

Predicting Perceptual Decisions Using Visual Cortical Population Responses and Choice History

Anna Ivic Jasper,^{1*} Seiji Tanabe,^{1*} and Adam Kohn^{1,2,3}

¹Dominick Purpura Department of Neuroscience, ²Department of Ophthalmology and Visual Sciences, and ³Department of Systems and Computational Biology, Albert Einstein College of Medicine, Bronx, New York 10461

Our understanding of the neural basis of perceptual decision making has been built in part on relating co-fluctuations of single neuron responses to perceptual decisions on a trial-by-trial basis. The strength of this relationship is often compared across neurons or brain areas, recorded in different sessions, animals, or variants of a task. We sought to extend our understanding of perceptual decision making in three ways. First, we measured neuronal activity simultaneously in early [primary visual cortex (V1)] and midlevel (V4) visual cortex while macaque monkeys performed a fine orientation discrimination perceptual task. This allowed a direct comparison of choice signals in these two areas, including their dynamics. Second, we asked how our ability to predict animals' decisions would be improved by considering small simultaneously-recorded neuronal populations rather than individual units. Finally, we asked whether predictions would be improved by taking into account the animals' choice and reward histories, which can strongly influence decision making. We found that responses of individual V4 neurons were weakly predictive of decisions, but only in a brief epoch between stimulus offset and the indication of choice. In V1, few neurons showed significant decision-related activity. Analysis of neuronal population responses revealed robust choice-related information in V4 and substantially weaker signals in V1. Including choice- and reward-history information improved performance further, particularly when the recorded populations contained little decision-related information. Our work shows the power of using neuronal populations and decision history when relating neuronal responses to the perceptual decisions they are thought to underlie.

Key words: choice signals; perceptual decision making; visual cortex

Significance Statement

Decades of research has provided a rich description of how visual information is represented in the visual cortex. Yet how cortical responses relate to visual perception remains poorly understood. Here we relate fluctuations in small neuronal population responses, recorded simultaneously in primary visual cortex (V1) and area V4 of monkeys, to perceptual reports in an orientation discrimination task. Choice-related signals were robust in V4, particularly late in the behavioral trial, but not in V1. Models that include both neuronal responses and choice-history information were able to predict a substantial portion of decisions. Our work shows the power of integrating information across neurons and including decision history in relating neuronal responses to perceptual decisions.

Introduction

One way to understand how the activity of sensory neurons contributes to perception is to measure neuronal responses and per-

ceptual reports simultaneously, and assess their covariation (Britten et al., 1992; Parker and Newsome, 1998). This approach has revealed that trial-to-trial fluctuations in individual visual cortical neurons are weakly predictive of animals' choices in a perceptual task. Such neuron–choice correlations have been observed in many cortical areas (Nienborg et al., 2012; Crapse and Basso, 2015). Theory suggests that the pattern of neuron–choice correlations can be used to infer how sensory representations are “read out” to reach perceptual decisions (Shadlen et al., 1996; Haefner et al., 2013; Pitkow et al., 2015). Although inactivation experiments have shown that neuron–choice correlations do not indicate that a neuron (or, rather, the area in which it resides) is necessary for behavior (Cohen and Newsome, 2009; Chen et al.,

Received Jan. 7, 2019; revised June 12, 2019; accepted June 18, 2019.

Author contributions: S.T. and A.K. designed research; S.T. performed research; A.I.J. analyzed data; A.I.J., S.T., and A.K. wrote the paper.

This work was supported by funding from the NIH (NEI EY016774 and EY028626), the Simons Collaboration on the Global Brain (364994 and 542999), and an Irma T Hirschl Career Scientist Award. We thank Amin Zandvakili for technical assistance, and Christopher Henry and James McIntosh for comments on an earlier version of this paper.

The authors declare no competing financial interests.

*A.I.J. and S.T. contributed equally to this work.

Correspondence should be addressed to Adam Kohn at adam.kohn@einstein.yu.edu.

<https://doi.org/10.1523/JNEUROSCI.0035-19.2019>

Copyright © 2019 the authors

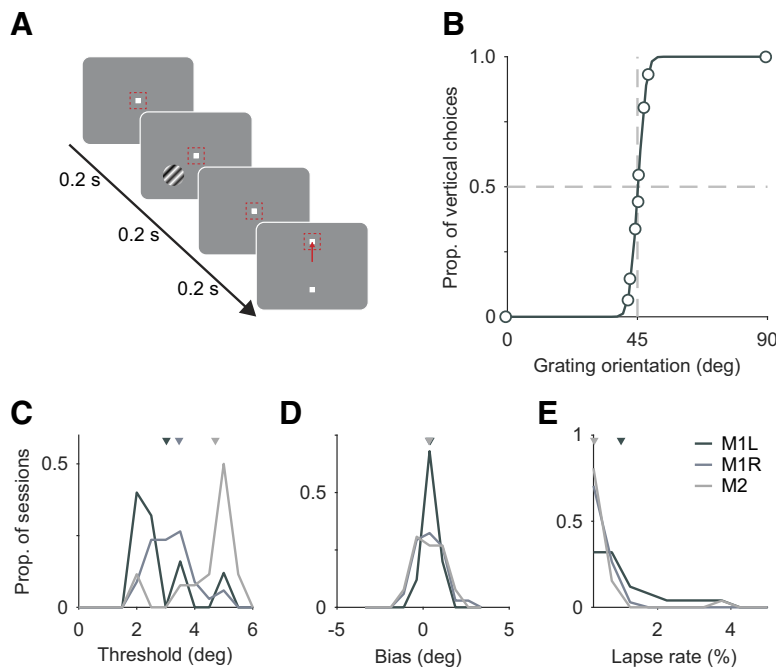


Figure 1. Behavioral task and psychophysical performance. **A**, Task design. **B**, Example psychometric function from one session in M1L. **C–E**, Psychometric thresholds (**C**), biases (**D**), and lapse rates (**E**) for all three cases. Triangles in **C–E** indicate mean. Prop., Proportion.

2016; Katz et al., 2016), correlative approaches remain an important ingredient for understanding perceptual decision making and for interpreting the consequences of “causal” perturbations (Smolyanskaya et al., 2015; Lakshminarasimhan et al., 2018; Yu and Gu, 2018).

To date, much of our progress in understanding neuron–choice correlations has relied on a stereotyped experimental approach: isolating individual neurons in an area of interest, tailoring the task in each session to the functional properties of the recorded neuron, and relating neuronal responses to perceptual reports using choice probability analysis (Britten et al., 1996). This approach has several limitations, which we sought to overcome in the current study.

