the animals' choice to a slightly greater extent than excitatory neurons. A slightly higher fraction of inhibitory neurons were choice selective (Fig. 2a,b; fraction choice selective: 15% of inhibitory neurons and 12% of excitatory neurons), and there was slightly stronger overall choice selectivity in inhibitory neurons (Fig. 2a,c,d).

The existence of similarly strong choice selectivity in excitatory and inhibitory neurons is a first hint that the connectivity between the two cell types preserves choice selectivity, suggesting non-random connections between inhibitory and excitatory neurons[7,24]. If choice selectivity is preserved because excitatory and inhibitory neurons with the same choice preference tend to be preferentially connected, one prediction is that excitatory and inhibitory neurons with the same choice selectivity will be more strongly correlated with each other compared to excitatory and inhibitory neurons with the opposite choice selectivity. This was indeed the case (Fig. 2e, dark green bars above light green bars). The stronger correlation between neurons with the same choice selectivity was also evident in pairs consisting of only excitatory or only inhibitory neurons (Extended Data Fig. 2e,f), in keeping with previous observations in V1 during passive viewing[12,14,15].

The animal's choice can be decoded with equal accuracy from both populations

The ability of individual inhibitory neurons to distinguish the animal's choice, to at least the same extent as excitatory neurons, argues against decision-making models that assume a non-selective inhibitory population. However, the small choice selectivity in single neurons (Fig. 2d) limits confidence in this conclusion. To further evaluate the discrimination ability of inhibitory neurons, we leveraged our ability to measure hundreds of neurons simultaneously. Specifically, we examined the ability of a linear classifier (support vector machines, SVM[41]) to predict the animal's choice from the population activity (cross-validated; L2 penalty; see Methods). We first performed this analysis on all neurons imaged simultaneously in a single session (Fig. 3a, left), training the classifier separately for every moment in the trial (97 ms bins).

111

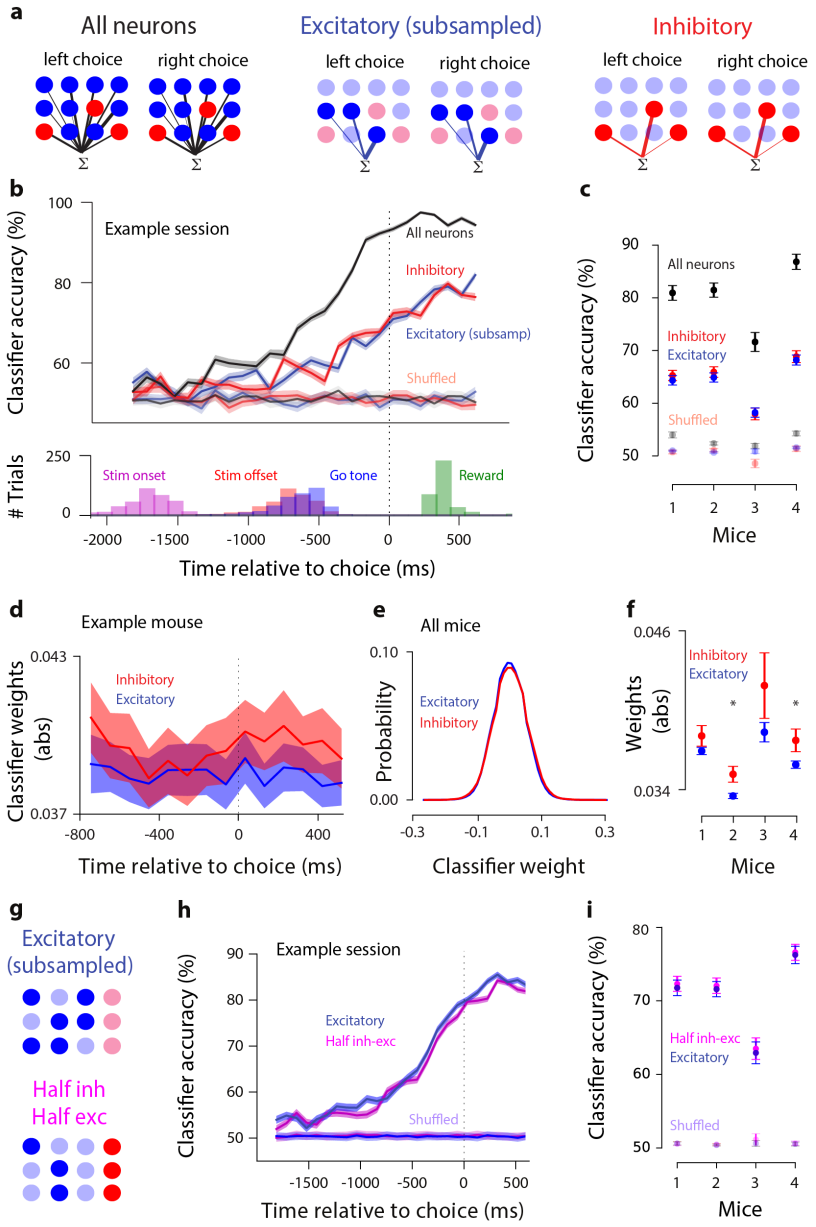


Figure 3. Linear classifiers can predict the animal's choice with equally high accuracy from the activity of either excitatory or inhibitory populations.

a, Schematic of decoding choice from the population activity of all neurons (left), only excitatory neurons (middle), subsampled to the same number as inhibitory neurons, and only inhibitory neurons (right). A linear SVM assigns weights of different magnitude (indicated by lines of different thickness) to each neuron in the population so that a weighted sum of population activity differs for trials preceding left vs. right choices. **b, Top:** classification accuracy of decoders trained on all neurons (black), subsampled excitatory neurons (blue), and inhibitory neurons (red) (cross-validated; decoders trained on every 97ms time bin; example session; mean+/-standard error across 50 cross-validated samples). Classification accuracy is lower for inhibitory or subsampled excitatory populations (red, blue) because of the smaller population size compared to all neurons (black). Classifier accuracy was similar for excitatory and inhibitory populations throughout the trial. Unsaturated lines show performance on shuffled trial labels. **Bottom:** distribution of stimulus onset, stimulus offset, go tone, and reward occurrence for the example session shown on the top. **c,** Classification accuracy during 97-0 ms before the choice for 4 animals on real (saturated points) and shuffled (unsaturated points) data.

Mean+/-standard error across days per mouse. **d-f,** When all neurons were included in the decoder (panel a, left), excitatory and inhibitory neurons were assigned weights of similar magnitude. **d,** Absolute value of weights for excitatory and inhibitory neurons in the decoders trained on all neurons, at every moment in the trial; example mouse; mean+/-standard error across days. **e,** Distribution of classifier weights (decoder training time: 97-0 ms before the choice) are overlapping for excitatory and inhibitory neurons. Neurons from all mice pooled (42,019 excitatory and 5,172 inhibitory neurons). **f,** Absolute value of weights in the classifier trained from 97-0ms before the choice for excitatory vs. inhibitory neurons, for each mouse. Mean+/-standard error across days. Star indicates $P<0.05$, t-test. **g-i,** No synergistic effect of excitatory and inhibitory neurons is evident when they are combined in a population. **g,** Schematic of decoding choice from a population of subsampled excitatory neurons (top) vs. a population of the same size but including half inhibitory and half excitatory neurons (bottom). **h,** Classifier accuracy of populations including only excitatory (blue) or half inhibitory, half excitatory neurons (magenta); example session. Classifier trained at each moment in the trial. Traces show mean+/-standard error across 50 cross-validated samples. **i,** Summary of each mouse (mean+/-standard error across days) for the decoders trained from 97-0 ms before the choice.

Classification accuracy gradually grew after stimulus onset and peaked at the time of the choice (Fig. 3b, black). Performance was at chance on a shuffle control in which trials were randomly assigned as left or right choice (Fig. 3b, shuffled). The ability of the entire population of PPC neurons to predict the animal's upcoming choice confirms previous observations[25,27,31,32]. As with the single neuron data (Extended Data Fig. 2c,d), analysis of error trials indicated that the animal's choice, rather than the stimulus, was generally the key feature driving the high classification accuracy we observed (Extended Data Fig. 3).

We then examined classifier accuracy for excitatory and inhibitory populations separately. For excitatory neurons, we subsampled the population so that the total number of neurons matched the number of inhibitory neurons in the same session (Fig. 3a, middle). As expected, overall classification accuracy was reduced due to the smaller population size; although performance was still well above chance and the temporal dynamics were the same as when all neurons were included (Fig. 3b, blue trace). Finally, we included all inhibitory neurons (Fig. 3a, right). Remarkably, the classification accuracy of inhibitory neurons closely mirrored that of excitatory neurons during the course of a trial (Fig. 3b, red and blue traces overlap). Similar classification accuracy for excitatory and inhibitory populations was observed in all subjects (Fig. 3c).

One possibility is that our analysis obscured a difference between excitatory and inhibitory neurons because it evaluated their performance separately, rather than considering how these neurons are leveraged collectively in a classifier that can take advantage of both cell types. To test this, we examined the classifier that was trained on all neurons (Fig. 3a left; Fig. 3b black), and compared the classifier weights assigned to excitatory vs. inhibitory neurons. If excitatory and inhibitory neurons contributed equally to the decoder, they should be assigned comparable weights by the classifier. This is indeed what we found: the weight magnitudes of excitatory and inhibitory neurons were matched for the entire course of the trial (Fig. 3d; absolute value of weights). Also the distributions of weights were overlapping (Fig. 3e: weights of all neurons of all mice. Fig. 3f: absolute weights). The comparable classifier weights for excitatory and

inhibitory neurons demonstrate that both cell types were similarly informative about the animal's upcoming choice. The average weights assigned by the classifier were slightly higher for inhibitory neurons (Fig. 3d-f), perhaps reflecting the slightly stronger choice selectivity in single neurons (Fig. 2d).

Finally, we tested whether excitatory and inhibitory populations are synergistic such that choice can be decoded more accurately from a mixed population. This could occur if excitatory and inhibitory populations have different sources of noise[42,43]. To assess this possibility we trained the classifier on a population that included half excitatory and half inhibitory neurons (Fig. 3g bottom), and compared its choice-prediction accuracy with the classifier that was trained on a population of the same size, but consisted only of excitatory neurons (Fig. 3g top). We found similar classification accuracy for both decoders during the entire trial (Fig. 3h,i), arguing against a synergistic effect of a mixed population and in favor of shared noise sources across all neurons.

Excitatory and inhibitory populations reflect the animals' choice with similar stability

If excitatory and inhibitory neurons are connected within subnetworks with frequent cross talk, the two populations should not only predict the animal's choice with similar accuracy, as shown above, but also with similar temporal dynamics. To assess this, we quantified each population's stability: the extent to which a classifier trained at one moment could successfully classify neural activity as preceding left vs. right choice at different moments. If population-wide patterns of activity are similar over time (e.g., all neurons gradually increase their firing rates), classifiers trained at one moment can accurately classify neural activity at different moments. Excitatory and inhibitory populations might differ in this regard, with one population more stable than the other.

As the gap between testing and training time increased, a gradual drop occurred in the classifier accuracy, as expected (Fig. 4a,b). This drop in accuracy occurred at a very similar rate for excitatory and inhibitory populations (Fig. 4b). To quantify this, we determined the time window over which the classifier accuracy was within 2 standard deviations of the accuracy at the same training-testing time window (Fig. 4c). This was indistinguishable for excitatory and inhibitory neurons (Fig. 4d; Extended Data Fig. 4a). An alternate method for assessing stability, computing the angle between pairs of classifiers trained at different time windows, likewise suggested that excitatory and inhibitory populations are similarly stable (Extended Data Fig. 4c).