First, tailoring the task (e.g., direction of motion in a motion discrimination task) to each recorded neuron complicates the comparison of choice signals across cells. Comparisons of choice signals in different brain areas are even more problematic, because these are often measured in different sessions, animals, or even tasks. Behavioral strategies can vary strongly across similar tasks and even across animals performing identical tasks, and these variations in strategy likely influence choice information (Goris et al., 2017; Bondy et al., 2018). To allow a direct comparison of choice signals in early and midlevel primate visual cortex, we therefore simultaneously recorded neurons with overlapping spatial receptive fields (RFs), in primary visual cortex (V1) and area V4 while monkeys performed a fine orientation discrimination task. The task was not adjusted to the functional properties of the sampled neurons, except for placing the behavioral stimuli in the aggregate spatial RF.

Second, with few exceptions (Bondy et al., 2018), neuronal correlates of perceptual decisions have been studied in single neurons, although sensory information is encoded by neuronal populations (Averbeck et al., 2006; Kohn et al., 2016). Here we record simultaneously from small neuronal populations (up to ~30 neurons) in V1 and V4 and read out their responses to

predict choice. We compare the view they provide about choice representation to that afforded by single neurons.

Finally, perceptual decisions are influenced not only by sensory evidence, but also by a host of other factors, including the history of choices and rewards (Busse et al., 2011; Fründ et al., 2014; Abrahamyan et al., 2016; Akrami et al., 2018). These history effects are well documented in humans and other animals but rarely considered when relating neuronal responses to choices (but see Dodd et al., 2001; Lueckmann et al., 2018). Here we assess how consideration of choice-history information improves our ability to predict choices on the current trial using neuronal responses, and whether the choice information available in neuronal responses directly reflects choice history.

Materials and Methods

Subjects. We used two male, adult cynomolgus macaques (*Macaca fascicularis*). All procedures were approved by the Institutional Animal Care and Use Committee of the Albert Einstein College of Medicine and were in compliance with the guidelines set forth in the National Institutes of Health *Guide for the Care and Use of Laboratory Animals*.

Animals were first familiarized with a restraining chair (Crist Instruments) and then implanted with a titanium headpost. Implantation was performed under isoflurane anesthesia, following strict sterile procedures. A postoperative analgesic (buprenorphine or flunixin) and antibiotic (enrofloxacin) were provided. Animals recovered for at least 6 weeks before the initiation of behavioral training.

Task and behavior. Animals viewed a calibrated monitor (Iiyama; 1024 × 768 resolution; 100 Hz refresh) from a distance of 57 cm. Visual stimuli were generated using custom OpenGL software (Expo; <http://sites.google.com/a/nyu.edu/expo>), which also controlled task contingencies. Eye position was monitored using a video eye-tracking system (SR Research) with a sampling rate of 1 kHz. Animals were positively reinforced with a drop of liquid reward.

Animals performed a two-alternative forced-choice orientation discrimination task. Trials began with subjects fixating a bright spot (0.15 × 0.15 deg²; 80 cd/m²) on a gray background (40 cd/m²). After 0.2 s, we presented a drifting, sinusoidal grating in the parafoveal visual field, for 0.2 s (Fig. 1A). Gratings were presented at full contrast, with a drift rate of 6 Hz and spatial frequency of 2–4 cpd; 0° gratings drifted downward; 90° gratings rightward. 0.2 s after stimulus offset, the fixation spot was replaced with two targets on the vertical meridian (eccentricity of 2.4–5.6°); the animal reported its choice by making a saccade to one of the targets. All orientations that were closer to vertical than 45° were associated with the top target; orientations closer to horizontal than 45° were associated with the bottom target. The range of orientations was adjusted for each animal to span perceptual threshold, and included the perceptually ambiguous stimulus (45°) which was rewarded randomly. In most sessions, the probability of presenting a 45° orientation was twice that of the other orientations. Reward was doubled after three consecutive correct responses; reward was reset to its base volume after an incorrect response. The intertrial interval was 2 s, but was extended following an erroneous choice. Eye position had to stay within a 1.4 × 1.4° window until the appearance of the targets, else the trial was aborted and discarded.

We measured the animal’s performance by fitting a cumulative Gaussian function to the psychometric function, using maximum likelihood for a Bernoulli process. We defined the bias as the difference between the

mean of the Gaussian and 45° , and the threshold as the Gaussian SD, equivalent to the point where the animal chooses a correct answer in 84% of the trials. Only sessions that had a threshold $\leq 6^\circ$ and a bias $\leq 3^\circ$ were considered further (82% of the available sessions exceeded these criteria).

Recording. Training continued until the subjects reached asymptotic performance, after which we implanted microelectrode arrays (Blackrock Systems) in V1 and V4. We targeted the arrays to matching retinotopic locations, relying on anatomical markers and previous mapping studies (Van Essen et al., 1984; Gattass et al., 1988). Each microelectrode array had a 6×8 arrangement (0.4 mm spacing; 1 mm electrode length). After array insertion, we sutured the dura over the arrays and covered it with a gelatin film. The craniotomy was filled with gelfoam or silicone elastomer, and covered with titanium mesh.

Extracellular voltage signals were filtered between 250 Hz and 7.5 kHz. Events were recorded at 30 kHz sampling rate when the extracellular voltage exceeded a user-defined threshold. Local field potentials on each channel were also recorded, although this signal was not used in this study. We sorted spikes off-line using Offline Sorter (Plexon) or custom MATLAB code (R. Kelly and M. A. Smith, University of Pittsburgh).

On the first days of recording, we mapped the spatial RFs of the sampled neurons, using small gratings (0.5° diameter, 2 cpd, 6 Hz drift rate, 100% contrast, 0.35 s presentation with 0.05 s interstimulus interval) presented at different spatial locations and orientations while the animal fixated. We then chose the size of the behavioral stimuli to cover the RFs of the V1 and the V4 populations, but not to impinge on the fixation spot. The animals were retrained for ~ 1 week at the new stimulus location (usually $1\text{--}3^\circ$ from the originally trained location).

Choice analyses. To ensure recording stationarity, we only analyzed units with a Fano factor ≤ 1.3 (81.2% of recorded units). In addition, unless otherwise noted, we required that analyzed units fired at least 2 spikes/s during stimulus presentation, on trials involving the perceptually ambiguous stimulus (45°) and directly adjacent stimulus orientations (typically $1\text{--}2^\circ$ above and below 45°). Almost half of the units (41.2%) meeting the stationarity criterion also met this additional responsiveness criterion.