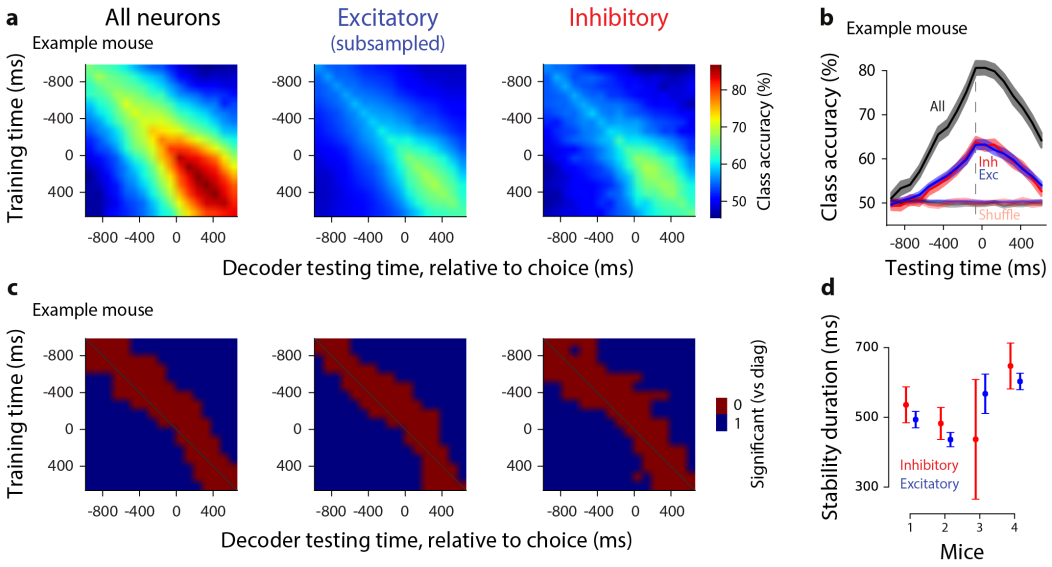


Figure 4. Classifiers, whether trained on excitatory or inhibitory neurons, show comparable stability during decision formation.
Cross-temporal generalization and stability of choice decoders. **a,** Classification accuracy of decoders for each pair of training and testing time points, using the population activity of all neurons (left), subsampled excitatory neurons (middle), or inhibitory neurons (right). Diagonal: same training, testing time (same as in Fig. 3). Example mouse, mean across 45 sessions. **b,** Example classification accuracy traces showing how classifiers trained at 97-0ms before choice generalize to other times in the trial. Excitatory and inhibitory neurons show the same time course of generalization. Same mouse as in (a), mean+/-standard error across days **c,** Decoders are stable in a short window outside their training time. Red indicates stability: classification accuracy of a decoder tested at a time different than its training time is within 2 standard deviation of the decoder tested at the same time as the training time. Example mouse; mean across days. **d,** Summary of stability duration for the decoder trained from 97-0 ms before the choice, using inhibitory neurons (red) or subsampled excitatory neurons (blue), for each mouse. Mean+/-standard error across days, per mouse.

Correlated variability is higher in the inhibitory population

Our single-neuron analysis (Fig. 2) demonstrated that choice selectivity is slightly higher in inhibitory neurons; however, our population analysis demonstrated similar fidelity of the population code for choice in inhibitory vs. excitatory neurons (Fig. 3b,c). This discrepancy could be due to different amounts of shared variability, or noise correlations, in the two populations. Noise correlations are activity patterns that are correlated among neurons and vary across repeated presentations of the same input. Our dataset, which included simultaneous activity from hundreds of neurons, was especially well-suited to assess noise correlations: correlations can have a large effect at the population level even when their effect at the level of neuron pairs is small[42,43].

To examine how noise correlations affected classification accuracy, we sorted neurons based on their individual choice selectivity and added them one by one to the population (from highest to

lowest choice selectivity defined as |AUC-0.5|). We measured classification accuracy for each of these neuronal ensembles of increasingly bigger size. Classification accuracy improved initially as more neurons were included in the decoder, but quickly saturated (Fig. 5a black; 97-0 ms before the choice).

To understand why classification accuracy saturates, we tested the effect of noise correlations on classification accuracy. Specifically, we created "pseudo populations", in which each neuron in the population was taken from a different trial (Fig. 5a gray). This removed noise correlations because those are shared across neurons within a single trial. Higher classification accuracy in pseudo populations compared to real populations indicates the presence of noise that overlaps with signal, constraining information. This is what we observed (Fig. 5a, gray trace above black trace). Across all mice, removing noise correlations resulted in a consistent increase in classification accuracy for the population including all neurons (Fig. 5b; filled vs. open circles). This establishes that correlations limit classification accuracy in the full population.

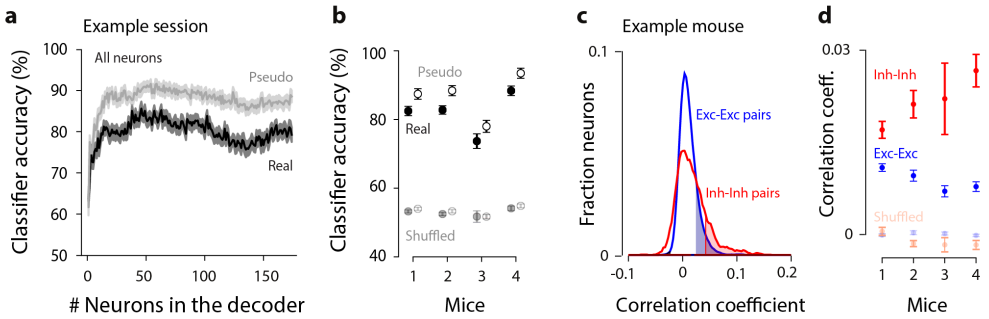


Figure 5. Noise correlations limit the efficacy of pooling and are higher among inhibitory neurons.
a, Classification accuracy for an example session (at time window 97-0 ms before the choice) on neural ensembles of increasingly bigger size, with the most choice-selective neurons added first. Mean+/-standard error across 50 cross validated samples. Gray: classification accuracy for pseudo-populations, in which noise correlations were removed by shuffling. Black: real populations. Both cell types were included ("All neurons"). **b,** Summary for each mouse; points show mean+/-standard error across days. Values were computed for the largest neuronal ensemble (the max value on the horizontal axis in (a)). **c,** Example mouse: distribution of Pearson's correlation coefficients of neural responses (97-0 ms before the choice) after subtracting out the mean activity, revealing noise correlations, for excitatory neurons (blue; n=11867) and inhibitory neurons (red; n=1583). Shaded areas indicate significant quantities compared to a shuffled control: trial orders were shuffled for each neuron to remove noise correlations. **d,** Summary of noise correlation coefficients for each mouse; mean+/-standard error across days.

To test the hypothesis that inhibitory neurons are more correlated than excitatory neurons, we compared the strength of pairwise noise correlations for excitatory vs. inhibitory neurons (Fig. 5c,d). Inhibitory pairs had significantly higher noise correlations compared to excitatory pairs (Fig. 5,d). Similar results were observed when we measured correlations during spontaneous activity (Extended Data Fig. 5b). Importantly, we obtained the same results even when we restricted the analysis to those inhibitory and excitatory neurons that had the same inferred spiking activity (Extended Data Fig. 5c,d). This was done because the higher spiking activity of

inhibitory neurons (Fig. 1h-j) could potentially muddle the comparison of pairwise noise correlations between excitatory and inhibitory neurons. These findings confirm previous studies that suggested the dense connectivity between inhibitory neurons leads to stronger correlations among inhibitory neurons compared to other cell types[12,23]. Also, similar to previous reports[12,23], we found intermediate correlations for pairs consisting of one inhibitory neuron and one excitatory neuron (Extended Data Fig. 5). The higher correlations in inhibitory neurons offer insight into why individual inhibitory neurons are slightly more choice selective than excitatory neurons (Fig. 2), yet at the ensemble level, both neuronal populations reflect choice with near-equal accuracy (Fig. 3).

Inhibitory neurons mirror excitatory neurons during the course of learning

Our observations thus far argue that excitatory and inhibitory neurons are similarly choice selective and similarly stable, in keeping with the hypothesis that they form selective subnetworks[24] according to their functional properties[7]. To distinguish whether these subnetworks are present in naïve animals or arise with experience, we measured neural activity as animals transitioned from novice to expert decision-makers. To achieve this, we leveraged the data from 3 mice in which the same neural population was imaged throughout learning (35-48 sessions; Extended Data Fig. 6a), to compare the dynamics of choice signal in excitatory and inhibitory populations over the course of learning.

Classification accuracy increased consistently as the animals became experts (Fig. 6a,d). This effect was significant in all animals and was present in both excitatory and inhibitory neurons (Fig. 6d). In all populations, classification accuracy was strongly correlated with the animal's performance across training days (Fig. 6b).

Accurate classification of choice also became more prompt: it appeared progressively earlier in the trial as the animals became experts. Initially, classification accuracy was high only after the choice (Fig. 6a; e.g. black arrow). As the animals gained experience, high classification accuracy was evident progressively earlier in the trial, long before the choice (Fig. 6a; e.g. gray arrow). This resulted in a negative correlation between animal's performance and the onset of super-threshold choice accuracy across training days (Fig. 6c). This effect was significant in all animals and was present in both excitatory and inhibitory neurons (Fig 6e).

Importantly, animal's licking or running behavior could not explain the learning-induced changes in the magnitude of choice classifier (Extended Data Fig. 7). The center-spout licks that preceded the left vs. right choices were overall similar over the course of learning (Extended Data Fig. 7a), and did not differ in early vs. late training days (Extended Data Fig. 7b).

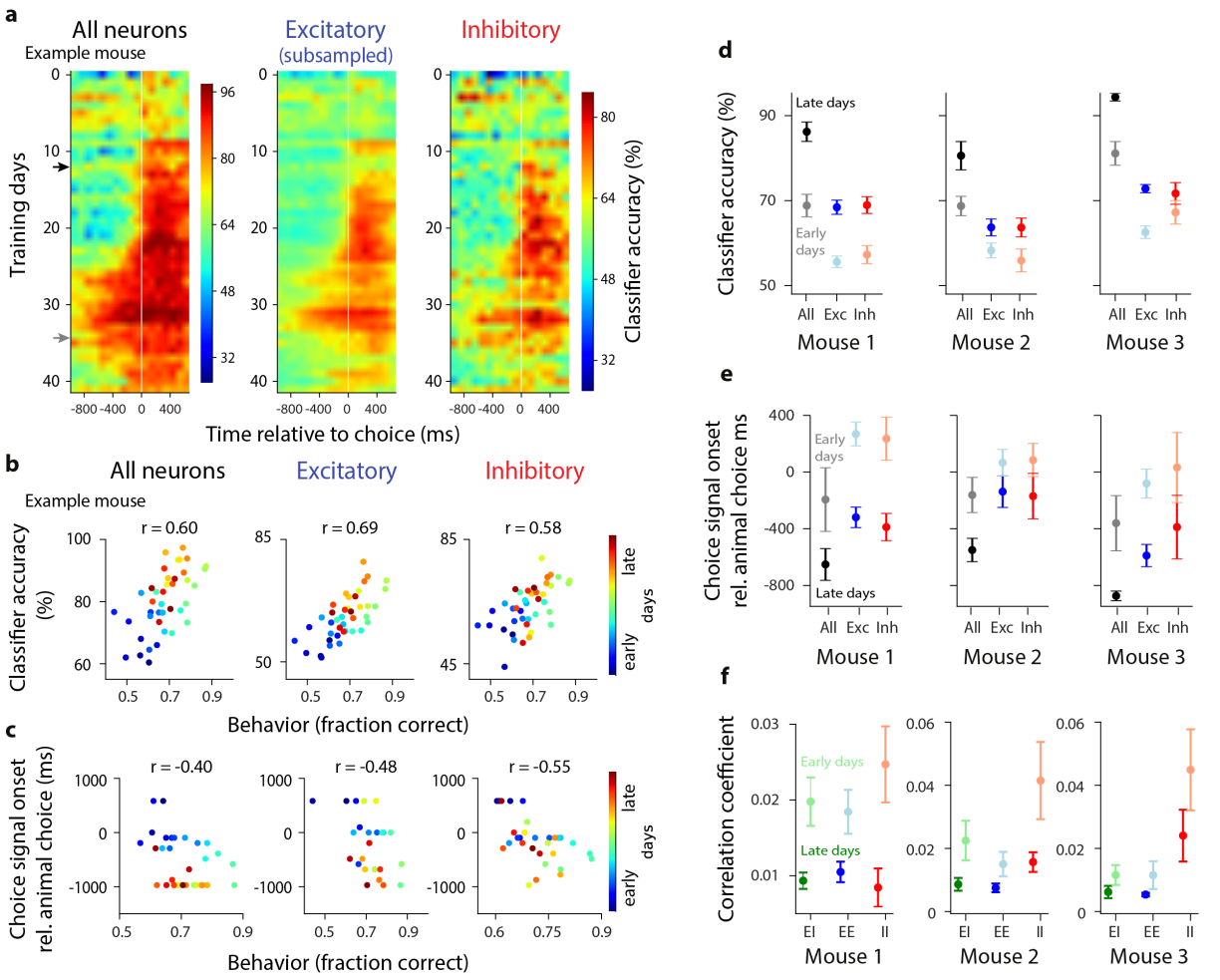


Figure 6. Learning leads to increased magnitude and advanced timing of the choice signal, as well as reduced noise correlations, in both excitatory and inhibitory populations.
a, Decoder accuracy is shown for each training session, for all neurons (left), subsampled excitatory (middle), and inhibitory neurons (right). White vertical line: choice onset. Each row: average across cross-validation samples; example mouse. Colorbar of the inhibitory plot applies to the excitatory plot too. **b,** Scatter plot of classifier accuracy at 97-0 ms before the choice vs. behavioral performance (fraction correct on easy trials), including all training days. Titles indicate r (Pearson correlation coefficient; $p<0.001$); same example mouse as in (a). Correlations for behavior vs. classification accuracy for all neurons, excitatory and inhibitory: 0.55, 0.35, 0.32 in mouse 2; 0.57, 0.63, 0.32 in mouse 3. Correlations for behavior vs. choice-signal onset for all neurons, excitatory and inhibitory: -0.60, -0.34, -0.38, in mouse 2; -0.60, -0.27, -0.28 in mouse 3. All values: $p<0.05$ **c,** Same as (b), except showing the onset of choice signal, i.e. the first moment in the trial that classifier accuracy was above chance (ms, relative to choice onset) vs. behavioral performance. **d,** Summary of each mouse, showing classification accuracy averaged across early (dim colors) vs. late (dark colors) training days. **e,** Same as (d), but showing choice signal onset (ms). **f,** Same as (d), but showing pairwise noise correlation coefficients. Early days were the first few training days in which the animal's performance was lower than the 20th percentile of animal's performance across all days. Late days included the last training days in which the animal's behavioral performance was above the 80th percentile of performance across all days.

233

The similarity in lick movements for early vs. late sessions stands in contrast to the changes in the choice classifier for early vs. late sessions (Fig. 6). We also assessed animals' running behavior during the course of learning (Extended Data Fig. 7c,d). In some sessions, the running distance differed preceding left vs. right choices (Extended Data Fig. 7c). Fortunately, when we restricted our analysis to days in which the running distance was indistinguishable for the two choices (97-0ms before the choice, t-test, $P>0.05$), we nonetheless were able to accurately classify the animal's choice using neural activity (Extended Data Fig. 7d). These observations provide reassurance that the population activity does not entirely reflect preparation of licking and running movements and argues instead that the population activity reflects the animal's stimulus-informed choice. Nevertheless, the correlation between movements and choice that we report here and others have observed[44] argues that the link between movements, neural activity and decision-making must be more deeply investigated.

Finally, we studied how cofluctuations changed over the course of training. Pairwise correlations in neural activity were overall higher in early training days, when mice were novices, compared to late training days, as they approached expert level behavior (Fig. 6f, unsaturated colors above saturated colors). This effect was observed for all combinations of neural pairs (Fig. 6f, green: excitatory-inhibitory; blue: excitatory-excitatory; red: inhibitory-inhibitory). These findings are in agreement with previous reports suggesting that learning results in reduced noise correlations[45-47], enhancing information that is encoded in neural populations[42,48]. To test if the learning-induced increase in classification accuracy (Fig. 6a,b,d) was all a consequence of the reduction in noise correlations (Fig. 6f), we studied how classification accuracy of pseudo populations (Fig. 5a,b), which lack noise correlations, changed with training. Interestingly, we still observed a significant increase in the classification accuracy of pseudo populations as a result of training (Extended Data Fig. 8a,b). Therefore, the reduction in noise correlations cannot alone account for the improved choice prediction that occurs in the population with training.

Discussion

Despite a wealth of studies assessing selectivity of inhibitory neurons during passive viewing of sensory stimuli, little is known about whether these principles extend to decision-making. This represents a critical gap in knowledge because many decision-making models rely on inhibitory neurons. To close this gap, we measured excitatory and inhibitory populations during perceptual decisions about multisensory stimuli.

We demonstrated that inhibitory populations can predict the animal's impending choice with the same fidelity as excitatory populations. This advocates for specific connectivity between excitatory and inhibitory neurons that preserves selectivity for choice[2]. These observations are significant because they argue that decision circuits share the subnetwork structure characterized in primary sensory areas in which excitatory neurons target specific inhibitory populations[24] with similar tuning[7].

Our additional observations likewise suggest commonalities between decision circuits in PPC and sensory circuits in V1. In V1, inhibitory neurons are more interconnected[20,39,49] and share more common input[12] compared to excitatory neurons. Further, noise correlations are proportional to the tuning similarity of neurons[14,15]. Similarly, we found that among inhibitory neurons, noise correlations were stronger than among excitatory neurons, and that in both populations noise correlations were stronger between neurons selective for the same choice.

As mice learned the decision-making task, noise correlations gradually decreased while classification accuracy increased (Fig. 6). Multiple studies have shown that correlated variability changes with cognitive processes, such as perceptual learning and attention[45-47,50]. Here, the reduced noise correlations may partially account for the improved classification accuracy with learning. However, the reduction in noise correlations is unlikely to be the sole mechanism supporting the behavioral improvement: even in the absence of noise correlations, choice selectivity in the population increased with learning (Extended Data Fig 6a-b). Future experiments using causal manipulations will reveal whether the increased choice selectivity we observed in PPC originates there or is inherited from elsewhere in the brain.

We have brought to decision-making an approach that has been instrumental for understanding how incoming inputs are processed in early sensory areas[7,9,11,12,21,22,36,51,52,10,53,54]. Our results advocate for models in which the connectivity preserves the choice selectivity in inhibitory neurons, even as they are barraged with input from excitatory neurons. This targeted connectivity may implement a canonical neural computation, as it is present in early sensory areas and oculomotor areas, and is implicated on theoretical grounds because of stability and robustness[2,7].

Methods

Imaging and behavioral dataset

Our simultaneous imaging and decision-making dataset includes 135 sessions from 4 mice (45, 48, 7, and 35 sessions per mouse). Median number of trials per session is 213, 253, 264, and 222, for each mouse. On average, 480 neurons were imaged per session, out of which 40 neurons were inhibitory and 330 were excitatory. Approximately 100 neurons per session were not classified as either excitatory or inhibitory since they did not meet our strict cell-type classification criteria (see below). In 3 of the mice, the same group of neurons was imaged throughout learning (35-48 training days).

Mice and surgical procedure

Gad2-IRES-CRE[55] mice were crossed with Rosa-CAG-LSL-tdTomato-WPRE (aka Ai14[56]) to create mice in which all GABAergic inhibitory neurons were labeled. Adult mice (~2-month old) were used in the experiments. Meloxicam (analgesic), dexamethasone (anti-inflammatory) and Baytril (enrofloxacin; anti-biotic) were injected 30min before surgery. Using biopsy punch, a circular craniotomy (diameter: 3mm) was made over the left PPC (stereotaxic coordinates: 2 mm

posterior, 1.7 mm lateral of bregma[26] under isoflurane (~5%) anesthesia. Pipettes (10-20 um in diameter, cut at an angle to provide a beveled tip) were front-filled with AAV9-Synapsin-GCaMP6f (U Penn, Vector Core Facility) diluted 2X in PBS (Phosphae-buffered saline). The pipette was slowly advanced into the brain (Narishige MO-8 hydraulic micro-manipulator) to make ~3 injections of 50nL, slowly at an interval of ~5-10min, by applying air pressure using a syringe. Injections ware made near the center of craniotomy at a depth of 250-350um below the dura. A glass plug consisting of a 5mm coverslip attached to a 3mm coverslip (using IR-curable optical bond, Norland) was used to cover the craniotomy window. Vetbond, followed by metabond, was used to seal the window. All surgical and behavioral procedures conformed to the guidelines established by the National Institutes of Health and were approved by the Institutional Animal Care and Use Committee of Cold Spring Harbor Laboratory.

Imaging

We used a 2-photon Moveable Objective Microscope with resonant scanning at 30 frames per second (Sutter Instruments, San Francisco, CA). A 16X, 0.8 NA Nikon objective lens was used to focus light on fields of view of size 512x512 pixels (~575 µm x ~575 µm). A Ti:sapphire laser (Coherent) delivered excitation light at 930nm (average power: 20-70 mW). Red (ET670/50m) and green (ET 525/50m) filters (Chroma Technologies) were used to collect red and green emission light. The microscope was controlled by Mscan (Sutter). In mice in which chronic imaging was performed during learning, the same plane was identified on consecutive days using both coarse alignment, based on superficial blood vessels, as well as fine alignment, using reference images of the red channel (tdTomato expression channel) at multiple magnification levels. For each trial, imaging was started 500ms before the trial-initiation tone, and continued 500ms after reward or time-out.

Decision-making behavior

Mice were gradually water restricted over the course of a week, and were weighed daily. Mice harvested at least 1 mL of water per behavioral/imaging session, and completed 100-500 trials per session. After approximately one week of habituation to the behavioral setup, 15-30 training days were required to achieve 75% correct choice. Animal training took place in a sound isolation chamber. The stimulus in all trials was multisensory, consisting of a series of simultaneous auditory clicks and visual flashes, occurring with Poisson statistics[57,58]. Stimulus duration was 1000ms. Each pulse was 5ms; minimum interval between pulses was 32ms, and maximum interval was 250ms. The average repetition rate of the pulses varied between 5 to 27 Hz. The category boundary for marking high-rate and low-rate stimuli was 16Hz, at which animals were rewarded randomly on either side. The highest stimulus rates used here are known to elicit reliable, steady state flicker responses in retinal ERG in mice[59,60].

Mice were on top of a cylindrical wheel and a rotary encoder was used to measure their running speed. Trials started with a 50ms initiation tone (Extended Data Fig. 1a). Mice had 5sec to

initiate a trial by licking the center waterspout, after which the multisensory stimulus was played for 1 second. If mice again licked the center waterspout, they received 0.5µL water on the center spout; also a 50ms go cue was immediately played. Animals had to report a choice by licking to the left or right waterspout within 2sec. Mice were required to confirm their choice by licking the same waterspout one more time within 300ms after the initial lick. If the choice was correct, mice received 2-4µL water on the corresponding waterspout. An incorrect choice was punished with a 2sec time-out. The experimenter-imposed inter-trial intervals (ITI) were drawn from a truncated exponential distribution, with minimum, maximum, and lambda equal to 1sec, 5sec, and 0.3sec, respectively. However, the actual ITIs could be much longer depending on when the animal initiates the next trial. Bcontrol[28] with a Matlab interface was used to deliver trial events (stimulus, reward, etc) and collect data**.**

Logistic regression model of behavior

A modified version of the logistic regression model in[61] was used to assess the extent to which the animal's choice depends on the strength of sensory evidence, i.e. how far the stimulus rate is from the category boundary at 16Hz, the previous choice outcome (success or failure) and ITI, i.e. the time interval between the previous choice and the current stimulus onset (Extended Data Fig. 1b).

$$p = \frac{1}{1+e^{-z}} \qquad \textit{eq. 1}$$

$$z = \beta_0 + (\beta_{r1} R_1 + \beta_{r2} R_2 + \beta_{r3} R_3 + \beta_{r4} R_4 + \beta_{r5} R_5 + \beta_{r6} R_6) + (\beta_{s1} S_1 + \beta_{s2} S_2) + (\beta_{f1} F_1 + \beta_{f2} F_2)$$

where p is the probability of choosing the left choice, and z is the decision variable. R, S and F are vectors of indicator variables; each element corresponds to 1 trial. Stimulus strength (R) was divided into 6 bins (R_1 to R_6). Previous success (S) was divided into 2 bins (S_1 to S_2): success after a long ITI (> 7sec) and success after a short ITI (< 7sec). Previous failure (F) was divided into 2 bins (F_1 to F_2): failure after a long and short ITI. For instance, if a trial had stimulus strength 3Hz, and was preceded by a success choice with ITI 5sec, we will set all R, S and F parameters to 0, except for R_2 and S_1, which will be set to 1, indicating that the trial's stimulus strength was in bin 2, and it was preceded by a success whose ITI was in bin 1 (Extended Data Fig. 1b).