We computed responsivity, single-neuron variability and noise correlations using spike counts measured 0–250 ms after stimulus onset. For noise correlations, we Z-scored the spike counts for each unit and each stimulus orientation separately, and computed the Pearson correlation coefficient from these Z-scored spike counts (Bair et al., 2001; Kohn and Smith, 2005). For measurements of correlations, we only included orientations for which the unit fired on average at least 2 spikes/s.

We assessed the relationship between single neuron responses and the animals' choices using receiver operating curve (ROC) analysis [computing choice probability (CP); Britten et al., 1996]. For all choice-related analysis, we used trials for the 45° stimulus as well as the two neighboring stimuli (always within 3° of 45°), and chose an equal number of trials with vertical and horizontal choices for each stimulus. Our results were similar when based only on responses to the 45° stimulus (data not shown). We only considered sessions that had at least 10 trials for each stimulus and choice, and a minimum total of 60 trials. 67% of the sessions that passed our behavioral criteria also provided a sufficient number of trials for choice-related analyses.

We computed CP using spike counts measured 0–250 ms after stimulus onset. We then Z-scored the neuronal responses to each stimulus separately, and computed CP on the combined distributions, as in the study by Britten et al. (1996). For each neuron, we used the slope of the orientation tuning curve at 45° , defined using all of the behavioral stimuli except 0° and 90° , to decide whether stronger neuronal responses would be associated with vertical or horizontal choices. To assess statistical significance, we created 1000 permutations by shuffling the choices of all trials, and computed CP on these permuted data. Measured CPs that lay outside the 2.5–97.5% interval of this permuted distribution were considered statistically significant. To assess CP dynamics, we computed the CP using an identical procedure, but measuring responses in sliding 100 ms windows with 50% overlap.

To compute the neurometric curve, we first performed ROC analysis on the distribution of spike counts for the ambiguous 45° stimulus compared with each of the other stimuli, excluding 0° and 90° (similar to the

neuron/anti-neuron procedure described by Britten et al., 1992). Responses were measured 0–250 ms after stimulus onset. To quantify neuronal threshold, we fit a cumulative Gaussian function to the results of the ROC analysis, and defined threshold as the SD of that function.

To assess the choice signal present in neuronal populations, we used binomial logistic regression. As with CP analysis, we selected a balanced number of choices for each stimulus condition, Z-scored the spike counts for each stimulus orientation separately, and then fit the data together. We fit the model using “cvglmnet” provided by the glmnet-package (Qian et al., 2013). We split our dataset into training and test datasets, and assessed performance on a pair of left out trials (consisting of one of each choice from the same orientation). Each trial was used at most once for assessing performance. To reduce overfitting, we used Lasso regularization, choosing the strength of the penalty term using 10-fold cross validation on the training set. Specifically, the weight of the penalty term was set to minimize model deviance.

Even with regularization, we found that the model frequently overfit when we used the full measured population. We therefore applied an alternative strategy. We rank-ordered the neurons based on their predictive performance (measured using binomial logistic regression models and the same leave-one-pair-out cross-validation procedure described above). We then fit a regularized binomial logistic regression model using the neuron with best predictive performance, and then fit a new model to the two best neurons, and so on. When prediction accuracy did not increase for two consecutive increments of population size, we defined the relevant population for that session as that which had provided the maximal prediction accuracy. Performance was again measured by leave one-pair-out cross validation. We used the same cross-validation folds for all increments.

We assessed the statistical significance of model performance by first randomly permuting the choices across trials, within each stimulus condition of each session. We then calculated the prediction accuracy for each neuron, and fit the population model as described above. Cross-validation folds were not shared between the ranking procedure and fitting the population responses, because these were also not shared in the original data. Note that because performance was defined as the maximum observed for populations of increasing size, chance performance for the randomly permuted data was $>50\%$. We repeated this procedure 1000 times for each session, and defined performance of a model based on the measured responses to be significant if its performance exceeded the 950th best model applied to permuted data.

To be sure our results did not depend strictly on the use of logistic regression, we also predicted behavioral choices from neuronal population responses using a support vector machine with a linear kernel (using libsvm; Chang and Lin, 2011) and found similar results (data not shown; Astrand et al., 2014).

Choice-history analysis. We estimated the influence of choice history on behavior using a probabilistic choice model (Busse et al., 2011; Abrahamyan et al., 2016). The model contained a matrix of 10 independent variables for each trial, and a constant (bias term). There were eight sensory parameters, one for each stimulus orientation (excluding the 45° orientation). On each trial, the relevant stimulus variable was either set to -1 (for stimuli $<45^\circ$) or 1 (for stimuli $>45^\circ$), or to 0 if that stimulus was not presented on that trial. There were two choice-history variables, capturing the choice on the previous trial. The first choice-history variable was set to 1 if the animal received a reward and made a vertical choice in the previous trial, -1 if the animal received a reward and made a horizontal choice in the previous trial and 0 if the animal was not rewarded in the previous trial. The second choice-history variable was set to 1 if the animal was not rewarded and made a vertical choice in the previous trial, -1 if the animal was not rewarded and made a horizontal choice in the previous trial and 0 if the animal was rewarded in the previous trial.

We used a binomial logistic regression model with Lasso regularization to relate these variables to the animals' choices. We fit the model (using “cvglmnet”; Qian et al., 2013) to a training dataset (80% of trials) to estimate the weights and the regularization parameter (chosen through 10-fold cross-validation on the training set). We used the remaining 20% of trials to estimate performance. This procedure was re-

peated 100 times, creating 100 folds per session; the reported values are the average of these 100 folds. To test whether the choice-history model performed better than a model that does not have the two choice-history variables, we also computed the likelihood for a “no-history” model. For each fold we computed twice the log likelihood ratio of the two models and compared the value to a χ^2 statistic with two degrees of freedom (Abrahamyan et al., 2016). We chose an α -level of 0.05 to reject the null hypothesis that the choice-history model is no better than the no-history model.

To predict choices using a model with both choice history and neuronal responses, we used the same ranking of neurons as in the population analysis described above. However, this time we fit the regression model to the most informative neuron including the two choice-history variables. We then added the next most informative neuron and fit a new model to these two neurons and the two choice-history variables, and so on. We stopped when prediction accuracy did not increase for the two consecutive increments of population size. Performance was measured using a leave-one-pair-out cross-validation procedure. We used the same cross-validation folds for each population size considered.

We tested whether choice signals in neuronal responses were because of choice-history effects by repeating our choice analyses, but considering separately trials in which the animal had previously made a vertical or horizontal choice. For these analyses, we required at least 10 trials of each choice for each stimulus (45° stimulus and one or two adjacent orientations) and at least 40 trials in total. We could not perform this analysis conditioning on reward outcome (the second aspect of choice history used in our choice model) because the animals’ success rate ($\sim 80\%$) produced too few error trials.