For each session the scalar coefficients β_0, β_{r1} to β_{r6}, β_{s1}, β_{s2}, β_{f1}, and β_{f2} were fitted using Matlab glmfit.m. Extended Data Fig. 1b left shows β_{r1} to β_{r6}. Extended Data Fig. 1b middle shows β_{s1} and β_{s2}, and Extended Data Fig. 1b right shows β_{f1} and β_{f2}.

ROI extraction and deconvolution

The recorded movies from all trials were concatenated and corrected for motion artifacts by cross-correlation using DFT registration[62]. Subsequently, active ROIs (sources) were extracted using the CNMF algorithm[34] as implemented in the CaImAn package[35] in MATLAB. The traces

of the identified neurons were ΔF/F normalized and then deconvolved by adapting the FOOPSI deconvolution algorithm[34,63] to a multi-trial setup. More specifically, for each component, the activity trace over all the trials was used to determine the time constants of the calcium indicator dynamics as in[34]. Then the neural activity during each trial was deconvolved separately using the estimated time constant and a zero baseline (since the traces were ΔF/F normalized). A second order exponential was used to simulate the rise and decay of the indicator. The deconvolved neural activity (spikes) was normalized such that a deconvolved value of 1 corresponded to a spike that could generate a calcium transient that reaches a maximum 100% ΔF/F value (Fig. 1e). Spikes with amplitudes less than 0.5x the noise level were discarded.

Neuropil Contamination removal

The CNMF algorithm demixes the activity of overlapping neurons, also from background neuropil activity by explicitly modeling the neuropil activity as a low rank spatiotemporal matrix[34]. In this study a rank two matrix was used to capture the neuropil activity. To evaluate its efficacy we compared the traces obtained from CNMF to the traces from a "manual" method similar to[36] (Extended Data Fig. 9): the set of spatial footprints (shapes) extracted from CNMF were binarized by thresholding each component at the 0.2x its maximum value level. The binary masks were then used to average the raw data and obtain an activity trace that also included neuropil effects. To estimate the background signal, an annulus around the binary mask was constructed with minimum distance 3 pixels from the binary mask and width 7 pixels (Extended Data Fig. 9a). The average of the raw data over the annulus defined the background trace, which was then subtracted from the activity trace. The resulted trace was then compared with the CNMF estimated temporal trace for this activity. The comparison showed a very high degree of similarity between the two traces (Extended Data Fig. 9; example component; r=0.96), with the differences between the components being attributed to noise and not neuropil related events. Note that this "manual" approach is only applicable in the case when the annulus does not overlap with any other detected sources. These results demonstrate the ability of the CNMF framework to properly capture neuropil contamination and remove it from the detected calcium traces.

ROI inclusion criteria

We excluded poor-quality ROIs identified by the CNMF algorithm based on a combination of criteria: 1) size of the spatial component, 2) decay time constant, 3) correlation of the spatial component with the raw ROI image built by averaging spiking frames, 4) correlation of the temporal component with the raw activity trace, and 5) the probability of fluorescence traces maintaining values above an estimated signal-to-noise level for the expected duration of a calcium transient[35] (GCaMP6f, frame rate: 30Hz). A final manual inspection was performed on the selected ROIs to validate their shape and trace quality.

Identification of inhibitory neurons

We used a two-step method to identify inhibitory neurons. First, we corrected for bleed-through from green to red channel by solving the following common-slope regression model ("common" because it finds a single slope for all ROIs):

$$r_i = \beta + s * g_i \qquad eq.\ 2$$

where r_i is the average of pixels that belong to ROI 'i' on the mean projection image of the red channel. Similarly g_i is the average of pixels of that same ROI, but on the mean projection image of the green channel and β is an offset. We solved this equation simultaneously for all ROIs to get s, a common slope, which reflected the fraction bleed-through from green to red channel. Then we used this slope (s) to compute the bleedthrough-corrected image of the red-channel (I):

$$I = R - s * G \qquad eq.\ 3$$

where R is the mean projection image of the red channel, and G is the mean projection of the green channel.

Next, we identified inhibitory neurons on the bleedthrough-corrected image (I) using two measures. 1) A measure of local contrast, by computing, on the red channel, the average pixel intensity inside each ROI mask relative to its immediate surrounding mask (width=3 pixels). Given the distribution of contrast levels, we defined two threshold levels, T_E and T_I. ROIs whose contrast measure fell above T_I were identified as inhibitory neurons. ROIs whose contrast measure fell below $T_E$were identified as excitatory neurons, and ROIs with the contrast measure in between T_E and T_I were not classified as either group ("unsure" class). 2) In addition to a measure of local contrast, we computed for each ROI the correlation between the spatial component (ROI image on the green channel) and the corresponding area on the red channel. High correlation values indicate that the ROI on the green channel has a high signal on the red channel too; hence the ROI is an inhibitory neuron. We used this correlation measure to further refine the neuron classes computed from the local contrast measure (i.e. measure 1 above). ROIs that were identified as inhibitory based on their local contrast (measure 1) but had low red-green channel correlation (measure 2), were reset as "unsure" neurons. Similarly, ROIs that were classified as excitatory (based on their local contrast) but had high red-green channel correlation were reclassified as unsure. Unsure ROIs were included in the analysis of all-neuron populations (Fig. 3a left); but were excluded from the analysis of excitatory only or inhibitory only populations (Fig. 3a middle, right). Finally, we manually inspected the ROIs identified as inhibitory to confirm their validity. This method resulted in 8.3% inhibitory neurons, which is fairly in agreement with previous studies, although slightly lower than some reports[37]. This was expected given our strict method for identifying inhibitory neurons. The inhibitory population is likely biased towards parvalbumin positive (PV) neurons, because 1) we found higher inferred spiking activity for our inhibitory neurons (Fig. 1h-j) which is expected for PV neurons[39]; 2) PV cells are the most prevalent interneurons in cortical circuits[64].

General analysis procedures

All analyses were performed on inferred spiking activity. Traces were down-sampled, so each bin was the non-overlapping moving average of 3 frames (97.1ms). Inferred spiking activity for each neuron was normalized so the max spiking activity for each neuron equaled 1. The trace of each trial was aligned to the time of the choice (i.e. the time of the 1st lick to either of the side waterspouts after the go tone). Two-tailed t-test was performed for testing statistical significance. Summary figures including all mice were performed on the time bin preceding the choice, i.e. 97-0ms before choice onset. All reported correlations are Pearson's coefficients. Analyses were performed in Python and Matlab.

ROC analysis

The area under the ROC curve (AUC) was used to measure the choice preference of single neurons. Choice selectivity was defined as the absolute deviation of AUC from chance level (0.5). To identify significantly choice-selective neurons, for each neuron we performed ROC on shuffled trial labels (i.e. left and right choices were randomly assigned to each trial). This procedure was repeated 50 times to create a distribution of shuffled AUC values for each neuron (Extended Data Fig. 2a, "shuffled"). A neuron's choice selectivity was considered to be significant if the probability of the actual AUC (Extended Data Fig. 2a, "real") being drawn from the shuffled AUC distribution was less than 0.05.

Decoding population activity

A linear SVM (Python sklearn package) was trained on each bin of the population activity in each session (non-overlapping 97ms time bins). To break any dependencies on the sequence of trials, we shuffled the order of trials for the entire population. To avoid bias in favor of one choice over the other, we matched the number of left- and right-choice trials used for classifier training. L2 regularization was used to avoid over-fitting. 10-fold cross validation was performed by leaving out a random 10% subset of trials to test the classifier performance, and using the remaining trials for training the classifier. This procedure was repeated 50 times. A range of regularization values was tested, and the one that gave the smallest error on the validation dataset was chosen as the optimal regularization parameter. Classifier accuracy was computed as the percentage of testing trials in which the animal's choice was accurately predicted by the classifier, and summarized as the average across the 50 repetitions of trial subsampling. A minimum of 10 correct trials per choice was required in order to run the SVM on a session. Inferred spiking activity of each neuron was z-scored before running the SVM.

When comparing classification accuracy for excitatory vs. inhibitory neurons, the excitatory population was randomly sub-sampled to match the population size of inhibitory neurons to enable a fair comparison (Fig. 3, blue vs. red). To compare the distribution of weights in the all-

neuron classifier (Fig. 3 black), the weight vector for each session was normalized to unity length (Fig. 3d-f).

Stability

To test the stability of the population code, decoders were trained and tested at different time bins[65] (Fig. 4). To avoid the potential effects of auto-correlation, we performed cross validation not only across time bins, but also across trials. In other words, even though the procedure was cross validated by testing the classifier at a time different from the training time, we added another level of cross-validation by testing on a subset of trials that were not used for training. This strict method allowed our measure of stability duration to be free of auto-correlation impacts.

As an alternative measure of stability, the angle between pairs of classifiers that were trained at different moments in the trial was computed (Extended Data Fig. 4c). Small angles indicate alignment, hence stability, of the classifiers. Large angles indicate misalignment, i.e. instability of the classifiers.

Noise correlations

Noise correlations were assessed at the level of the population as well as the neuron pairs. To estimate noise correlations at the population level, the order of trials was shuffled for each neuron independently. Shuffling was done within the trials of each choice, hence retaining the choice signal, while de-correlating the population activity to remove noise correlations. Then we classified population activity in advance of left vs. right choice (at time bin 97–0 ms before the choice) using the de-correlated population activity. This procedure was performed on neural ensembles of increasingly bigger size, with the most selective neurons (|AUC-0.5|) added first (Fig. 5a). To summarize how noise correlations affected classification accuracy in the population (Fig. 5b), we computed, for the largest neural ensemble (Fig. 5a, max value on the horizontal axis), the change in classifier accuracy in the de-correlated data ("pseudo populations") vs. the original data. This analysis was only performed for the entire population; the small number of inhibitory neurons in each session prevented a meaningful comparison of classification accuracy on real vs. pseudo populations.

To measure pairwise noise correlations, we subtracted the trial-averaged response to a particular choice from the response of single trials of that choice. This allowed removing the effect of choice on neural responses. The remaining variability in trial-by-trial responses can be attributed to noise correlations, measured as the Pearson correlation coefficient for neuron pairs. We also measured noise correlations using the spontaneous activity defined as the neural responses in 97-0ms preceding the trial initiation tone (Extended Data Fig. 2f; Extended Data Fig. 5b). We computed the pairwise correlation coefficient (Pearson) for a given neuron with each other neuron within an ensemble (e.g., excitatory neurons). The resulting coefficients were then

averaged to generate a single correlation value for that neuron. This was repeated for all neurons within the ensemble (Fig. 5c).

To compute pairwise correlations on excitatory and inhibitory neurons with the same inferred spiking activity (Extended Data Fig. 5c,d), we computed the median of inferred spiking activity across trials for individual excitatory and inhibitory neurons in a session. The medians were then divided into 50 bins. The firing-rate bin that included the maximum number of inhibitory neurons was identified ("max bin"); inhibitory and excitatory neurons whose firing rate was within this "max bin" were used for the analysis. The firing rates were matched for these neurons because their median firing rate was within the same small bin of firing rates. Pairwise correlations were then computed as above.

Learning analysis

In 3 of the mice, the same field of view was imaged each session during learning. This was achieved in two ways. First, the vasculature allowed a coarse alignment of the imaging location from day to day. Second, the image from the red channel was used for a finer alignment. Overall, most neurons were stably present across sessions (Extended Data Fig. 6). This allowed us to assess the population activity of a similar group of neurons during learning.

"Early days" (Fig. 6; Extended Data Fig. 8) included the initial training days in which the animal's performance, defined as the fraction of correct choices on easy trials, was lower than the 20th percentile of performance across all days. "Late days" (Fig. 6; Extended Data Fig. 8) included the last training days in which the animal's behavioral performance was above the 80^{th} percentile of performance across all days.

To measure the timing of decision-related activity (Fig. 6c,e), we identified all sessions in which classifier accuracy was significantly different than the shuffle (t-test, $p<0.05$) over a window of significance that was at least 500 ms long. We defined the "choice signal onset" (Fig. 6c,e) as the trial time corresponding to the first moment of that window. Sessions in which the 500 ms window of significance was present are included in Fig. 6c. The number of points (and hence the relationship between session number and color in Fig. 6c) differs slightly across the three groups. This is because on some sessions, the window of significance was present in one group but not another. For example, in a session the population including all neurons might have a 500 ms window of significance, hence it will contribute a point to Fig. 6c left, while the population with only inhibitory neurons might be only transiently significant for <500ms, hence it will be absent from Fig. 6c right.

Data and Software Availability

The CNMF algorithm is publicly available on github:
https://github.com/flatironinstitute/CaImAn-MATLAB

Code availability

559 All codes will be made publicly available on github before publication.

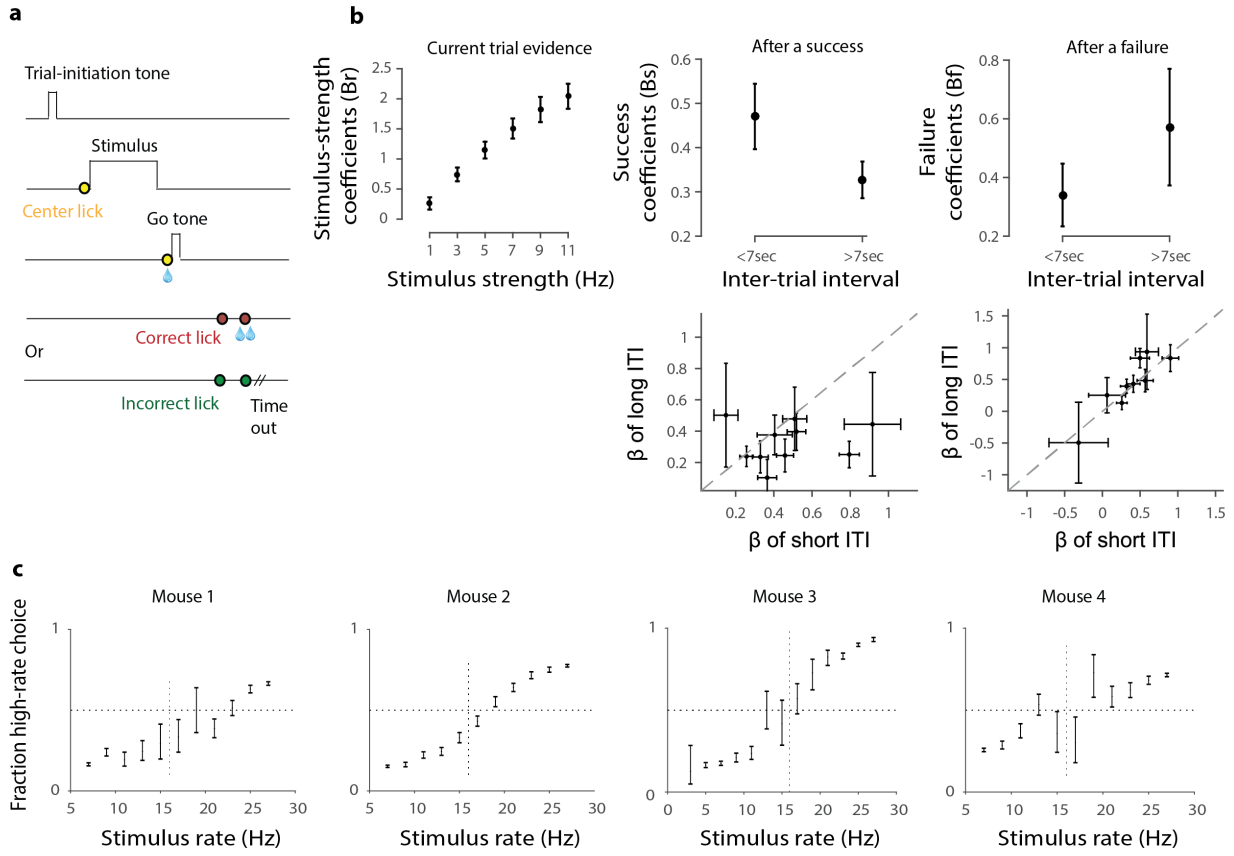
560 **Author Contributions**

561 Conceptualization and Writing: FN and AKC. Experiments and Analysis: FN. Decoding
562 methodology and common-slope regression model: GFE, JPC and FN. Spike-inference
563 methodology: EAP. Funding Acquisition, Resources and Supervision: AKC.

564 **Acknowledgements**

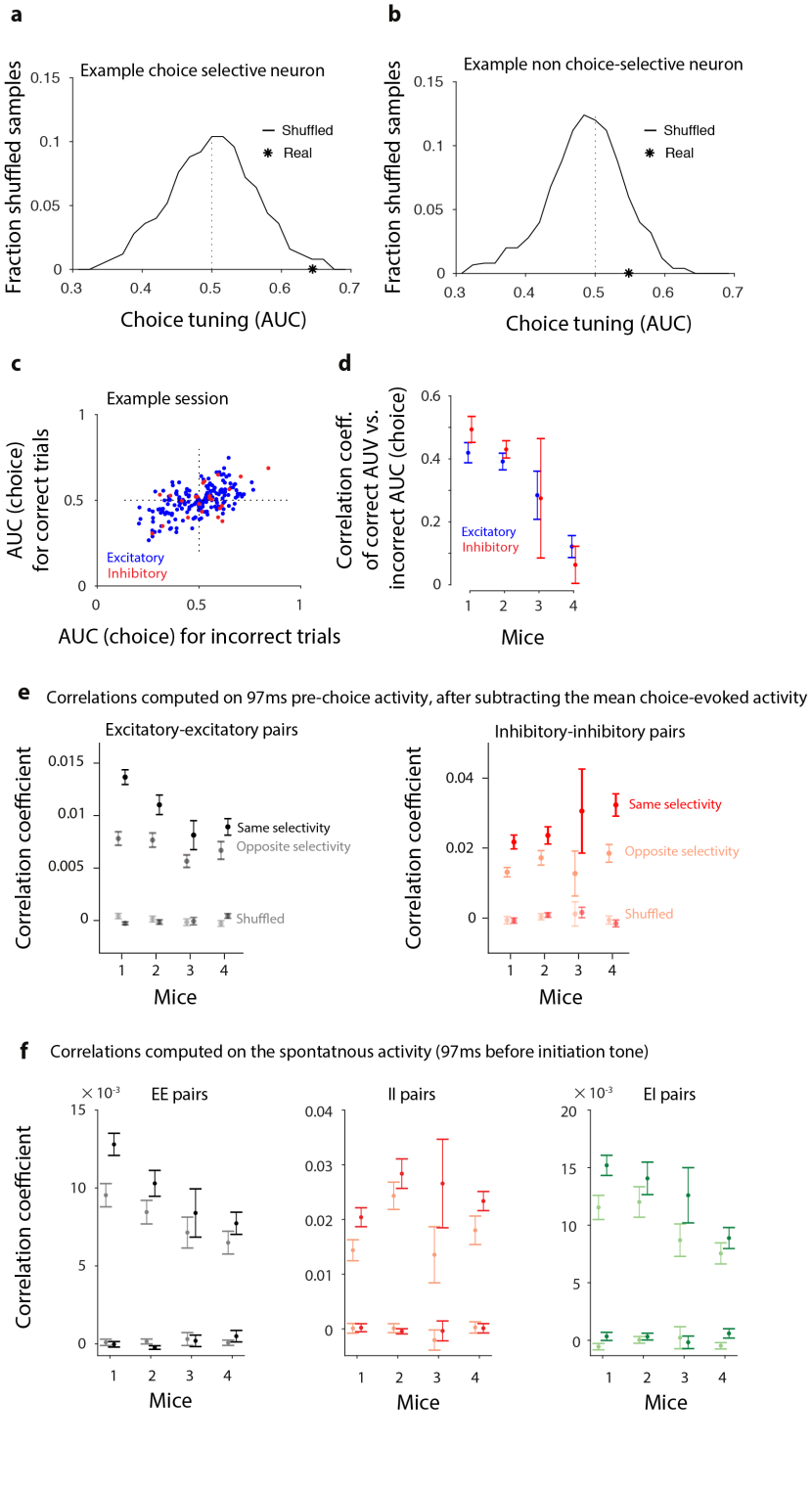
565 We thank Hien Nguyen for help with training mice, Matt Kaufman, Kachi Odoemene, Fred
566 Marbach for technical assistance and thoughtful conversations. We thank Peter Latham, Robin
567 Cao for helpful discussions on decision-making models. We thank Andrea Giovannucci for help
568 with ROI inclusion criteria #5. We thank Ashley Juavinett, Simon Musall, and Sashank Pisupati
569 for helpful discussions and feedback on the manuscript. The work was supported by the Simons
570 Collaboration on the Global Brain, ONR MURI, the Klingenstein-Simons Foundation and the
571 Pew Charitable Trust.

Extended Data Figures



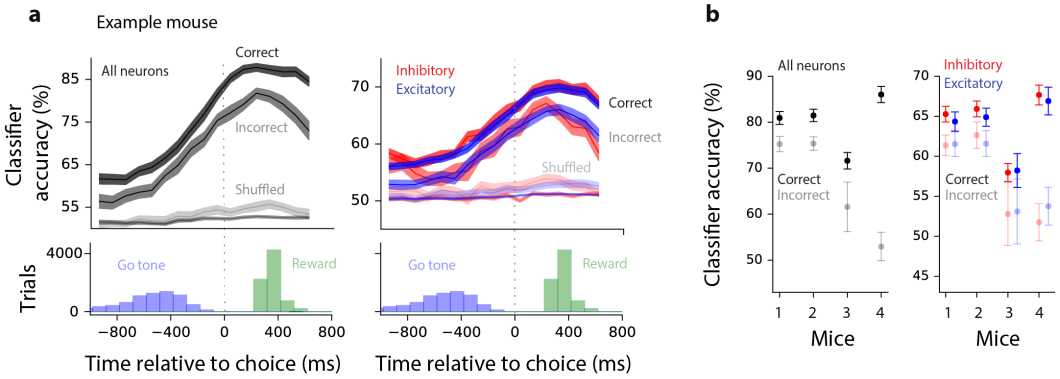
Extended data Figure 1. Perceptual decisions about stimulus rate reflect current evidence, previous trial's outcome, and the time passed since the previous trial.