We found that in a small fraction of trials (8.3%) animals had been allowed to complete trials in which their gaze briefly left the fixation window. These trials were excluded from our neuron–choice analyses. However, as animals had indicated their choice on these trials and received reward when correct, these trials were relevant for assessing choice and reward history effects. We therefore included these trials in our choice-history analysis, when they were followed by a valid trial (i.e., in which gaze remained within the fixation window).

Microsaccade detection. To control for contamination of the recorded responses by eye movements, we reanalyzed our data after discarding trials in which at least one microsaccade was detected in an epoch beginning -150 ms before stimulus onset, and extending to the end of each trial. Microsaccade detection was performed using a method based on the study by Horwitz and Albright (2003). In brief, we smoothed the eye position time series with a Gaussian filter kernel ($SD = 6$ ms) and computed its derivative. A microsaccade was defined as an event with velocity of $\geq 10^\circ/s$ lasting for at least 8 ms. The vast majority of detected events ($>99\%$) had an amplitude $>0.1^\circ$. Previous work suggests that $\sim 90\%$ of the events of this amplitude detected with video tracking are true microsaccades (Kimmel et al., 2012).

To ensure we could still detect a relationship between neuronal responses and choice in the remaining trials, we required that the session still contained at least 10 trials of each choice for each stimulus, and at least 40 trials in total. These criteria left us with 41 sessions of data for the CP analysis, 27 sessions for the V1 population analysis, and 28 for the V4 population analysis, with an average of 72.3 ± 28.2 trials used per session. We compared the results of analyzing trials without microsaccades with analyses performed on an equal number of trials containing microsaccades, to ensure equal statistical power for the two analyses.

Statistical analyses. All indications of variability are SDs, unless otherwise indicated.

Results

We trained two male macaque monkeys to perform a fine orientation discrimination task (Fig. 1A; see Materials and Methods). We presented a sinusoidal drifting grating in the parafoveal visual field for 200 ms, beginning 200 ms after the establishment of fixation. Two choice targets appeared 200 ms after stimulus offset. The animal indicated its decision by making a saccade to one of the targets. Correct choices (upward saccades for stimuli $>45^\circ$,

Table 1. Behavioral data

	M1L	M1R	M2
No. of trials per session	1021 \pm 412 (SD)	669 \pm 262	465 \pm 115
Threshold	3.02 \pm 0.98	3.46 \pm 0.74	4.97 \pm 1.23
Lapse rate	1.0 \pm 1.03%	0.23 \pm 0.39%	0.27 \pm 0.75%
Bias	0.42 \pm 0.36°	0.35 \pm 0.75°	0.39 \pm 0.87°
Trials available for choice-related analysis	342 \pm 134	262 \pm 52	147 \pm 39
No. of sessions analyzed	25	34	26

termed *vertical choices*, and downward for stimuli $<45^\circ$, termed *horizontal choices*) were rewarded with a drop of liquid. Trials in which the stimulus orientation was 45° were rewarded randomly.

Both animals became expert at the task, as illustrated with the psychometric function for a sample session in Figure 1B. To quantify the animals’ performance, we measured the bias, threshold and lapse rate for each session (see Materials and Methods; Table 1). On average, thresholds were $3\text{--}5^\circ$ (Fig. 1C), biases were negligible (Fig. 1D), and lapse rates were $<1\%$ (Fig. 1E).

While animals performed the task, we recorded from neuronal populations in V1 and V4 using two 48-channel arrays. After completing recordings in the left hemisphere of one of the animals [Monkey 1 L (M1L)], we retrained him to perform the task in left visual hemifield and implanted the right hemisphere (M1R), yielding three datasets. The spatial RFs of the recorded V1 and V4 populations in all three cases were overlapping (Fig. 2A). We chose the size (1° for M1L, 4° for M1R, and 5° for M2) and position of the behavioral stimuli to cover the RFs of the majority of the recorded units (Fig. 2A, yellow circles). In one case, the yield of V1 neurons was small (M1L) and, additionally, the stimulus was not well placed in the RF of these neurons. This V1 dataset was thus excluded.

In total, we recorded from 891 V1 units and 1957 V4 units, of which 385 V1 units and 568 V4 units, met our inclusion criterion (see Materials and Methods). The term “units” refers to well isolated single units (26% of units, having a signal-to-noise ratio >3.5 ; Kelly et al., 2007) and to multiunit clusters consisting of waveforms from several neurons. We found no notable differences between results for single units and multiunits, and thus considered the two types of recording together (for detailed comparison of the response statistics of our single unit and multiunit recordings, see Wissig and Kohn, 2012).

The behavioral stimuli evoked measurable firing rates in the analyzed units, but because stimuli were not tailored to each cell’s preference, these rates were modest (Fig. 2B; Table 2). Neuronal responses had low trial-to-trial variability, with Fano factors close to 1 in both V1 and V4 (Table 2). Spike count correlations (r_{sc} ; Zohary et al., 1994; Cohen and Kohn, 2011) were slightly higher between pairs of neurons in the same cortical area (V1–V1: 0.023 ± 0.152 ; V4–V4: 0.009 ± 0.146) than for inter-areal pairs (0.005 ± 0.142 ; for further details, see Table 2).

Single neuron choice signals

To determine whether the recorded units encode a choice-related signal, we computed CP (Britten et al., 1996) for each unit, using spike counts measured 0–250 ms after stimulus onset (extending slightly beyond the 200 ms stimulus presentation to account for neuronal response latency). CP measures the degree to which an ideal observer would be able to predict the animal’s choices on a given trial, using the observed spike count. A CP of 0.5 indicates chance performance. Values >0.5 indicate that stronger neuronal responses are associated with choices aligned with the neuron’s preferred stimulus; values <0.5 indicate the counter-intuitive

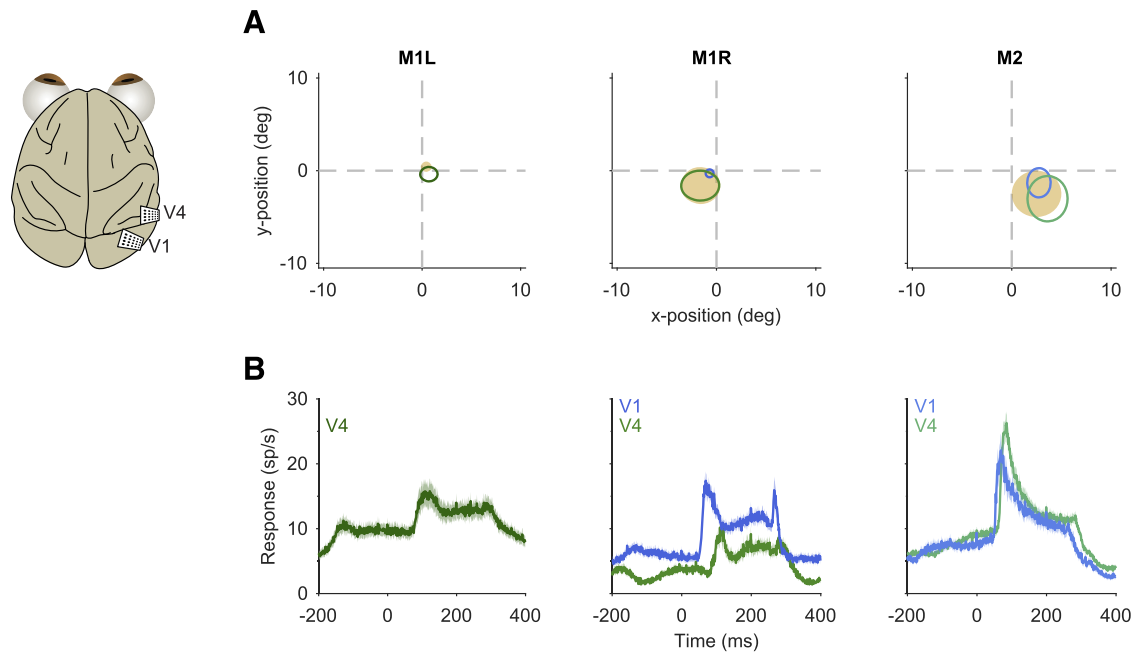


Figure 2. V1 and V4 physiology. **A**, Population spatial RFs for all three cases. Lines indicate the 75% contour line of a 2-dimensional Gaussian fit to the aggregate RF. Blue, V1; green, V4. Stimulus position and size is indicated by yellow shading. **B**, PSTHs for all three cases. PSTHs were calculated from responses to all behavioral stimuli, using all neurons included in choice-related analyses. Zero milliseconds indicates stimulus onset.

Table 2. Physiological data

	M1L	M1R	M2
No. of neurons			
V1	—	178	207
V4	127	56	385
Size of simultaneously recorded populations			
V1	—	0–11 (5.2 ± 2.8)	2–23 (8.0 ± 5.3)
V4	0–11 (5.1 ± 3.0)	0–4 (1.6 ± 1.2)	6–28 (14.8 ± 7.4)
Firing rate V1			
Fixation	—	6.1 ± 7.53 sp/s	6.91 ± 8.35 sp/s
Stimulus	—	10.5 ± 9.23 sp/s	12.38 ± 13.93 sp/s
Firing rate V4			
Fixation	9.17 ± 9.7 sp/s	2.92 ± 2.42 sp/s	7.15 ± 7.06 sp/s
Stimulus	12.16 ± 13.85 sp/s	5.75 ± 4.39 sp/s	13.69 ± 15.26 sp/s
Fano factor			
V1	—	1.0 ± 0.2	1.0 ± 0.2
V4	1.0 ± 0.2	1.1 ± 0.2	1.0 ± 0.2
V1 versus V4 (<i>t</i> test)	—	$p = 0.01$	$p = 0.57$
Spike count correlations			
V1	—	0.056 ± 0.162 $p < 0.001, N = 4386$	0.008 ± 0.144 $p < 0.001, N = 9417$
V4	0.010 ± 0.114 $p < 0.001, N = 3224$	0.035 ± 0.125 $p < 0.001, N = 325$	0.009 ± 0.149 $p < 0.001, N = 28985$
V1–V4	—	0.014 ± 0.122 $p = 0.002, N = 2675$	0.005 ± 0.142 $p < 0.001, N = 32310$
V1 vs V4 (<i>t</i> test)		$p = 0.02$	$p = 0.87$
V1 vs V1V4		$p \leq 0.001$	$p = 0.03$
V4 vs V1V4		$p = 0.002$	$p = 0.0007$

relationship that stronger responses predict choices opposite to the cell's preference.

We found little choice-related activity in individual V1 or V4 units (Fig. 3). The mean CP was only statistically different from chance in V4 of one animal (M1L; $p = 0.004$, one sample *t* test). For all other datasets it was indistinguishable from chance in both V1 (Fig. 3, blue) and V4 (green; $p > 0.05$; one sample *t* test).

Weak mean CPs might arise because the assumed relationship between neuronal tuning and choice is invalid, rather than because of an absence of choice-related signals. That is, if choice signals are not related to tuning preference in the expected manner (e.g., assuming that stronger responses in a vertical-prefering neuron result in more vertical choices), the measured CP would be equally likely to be above or below 0.5, resulting in a

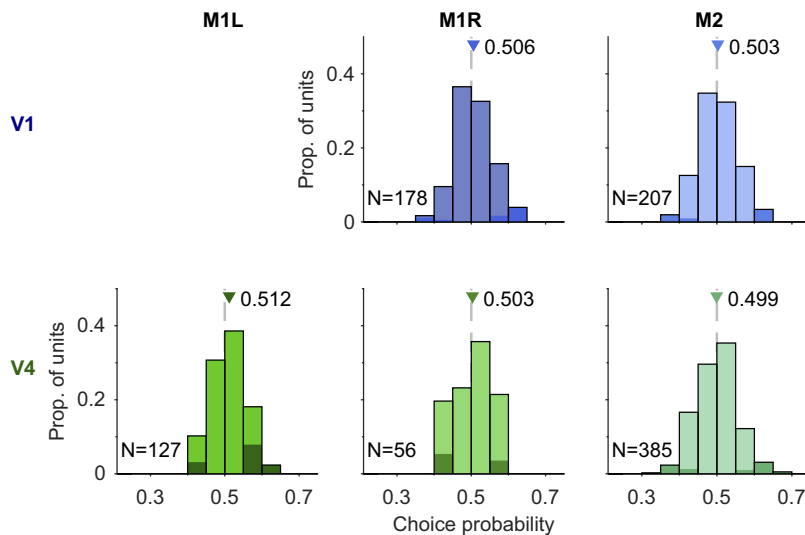


Figure 3. Distribution of choice probabilities for V1 (top row, blue) and V4 (bottom row, green). Dark filled bars indicate units whose CP was significantly different from 0.5 (see Materials and Methods). Triangle indicates the mean CP. CP is based on responses measured 0–250 ms after stimulus onset. Prop., Proportion.

mean near 0.5. To assess this possibility, we evaluated whether the CP of each unit was different from 0.5 using a permutation test. We found only a small proportion of units, 6.7% in V1 and 9.7% in V4 (shaded cases in Fig. 3), with statistically significant choice signals, a proportion not notably larger than expected by chance given our statistical criterion (α level = 0.05).