Refers to Fig. 1. **a,** Trial structure. In each trial, first a brief tone is presented to indicate to the animal to initiate the trial ("trial-initiation tone"). Once the animal licks to the center waterspout (row 2: yellow circle), the stimulus is presented for 1 sec. At the end of the stimulus, the animal is required to lick again in the center (row 3: yellow circle). This will result in: 1) a small water reward in the center, 2) a "go tone" that indicates to the animal to make his choice. If the animal licks to the correct side (row 4, 1st red circle), and confirms his lick (row 4, 2nd red circle), he will receive water as a reward. If the animal licks to the wrong side (last row, 1st green circle), and confirms his licks (last row, 2nd green circle), there will be a time-out, i.e. longer time before the next trial can start. **b**, A logistic regression model was used to assess the extent to which the animal's choice depends on stimulus strength, i.e. how far the stimulus rate is from the category boundary at 16Hz, previous choice outcome, and the time interval since the previous choice. Stimulus strength was divided into 6 bins (**left**); previous success was divided into 2 bins: success after a long ITI and success after a short ITI (**middle**); previous failure was also divided into 2 bins: failure after a long ITI and failure after a short ITI (**right**). Plots in top row show Beta values averaged across animals (same 10 animals as in Fig. 1c). Error bars: S.E.M across subjects. **Top left**: strength of the sensory evidence impacts animal's choices: the stronger the evidence, the higher the impact. **Top middle**: Success of a previous trial also impacts animal's decision; the impact is stronger when the previous success occurs after a short ITI (<7sec). **Top Right**: Same but for previous incorrect trials, except the effect of ITI after a failure was not significant. Plots in **bottom** row show success (left) and failure (right) Beta values for individual mice. Error bars: S.E.M returned from glmfit.m in Matlab. **c,** Behavioral performance of the four mice in which we imaged excitatory and inhibitory activity during decision-making. In mice 1, 2, and 4, imaging was performed throughout learning by tracking the same group of neurons. Plots reflect data from all sessions. Errors bars: Wilson Binomial Confidence Interval.



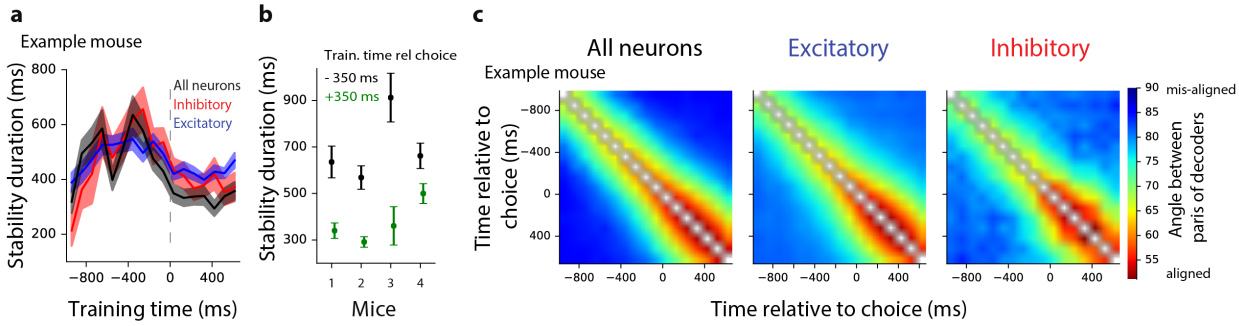
Extended data Figure 2. Single neuron measures reveal similar choice selectivity in excitatory and inhibitory neurons

Refers to Fig. 2. **a,b,** Example neurons to illustrate the method for assessing significant choice selectivity in individual neurons. In both panels, the solid line shows the distribution of values for the area under the ROC curve (AUC) generated by 50 different trial shuffles in which trials were randomly assigned to a left vs. right choice. Star indicates the real AUC value of the neuron. Significance was assessed from the probability of observing the real AUC in the shuffle distribution. When probabilities were <0.05, neurons were considered choice selective. Only the neuron in (a) has significant choice selectivity. **c,** Choice selectivity was computed on correct trials (vertical axis) as well as error trials (horizontal axis), and was correlated between the two conditions. Data is from a single session, each point shows an individual neuron whose cell type is indicated by the color. The positive correlation indicates that choice selectivity was overall similar on correct and error trials (Pearsons' correlation coefficient, excitatory neurons: $r=0.58$; $p<0.001$; inhibitory neurons: $r=0.55$, $p=0.007$). The small number of points in quadrants 2 and 4 indicate less frequent neurons that showed opposite selectivity on correct vs. error trials. **d,** Summary of correlation coefficient of AUC on correct vs. AUC on incorrect trials, mean across sessions for each animal. Error bars: S.E.M. across sessions. The weaker correlation in mouse 4 indicates that this animal had a mixture of cells selective for the stimulus and cells selective for the choice. Note that although the center of the imaging window was identical in all animals, the imaging location within the window of this animal was slightly posterior to the others. The enrichment of cells selective for the stimulus, in this mouse compared to other mice, may reflect that the region we imaged in mouse 4 was closer to primary visual cortex. **e,** Noise correlations for pairs of excitatory neurons (left) and pairs of inhibitory neurons (red). Noise correlations during the time period 97 ms before the choice. Signal correlations were removed by subtracting off the across-trial mean response from each trial, separately done for trials of each choice. Noise correlations were overall much stronger for inhibitory neurons (note different scaling on vertical axes for left vs. right plots). For both cell types, correlations were stronger when cells had the same choice selectivity (saturated colors) vs. opposite choice selectivity (unsaturated colors). Unsaturated points are for shuffled trials in which the trial order is scrambled for each neuron independently. Mean +/- standard error across sessions. Correlations for excitatory-inhibitory pairs are shown in Fig. 2e. **f,** Noise correlations for three combinations of cell types: excitatory-excitatory (left), inhibitory-inhibitory (middle) and excitatory-inhibitory (right). Signal correlations were not present because the data are from the inter-trial interval (97 ms before the initiation tone) and therefore only reflect spontaneous activity.



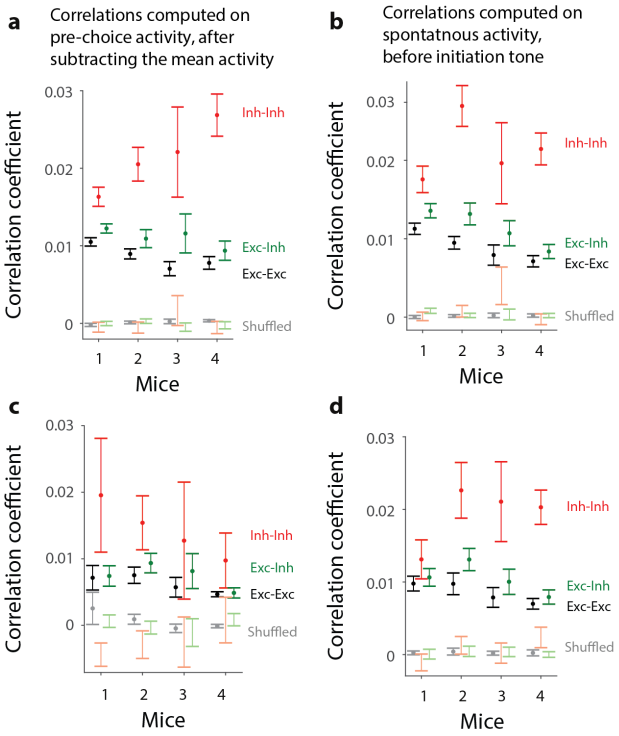
Extended data Figure 3. Classification accuracy on error trials varies across mice but is consistently similar for excitatory and inhibitory populations.

Refers to Fig. 3. **a,** Data from an example animal (48 sessions). **Top**: Classification accuracy of decoders trained on all neurons (left), subsampled excitatory neurons (right, blue trace), and inhibitory neurons (right, red trace). In all cases, classifiers were trained on correct trials; however they were tested either on correct (dark lines: "Correct") or incorrect (dim lines: "Incorrect") trials. Unsaturated lines show performance on shuffled trial labels. Classification accuracy was high, yet lower on incorrect trials, indicating that population activity primarily represents the animal's choice, yet it differs at least slightly for correct and incorrect trials. This reduction was similar for excitatory and inhibitory neurons (blue are red traces are overlapping in the right panel). **Bottom**: As in Fig 3b, histograms reflect the across-trial distribution of go tones and reward delivery. These are variable because the data are aligned to the animal's choice, which was under the animal's control and thus occurred at a variable time relative to go-tone or reward. Left and right panels are the same and are duplicated to facilitate alignment to each corresponding plot above. **b,** Summary across all mice for all neurons (left) and excitatory and inhibitory neurons separately (right). Classifier performance on correct (dark colors) and incorrect (dim colors) trials is shown. Mouse 4 had the largest difference in classification accuracy for correct vs. error trials. As with the single-neuron analysis, this difference likely reflects that the imaging region was slightly posterior within the window for this animal (see Extended Data Fig 2d, legend). Importantly, for all mice, the change in classification accuracy was quite similar for excitatory and inhibitory neurons, indicating that both populations reflect choice vs. stimulus to a comparable degree.



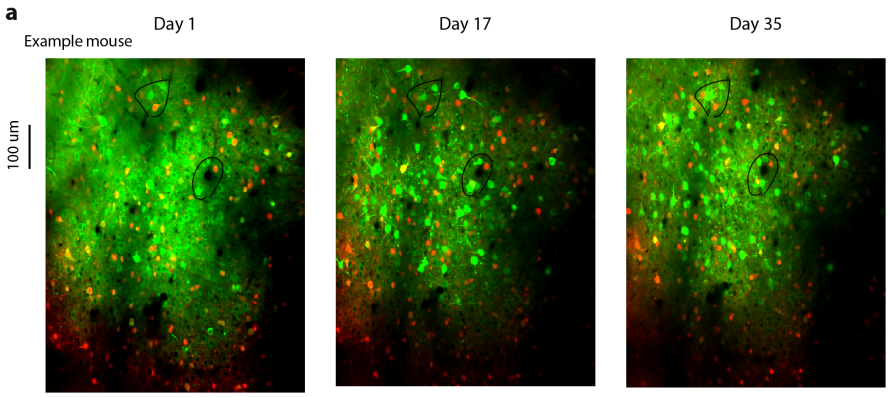
Extended data Figure 4. Additional analyses provide more evidence for similar stability of the choice decoder in excitatory and inhibitory populations.

Refers to Fig. 4. **a,** In an example mouse, population activity that predicts the animal's choice is similarly stable for excitatory and inhibitory neurons during the course of a trial. The vertical axis shows the stability duration for decoders trained at different times during the trial. Stability duration is defined as the width of the testing window over which decoder accuracy is not statistically different (red regions of Fig. 4c) from that obtained by using the same training and testing times (diagonal of Fig. 4a). Error bars: S.E.M. across sessions. Summary data for all mice at training time -97-0ms (dashed line) are shown in Fig. 4d. **b,** Stability duration of the all-neuron decoder (black in panel a) is compared for decoders trained 350ms before the choice (black), and 350ms after the choice (green). Population stability was lower after the choice than before the choice. This may be due to additional events, e.g. reward delivery and repeated licking, which follow the choice. **c**, Another measurement of stability likewise suggests similar temporal stability for excitatory and inhibitory populations. Stability was assessed by measuring the angle between pairs of decoders trained at different points in the trial. If a similar pattern of population activity represents choice from moment t_1 to moment t_2, the choice classifiers trained at these times will be aligned, i.e. the angle between the two classifiers will be small. The colors indicate the angle between pairs of decoders trained at different moments in the trial. Small angles (hot colors) indicate alignment of choice decoders; hence stable activity patterns across neurons. left: all neurons; middle: excitatory neurons (subsampled to match the number of inhibitory neurons); right: inhibitory neurons. As with our other method (Fig. 4), the time course of stability was similar for excitatory and inhibitory neurons.

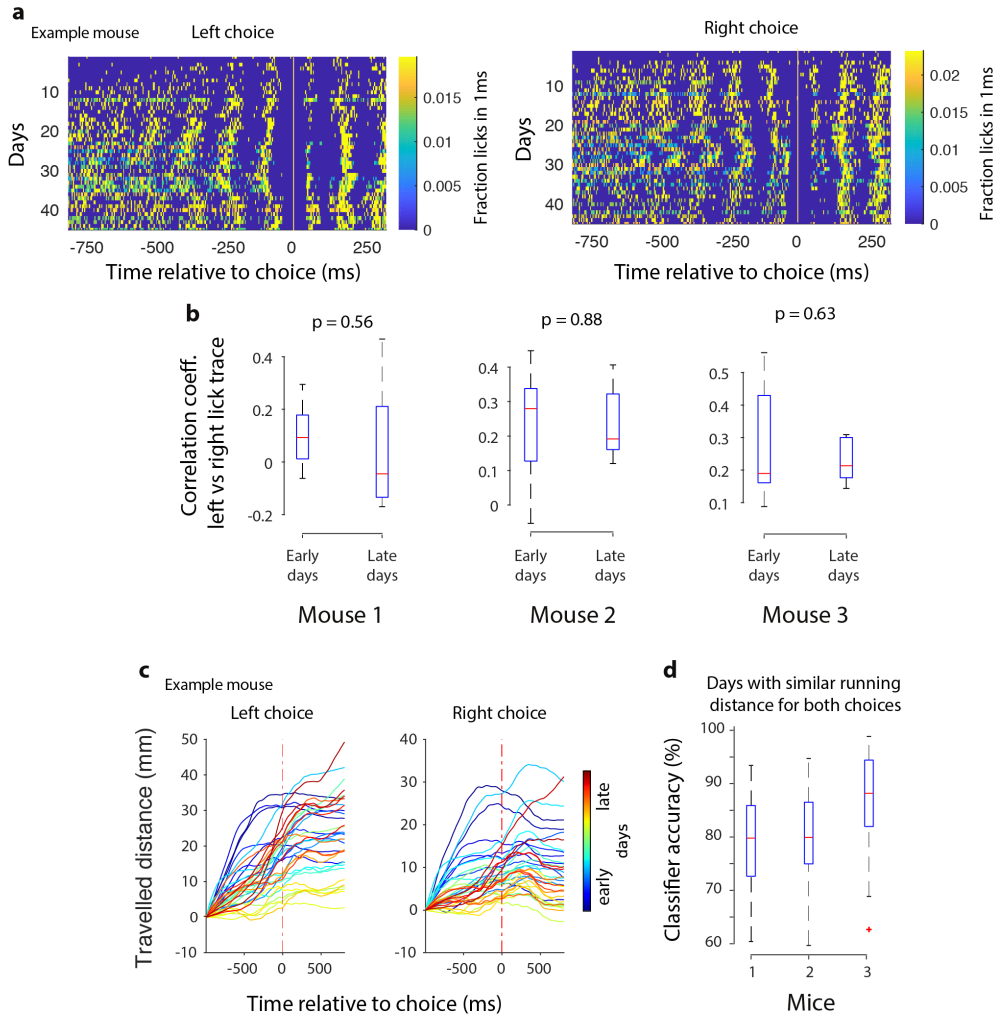


Extended data Figure 5. Higher noise correlations among inhibitory neurons.

Refers to Fig. 4. **a,** Noise correlations for pairs of excitatory neurons (black), pairs of inhibitory neurons (red) or excitatory-inhibitory pairs (green) during the time period 97-0ms before the choice. Signal correlations were removed by subtracting off the mean response from each trial (as in Extended Data Fig 2e). Noise correlations were overall much stronger for inhibitory-inhibitory pairs, and had intermediate values for excitatory-inhibitory pairs. **b,** Same as (a) but for the time period 97-0ms before the trial initiation tone. As in Extended Data Fig. 2f, signal correlations were not present because there is no stimulus and all activity is spontaneous. **c,d,** same as in a,b, except correlations were computed only on those excitatory and inhibitory neurons with the same median spiking activity (Methods).

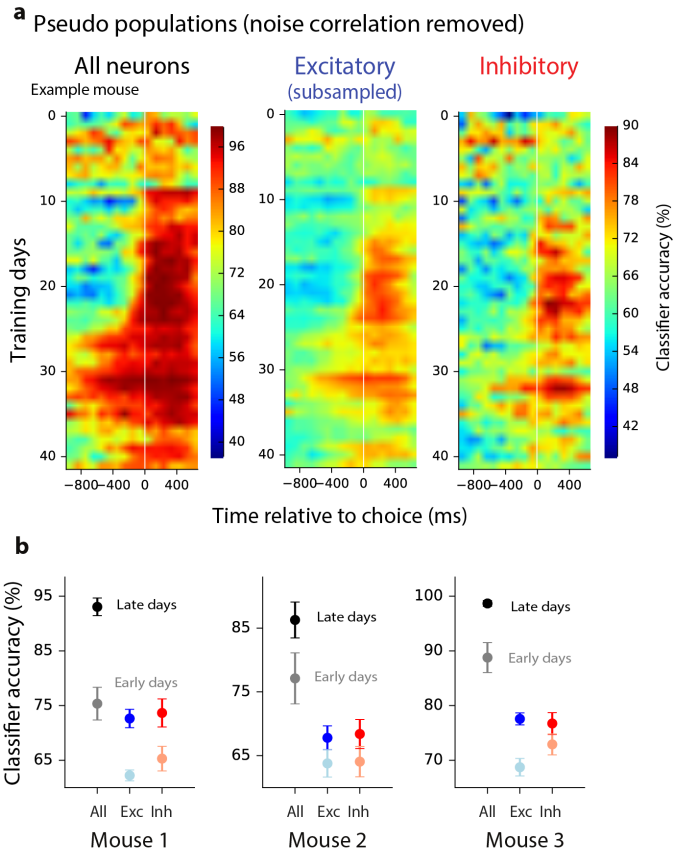


Extended data Figure 6. The same field of view was imaged during learning.
Refers to Fig. 6. **a,** Field of view from three example sessions of a mouse: 1st days of imaging (left), a middle imaging session (middle), and last day of imaging (right). Left to right panels span 60 days, out of which 35 days were imaged. Black circles mark example areas that can be easily matched among the sessions. Each panel is an average projection image of all the frames imaged in the session. Green and red (bleedthrough corrected) images were merged.



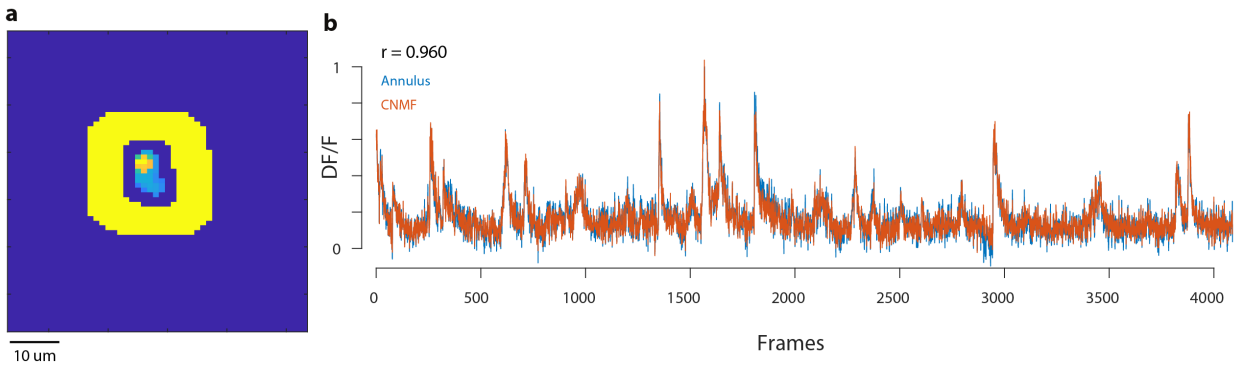
Extended data Figure 7. Further analysis of learning-induced changes in the population activity: changes in licking and running movements are unlikely to account for improved classifier accuracy during learning.

Refers to Fig. 6. **a,** Licking was similar in advance of high rate vs. low rate choices, both early and late in training. Licks that occur before the choice are to the center waterspout, and licks that occur after the choice are to the side waterspouts; example mouse. **b,** Each plot shows the Pearson's correlation coefficient between licking patterns for left and right choices, calculated 250ms before the choice. These correlations were typically similar for early vs. late training days, indicating that animal's licking pattern preceding left vs. right choices did not change drastically over the course learning. **c,** Distance that the animal travelled during the decision (as measured by the rotary encoder on the running wheel) was similar in advance of left vs. right choices; example mouse; each line represents a session (cold colors: early sessions; hot colors: late sessions). **d,** Classifier accuracy (97-0ms before the choice) of the full population was high even when the analysis was restricted to sessions in which the distance travelled was not significantly different (t-test, P>0.05; time 97-0ms before the choice) for left vs. right choices. This analysis was necessary because for some mice in some sessions, there were idiosyncratic differences between the distances travelled in advance of left vs. right choices. In (b) and (d), median (red horizontal line), inter-quartile range (blue box), and the entire range of data (dashed black lines) are shown.



Extended data Figure 8. Further analysis of learning-induced changes in the population activity: the reduction in noise correlations is insufficient to account for the improved classification accuracy during learning.

Refers to Fig. 6. **a,** Classification accuracy for each training session (average of cross-validation samples), for all neurons (left), subsampled excitatory (middle), and inhibitory neurons (right); example mouse. White vertical line: choice onset. This format is the same as Fig. 6a, but here the noise correlations are removed by making pseudo populations. **b,** Summary of each mouse, showing classification accuracy averaged across early (unsaturated colors) vs. late (saturated colors) training days. As in (a), data are based on pseudo-populations in which the noise correlations are removed. The learning-induced improvement in the classifier accuracy in pseudo populations indicates that reduced noise correlations (Fig. 6f) cannot solely account for the enhanced classifier accuracy during learning (Fig. 6a).



Extended data Figure 9. Removing neuropil contamination with CNMF or manually using an annulus leads to the same results.

Refers to Methods section "Neuropil Contamination removal". **a,** An example spatial component in the FOV and its surrounding annulus (yellow). **b,** ΔF/F trace for the same component obtained by manually subtracting the neuropil activity averaged over the annulus region (blue trace) or by using the output of the CNMF processing pipeline (red trace). The two traces look nearly identical as also demonstrated by their high correlation coefficient (r = 0.96; the traces are not denoised). These results demonstrate the ability of the CNMF framework to properly capture neuropil contamination and remove it from the detected calcium traces.

References

1 Wang, X. J. Probabilistic decision making by slow reverberation in cortical circuits. *Neuron* **36**, 955-968 (2002).

2 Lim, S. & Goldman, M. S. Balanced cortical microcircuitry for maintaining information in working memory. *Nat Neurosci* **16**, 1306-1314, doi:10.1038/nn.3492 (2013).

3 Machens, C. K., Romo, R. & Brody, C. D. Flexible control of mutual inhibition: a neural model of two-interval discrimination. *Science* **307**, 1121-1124, doi:10.1126/science.1104171 (2005).

4 Bogacz, R., Brown, E., Moehlis, J., Holmes, P. & Cohen, J. D. The physics of optimal decision making: a formal analysis of models of performance in two-alternative forced-choice tasks. *Psychol Rev* **113**, 700-765, doi:10.1037/0033-295X.113.4.700 (2006).

5 Deneve, S., Latham, P. E. & Pouget, A. Reading population codes: a neural implementation of ideal observers. *Nat Neurosci* **2**, 740-745, doi:10.1038/11205 (1999).

6 Wang, X. J. Decision making in recurrent neuronal circuits. *Neuron* **60**, 215-234, doi:10.1016/j.neuron.2008.09.034 (2008).

7 Znamenskiy, P. *et al. bioRxiv*, doi:10.1101/294835 (2018).

8 Aksay, E. *et al.* Functional dissection of circuitry in a neural integrator. *Nat Neurosci* **10**, 494-504, doi:10.1038/nn1877 (2007).

9 Sohya, K., Kameyama, K., Yanagawa, Y., Obata, K. & Tsumoto, T. GABAergic neurons are less selective to stimulus orientation than excitatory neurons in layer II/III of visual cortex, as revealed by in vivo functional Ca2+ imaging in transgenic mice. *J Neurosci* **27**, 2145-2149, doi:10.1523/JNEUROSCI.4641-06.2007 (2007).

10 Niell, C. M. & Stryker, M. P. Highly selective receptive fields in mouse visual cortex. *J Neurosci* **28**, 7520-7536, doi:10.1523/JNEUROSCI.0623-08.2008 (2008).