Another possible explanation for our weak CP values is that the representation of choice in V1 and V4 is dynamic (Britten et al., 1996; Nienborg and Cumming, 2009; Shiozaki et al., 2012; Wimmer et al., 2015; Goris et al., 2017), and perhaps more prevalent outside our chosen analysis window (0–250 ms after stimulus onset). We thus performed CP analysis in sliding temporal windows of 100 ms with 50% overlap. This analysis revealed significant CP in V4, during the epoch between stimulus offset and the appearance of the choice targets (Fig. 4, green; ranging from 0.514 to 0.523 in 150–400 ms in M1L; 0.523 in 200–300 ms and falling significantly below 0.5, to 0.48, in the 300–400 ms window in M1R; and 0.51 in 200–300 ms for M2). During these epochs, we also observed more neurons with significant CP values, with a peak of 23.6% for the 250–350 ms epoch in M1L. We found no significant choice signal in V1 of either animal, in any epoch (mean values between 0.49 and 0.51; percentage of significant cases varied from 2.4 to 9.0%; Fig. 4, blue). Thus, there is choice-related information in some V4 cells, but it occurs in the epoch between stimulus offset and the appearance of the choice targets.

Finally, we considered that the choice signals in our recordings might be weak because we analyzed all responsive neurons provided by the array and had each animal perform an identical task in all sessions. Our approach provided many neurons whose functional properties, namely, orientation tuning preference and selectivity, were largely irrelevant for the task. Because choice signals are often found to be strongest in neurons with greatest sensitivity for task stimuli (Nienborg et al., 2012), the inclusion of many task-irrelevant neurons could result in weaker choice signals in our data.

To test this possibility, we assessed the relationship between the threshold of each unit for the behaviorally-relevant stimulus orientations (the inverse of sensitivity), and the strength of the choice signal. We first calculated a neurometric function for each unit, as shown for an example unit in Figure 5A. Specifically, we

measured the discriminability between responses to the 45° stimulus and those to each of the other behavioral stimuli (except the 0° and 90° orientations), using ROC analysis. We then defined neuronal threshold as the SD of a cumulative Gaussian fit to these discriminability values. As expected, given that the behavioral stimuli were not adjusted to match neuronal preferences, the mean threshold was high in both V1 ($108.9 \pm 87.9^\circ$, $N = 385$) and V4 ($103.1 \pm 84.7^\circ$, $N = 568$).

To assess whether neurons sensitive to variations in orientation near 45° carried stronger choice signals, we compared CP values in neurons with low and high thresholds. We quantified choice information as the deviation of neuronal CP from 0.5, termed here absCP, to capture the strength of choice signals, without assuming a relationship to stimulus preference. During stimulus presentation (0–200 ms after stimulus onset), the mean

absCP for neurons with higher threshold ($\geq 50^\circ$, V1: 0.038 ± 0.029 , $N = 255$, V4: 0.038 ± 0.031 , $N = 318$) was not different from those with lower threshold ($< 50^\circ$, V1: 0.038 ± 0.028 , $N = 130$, V4: 0.041 ± 0.029 , $N = 187$; $p > 0.1$ for both comparisons, rank sum test). Further, there was no significant correlation between threshold and the strength of choice-related activity in either V1 ($r = 0.07$, $p = 0.33$ for M1R; $r = 0.004$, $p = 0.95$ for M2) or V4 ($r = -0.12$, $p = 0.17$ for M1L, $r = -0.07$, $p = 0.64$ for M1R, and $r = -0.06$, $p = 0.24$ for M2).

In the epoch after stimulus offset (200–400 ms), when choice signals were most evident, the mean absCP was slightly larger for neurons with low ($< 50^\circ$, V1: 0.045 ± 0.035 , $N = 255$, V4: 0.047 ± 0.036 , $N = 187$) than high threshold ($\geq 50^\circ$, V1: 0.035 ± 0.028 , $N = 255$, V4: 0.047 ± 0.036 , $N = 381$, $p(V1) = 0.03$, $p(V4) = 0.03$, rank sum test). There was a strong correlation between these two variables in V4 of one animal (Fig. 5B; M1L: $r = -0.32$, $p < 0.001$, permutation test) with the remaining cases showing no relationship ($p \geq 0.06$ for remaining cases).

We conclude that V4 contains stronger choice signals than V1, and that these signals are slightly more robust in neurons that have a lower threshold for task stimuli. Choice signals are most evident late in the trial, in the epoch between stimulus offset and the appearance of the targets used to report the decision.

Choice signal in neuronal populations

We next leveraged our simultaneous recordings to investigate the relationship between the animals' decisions and the responses of small neuronal populations in V1 and V4. Specifically, we attempted to predict the animals' choices on a trial-by-trial basis using a regularized, logistic regression model fit to the neuronal population responses (see Materials and Methods). We measured neuronal responses in the epoch 200–400 ms after stimulus onset, as single neuron choice signals were most evident at this time and population analysis did not reveal significant choice information in earlier epochs. For this analysis, we only considered sessions in which more than one neuron was recorded in each area. Performance was quantified by the percentage of correctly predicted choices on held-out trials.

To fit the logistic regression model, we first determined the ability of each neuron to predict choice. We then chose the unit

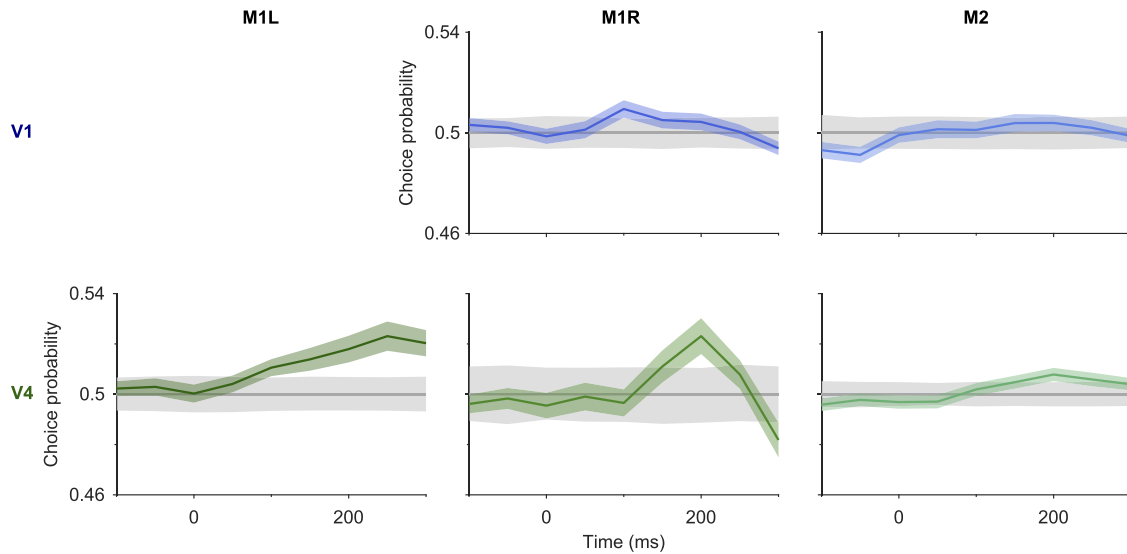


Figure 4. CP dynamics. Average CP, measured using a time window of 100 ms with an overlap of 50%. Top row, V1 neurons; bottom row, V4 neurons. Green and blue shaded areas denote SEM. Shaded gray areas denote the boundaries (95%) of the null distribution. Time of each bin is defined by the starting time of the time window. Zero milliseconds indicates stimulus onset.

with the highest performance and added units to the model in descending order of predictive performance, until model performance ceased improving (see Materials and Methods). Using this approach, we could predict the animal’s choice on $54.1 \pm 4.5\%$ of trials in V1 (Fig. 6A, blue; M1R: $53.7 \pm 4.5\%$, M2: $54.6 \pm 4.5\%$), exceeding the average 95th percentile confidence interval for chance performance (defined as model performance on permuted data) of 52.0%. For V1 populations, performance was significantly better than chance in 13 of 58 sessions (Fig. 6A; dark blue bars; permutation test; see Materials and Methods). In V4, model performance was more robust, predicting the decisions correctly on $58.6 \pm 6.5\%$ of trials (Fig. 6A, green; M1L: $60.7 \pm 6.7\%$; M1R: $53.5 \pm 4.1\%$, M2: $60.2 \pm 6.1\%$), with the 95% confidence interval for chance performance being 51.6%. Performance of models using V4 responses were significantly above chance in 40 of 67 sessions (Fig. 6A, dark green bars). Notably, we observed robust choice information in V4 of all three datasets, though choice signals were barely evident in the individual units recorded in one of the animals (M2; Figs. 3, 4). Differences between V1 and V4 choice signals were also evident when we excluded the V4 data from M1L (for which there were no corresponding V1 responses).

In additional analyses, we attempted to predict choices using models that considered the V1 and V4 populations together, but these models did not out-perform models based solely on V4 responses (data not shown). We also tested whether the performance of V1 and V4 populations was correlated on a session-by-session basis (e.g., whether

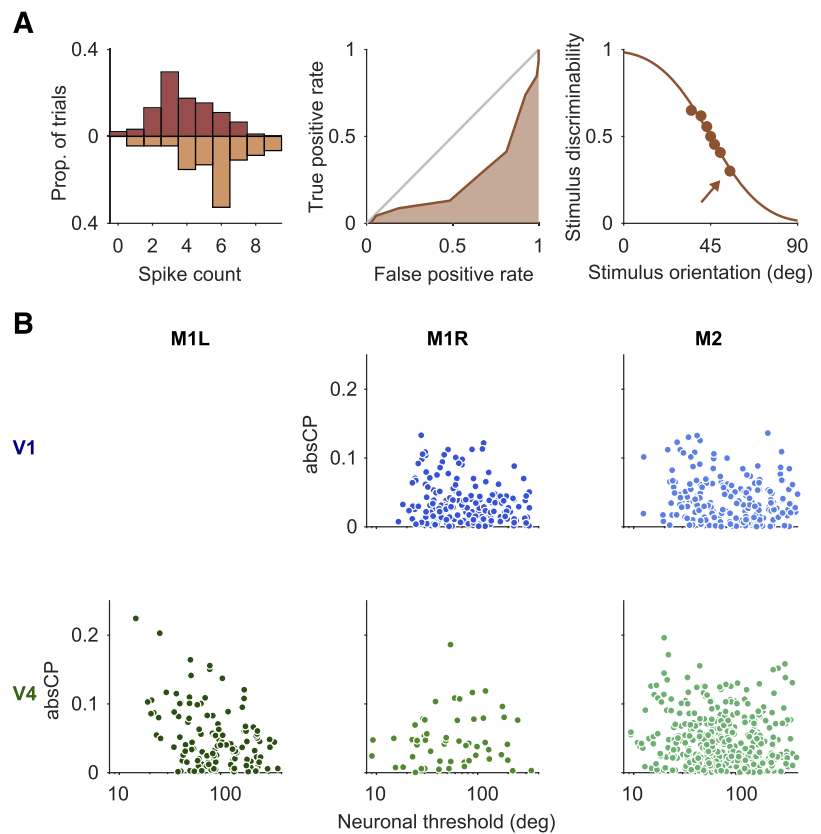


Figure 5. Relationship between neuronal threshold and choice probability. **A**, Definition of the neurometric function. Left, Spike count distributions for a 45° stimulus (dark brown) and a 55° stimulus (light brown). Middle, ROC curve for the two given spike count distributions in the left figure. Right, Stimulus discriminability for all orientations (dots) and fitted cumulative Gaussian (line). This neuron had a threshold of 20.9°. Arrow indicated the stimulus orientation that is discriminated from the 45° stimulus. **B**, absCP as a function of neuronal threshold for all V1 (top row, blue) and V4 (bottom row, green) neurons. Neuronal thresholds are based on responses measured 0–250 ms after stimulus onset; CP on responses 200–400 ms after stimulus onset.

sessions in which choice signals were particularly robust in V4 were also those sessions in which V1 contained stronger choice signals). We found a weak correlation in population performance in M2 ($r = 0.31$, $p = 0.05$) and no relationship in M1R ($r = -0.35$; $p > 0.1$).