11 Kerlin, A. M., Andermann, M. L., Berezovskii, V. K. & Reid, R. C. Broadly tuned response properties of diverse inhibitory neuron subtypes in mouse visual cortex. *Neuron* **67**, 858-871, doi:10.1016/j.neuron.2010.08.002 (2010).

12 Hofer, S. B. *et al.* Differential connectivity and response dynamics of excitatory and inhibitory neurons in visual cortex. *Nat Neurosci* **14**, 1045-1052, doi:10.1038/nn.2876 (2011).

13 Ch'ng, Y. H. & Reid, R. C. Cellular imaging of visual cortex reveals the spatial and functional organization of spontaneous activity. *Front Integr Neurosci* **4**, doi:10.3389/fnint.2010.00020 (2010).

14 Cossell, L. *et al.* Functional organization of excitatory synaptic strength in primary visual cortex. *Nature* **518**, 399-403, doi:10.1038/nature14182 (2015).

15 Ko, H. *et al.* Functional specificity of local synaptic connections in neocortical networks. *Nature* **473**, 87-91, doi:10.1038/nature09880 (2011).

16 Yoshimura, Y., Dantzker, J. L. & Callaway, E. M. Excitatory cortical neurons form fine-scale functional networks. *Nature* **433**, 868-873, doi:10.1038/nature03252 (2005).

17 Ringach, D. L. *et al.* Spatial clustering of tuning in mouse primary visual cortex. *Nat Commun* **7**, 12270, doi:10.1038/ncomms12270 (2016).

18 Cardin, J. A., Palmer, L. A. & Contreras, D. Stimulus feature selectivity in excitatory and inhibitory neurons in primary visual cortex. *J Neurosci* **27**, 10333-10344, doi:10.1523/JNEUROSCI.1692-07.2007 (2007).

19 Hirsch, J. A. *et al.* Functionally distinct inhibitory neurons at the first stage of visual cortical processing. *Nat Neurosci* **6**, 1300-1308, doi:10.1038/nn1152 (2003).
20 Packer, A. M. & Yuste, R. Dense, unspecific connectivity of neocortical parvalbumin-positive interneurons: a canonical microcircuit for inhibition? *J Neurosci* **31**, 13260-13271, doi:10.1523/JNEUROSCI.3131-11.2011 (2011).
21 Li, L. Y. *et al.* Differential Receptive Field Properties of Parvalbumin and Somatostatin Inhibitory Neurons in Mouse Auditory Cortex. *Cereb Cortex* **25**, 1782-1791, doi:10.1093/cercor/bht417 (2015).
22 Runyan, C. A. *et al.* Response features of parvalbumin-expressing interneurons suggest precise roles for subtypes of inhibition in visual cortex. *Neuron* **67**, 847-857, doi:10.1016/j.neuron.2010.08.006 (2010).
23 Khan, A. G. *et al.* Distinct learning-induced changes in stimulus selectivity and interactions of GABAergic interneuron classes in visual cortex. *Nature Neuroscience*, doi:10.1038/s41593-018-0143-z (2018).
24 Yoshimura, Y. & Callaway, E. M. Fine-scale specificity of cortical networks depends on inhibitory cell type and connectivity. *Nat Neurosci* **8**, 1552-1559, doi:10.1038/nn1565 (2005).
25 Goard, M. J., Pho, G. N., Woodson, J. & Sur, M. Distinct roles of visual, parietal, and frontal motor cortices in memory-guided sensorimotor decisions. *Elife* **5**, doi:10.7554/eLife.13764 (2016).
26 Harvey, C. D., Coen, P. & Tank, D. W. Choice-specific sequences in parietal cortex during a virtual-navigation decision task. *Nature* **484**, 62-68, doi:10.1038/nature10918 (2012).
27 Driscoll, L. N., Pettit, N. L., Minderer, M., Chettih, S. N. & Harvey, C. D. Dynamic Reorganization of Neuronal Activity Patterns in Parietal Cortex. *Cell* **170**, 986-999 e916, doi:10.1016/j.cell.2017.07.021 (2017).
28 Raposo, D., Kaufman, M. T. & Churchland, A. K. A category-free neural population supports evolving demands during decision-making. *Nat Neurosci* **17**, 1784-1792, doi:10.1038/nn.3865 (2014).
29 Licata, A. M. *et al.* Posterior Parietal Cortex Guides Visual Decisions in Rats. *J Neurosci* **37**, 4954-4966, doi:10.1523/JNEUROSCI.0105-17.2017 (2017).
30 Hwang, E. J., Dahlen, J. E., Mukundan, M. & Komiyama, T. History-based action selection bias in posterior parietal cortex. *Nat Commun* **8**, 1242, doi:10.1038/s41467-017-01356-z (2017).
31 Funamizu, A., Kuhn, B. & Doya, K. Neural substrate of dynamic Bayesian inference in the cerebral cortex. *Nat Neurosci* **19**, 1682-1689, doi:10.1038/nn.4390 (2016).
32 Morcos, A. S. & Harvey, C. D. History-dependent variability in population dynamics during evidence accumulation in cortex. *Nat Neurosci* **19**, 1672-1681, doi:10.1038/nn.4403 (2016).
33 Song, Y. H. *et al.* A Neural Circuit for Auditory Dominance over Visual Perception. *Neuron* **93**, 940-954 e946, doi:10.1016/j.neuron.2017.01.006 (2017).
34 Pnevmatikakis, E. A. *et al.* Simultaneous Denoising, Deconvolution, and Demixing of Calcium Imaging Data. *Neuron* **89**, 285-299, doi:10.1016/j.neuron.2015.11.037 (2016).

35 Giovannucci, A. *et al.* CaImAn: An open source tool for scalable Calcium Imaging data Analysis. *bioRxiv*, doi:10.1101/339564 (2018).

36 Chen, T. W. *et al.* Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature* **499**, 295-300, doi:10.1038/nature12354 (2013).

37 Sahara, S., Yanagawa, Y., O'Leary, D. D. & Stevens, C. F. The fraction of cortical GABAergic neurons is constant from near the start of cortical neurogenesis to adulthood. *J Neurosci* **32**, 4755-4761, doi:10.1523/JNEUROSCI.6412-11.2012 (2012).

38 Runyan, C. A., Piasini, E., Panzeri, S. & Harvey, C. D. Distinct timescales of population coding across cortex. *Nature* **548**, 92-96, doi:10.1038/nature23020 (2017).

39 Kwan, A. C. & Dan, Y. Dissection of cortical microcircuits by single-neuron stimulation in vivo. *Curr Biol* **22**, 1459-1467, doi:10.1016/j.cub.2012.06.007 (2012).

40 Green, D. M. & Swets, J. A. *Signal detection theory and psychophysics*. (Wiley, 1966).

41 Hofmann, T., Scholkopf, B. & Smola, A. J. Kernel methods in machine learning. *Ann. Statist.* **36**, 1171-1220, doi:10.1214/009053607000000677 (2008).

42 Moreno-Bote, R. *et al.* Information-limiting correlations. *Nat Neurosci* **17**, 1410-1417, doi:10.1038/nn.3807 (2014).

43 Averbeck, B. B., Latham, P. E. & Pouget, A. Neural correlations, population coding and computation. *Nat Rev Neurosci* **7**, 358-366, doi:10.1038/nrn1888 (2006).

44 Musall, S., Kaufman, M. T., Gluf, S. & Churchland, A. K. *bioRxiv*, doi:10.1101/308288 (2018).

45 Gu, Y. *et al.* Perceptual learning reduces interneuronal correlations in macaque visual cortex. *Neuron* **71**, 750-761, doi:10.1016/j.neuron.2011.06.015 (2011).

46 Jeanne, J. M., Sharpee, T. O. & Gentner, T. Q. Associative learning enhances population coding by inverting interneuronal correlation patterns. *Neuron* **78**, 352-363, doi:10.1016/j.neuron.2013.02.023 (2013).

47 Ni, A. M., Ruff, D. A., Alberts, J. J., Symmonds, J. & Cohen, M. R. Learning and attention reveal a general relationship between population activity and behavior. *Science* **359**, 463-465, doi:10.1126/science.aao0284 (2018).

48 Zohary, E., Shadlen, M. N. & Newsome, W. T. Correlated neuronal discharge rate and its implications for psychophysical performance. *Nature* **370**, 140-143, doi:10.1038/370140a0 (1994).

49 Galarreta, M. & Hestrin, S. A network of fast-spiking cells in the neocortex connected by electrical synapses. *Nature* **402**, 72-75, doi:10.1038/47029 (1999).

50 Cohen, M. R. & Maunsell, J. H. Attention improves performance primarily by reducing interneuronal correlations. *Nat Neurosci* **12**, 1594-1600, doi:10.1038/nn.2439 (2009).

51 Liu, B. H. *et al.* Visual receptive field structure of cortical inhibitory neurons revealed by two-photon imaging guided recording. *J Neurosci* **29**, 10520-10532, doi:10.1523/JNEUROSCI.1915-09.2009 (2009).

52 Atallah, B. V., Bruns, W., Carandini, M. & Scanziani, M. Parvalbumin-expressing interneurons linearly transform cortical responses to visual stimuli. *Neuron* **73**, 159-170, doi:10.1016/j.neuron.2011.12.013 (2012).

53 Wehr, M. & Zador, A. M. Balanced inhibition underlies tuning and sharpens spike timing in auditory cortex. *Nature* **426**, 442-446, doi:10.1038/nature02116 (2003).

54 Moore, A. K. & Wehr, M. Parvalbumin-expressing inhibitory interneurons in auditory cortex are well-tuned for frequency. *J Neurosci* **33**, 13713-13723, doi:10.1523/JNEUROSCI.0663-13.2013 (2013).

55 Taniguchi, H. *et al.* A resource of Cre driver lines for genetic targeting of GABAergic neurons in cerebral cortex. *Neuron* **71**, 995-1013, doi:10.1016/j.neuron.2011.07.026 (2011).

56 Madisen, L. *et al.* A robust and high-throughput Cre reporting and characterization system for the whole mouse brain. *Nat Neurosci* **13**, 133-140, doi:10.1038/nn.2467 (2010).

57 Brunton, B. W., Botvinick, M. M. & Brody, C. D. Rats and humans can optimally accumulate evidence for decision-making. *Science* **340**, 95-98, doi:10.1126/science.1233912 (2013).

58 Odoemene, O., Pisupati, S., Nguyen, H. & Churchland, A. K. *bioRxiv*, doi:10.1101/195792 (2017).

59 Tanimoto, N. *et al.* Electroretinographic assessment of rod- and cone-mediated bipolar cell pathways using flicker stimuli in mice. *Sci Rep* **5**, 10731, doi:10.1038/srep10731 (2015).

60 Krishna, V. R., Alexander, K. R. & Peachey, N. S. Temporal properties of the mouse cone electroretinogram. *J Neurophysiol* **87**, 42-48, doi:10.1152/jn.00489.2001 (2002).

61 Busse, L. *et al.* The detection of visual contrast in the behaving mouse. *J Neurosci* **31**, 11351-11361, doi:10.1523/JNEUROSCI.6689-10.2011 (2011).

62 Guizar-Sicairos, M., Thurman, S. T. & Fienup, J. R. Efficient subpixel image registration algorithms. *Opt Lett* **33**, 156-158 (2008).

63 Vogelstein, J. T. *et al.* Fast nonnegative deconvolution for spike train inference from population calcium imaging. *J Neurophysiol* **104**, 3691-3704, doi:10.1152/jn.01073.2009 (2010).

64 Rudy, B., Fishell, G., Lee, S. & Hjerling-Leffler, J. Three groups of interneurons account for nearly 100% of neocortical GABAergic neurons. *Dev Neurobiol* **71**, 45-61, doi:10.1002/dneu.20853 (2011).

65 Kimmel, D., Elsayed, G. F., Cunningham, J. P., Rangel, A. & Newsome, W. T. Encoding of value and choice as separable, dynamic neural dimensions in orbitofrontal cortex. *Cosyne* (2016).