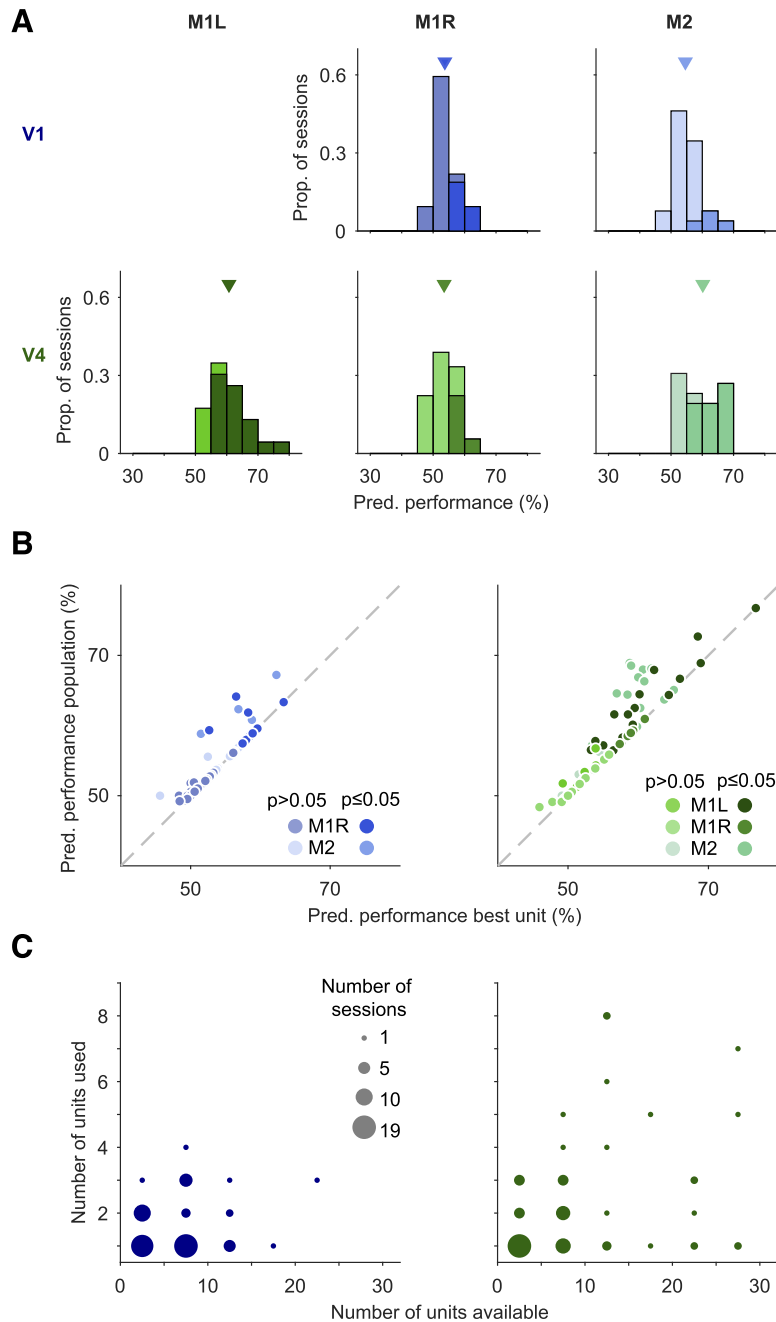


Figure 6. Choice signals in neuronal populations. **A**, Histogram of predictive (Pred) performance for each session in V1 (top row, blue) and V4 (bottom row, green). Dark bars indicate decoding performance, which is statistically significantly above chance performance. **B**, Performance of the neuronal population compared with that of the best unit within that population, for V1 (left, blue) and V4 (right, green). Dark dots indicate performance that is significantly above chance. **C**, Number of units included in the population analysis as a function of the available population size, for V1 (left, blue) and V4 (right, green). Size of the dots indicate the number of sessions. Decoding performance is based on responses measured 200–400 ms after stimulus onset.

We next compared the performance of single neurons and populations, on a session-by-session basis (Fig. 6B). By definition, the population performance was equal to, or better than, the performance of the best single neuron of each session. The improvement was generally modest, with an average increase in performance from 53.1 ± 3.9 to $54.1 \pm 4.5\%$ for V1 (M1R: 52.9 ± 3.9 to 53.7 ± 4.5 , M2: 53.4 ± 4.0 to 54.6 ± 4.5) and from 56.8 ± 5.6 to $58.6 \pm 6.5\%$ for V4 (M1L: 58.8 ± 6.5 to 60.7 ± 6.6 , M1R: 53.3 ± 4.4 to 53.5 ± 4.1 , M2: 57.5 ± 4.4 to 60.2 ± 6.1). The limited improvement was due in part to models using only a

small number of neurons (Fig. 6C; in V1, M1R: 1.6 ± 0.9 , M2: 1.5 ± 0.8 ; in V4, M1L: 2.4 ± 1.6 ; M1R: 1.2 ± 0.5 , M2: 2.5 ± 2.0). When larger populations were available, we observed more sizable improvements. For instance, when the population decoder considered three or more neurons, performance increased to $60.0 \pm 7.5\%$, relative to a performance of $55.8 \pm 6.0\%$ provided by the single best neuron.

We conclude that V4 population responses are predictive of perceptual decisions in a fine orientation discrimination task, whereas V1 population responses contain much weaker signals. Decisions can be predicted from the responses of small populations, even when choice signals in individual neurons are, on average, barely evident.

Influence of choice history

Animals' decisions are known to be influenced by their choice and reward history, especially if the task stimuli are near threshold (Seidemann, 1998; Gold et al., 2008; Busse et al., 2011; Fründ et al., 2014; Abrahamyan et al., 2016; Lueckmann et al., 2018). For instance, in the absence of strong sensory information, a subject might be more likely to repeat a decision if it led to a successful outcome on the previous trial, or to switch decisions if the previous choice was wrong. Although these choice-history effects are well known in the perceptual literature, they are seldom taken into account when attempting to understand the relationship between neuronal responses and perceptual reports (but see Dodd et al., 2001; Lueckmann et al., 2018).

We used a modified version of a regression analysis, previously applied to rodent (Busse et al., 2011) and human (Abrahamyan et al., 2016) psychophysical experiments, to evaluate the degree to which our animals were influenced by their previous decisions. Our model consisted of sensory terms (one covariate for each stimulus orientation) and history terms (one covariate for previously rewarded trials and one for previously unrewarded trials), as well as a bias term (see Materials and Methods). A sense of the relative importance of the previous trial's

choice is provided by comparing the history weights to the sensory weights, as shown in Figure 7A for an example session. The sensory weights for stimuli far away from the decision boundary are large (e.g., 0 and 90°), so the choice on trials involving these stimuli was driven almost entirely by the sensory evidence. However, for trials involving stimuli close to the decision boundary (e.g., 46°), the animals' decision was influenced to a comparable degree by the choice and reward of the previous trial as by the sensory stimulus itself (β_{reward} was -0.31 , $\beta_{\text{no reward}}$ was 0.29 and weight for the 46° stimulus was 0.68).

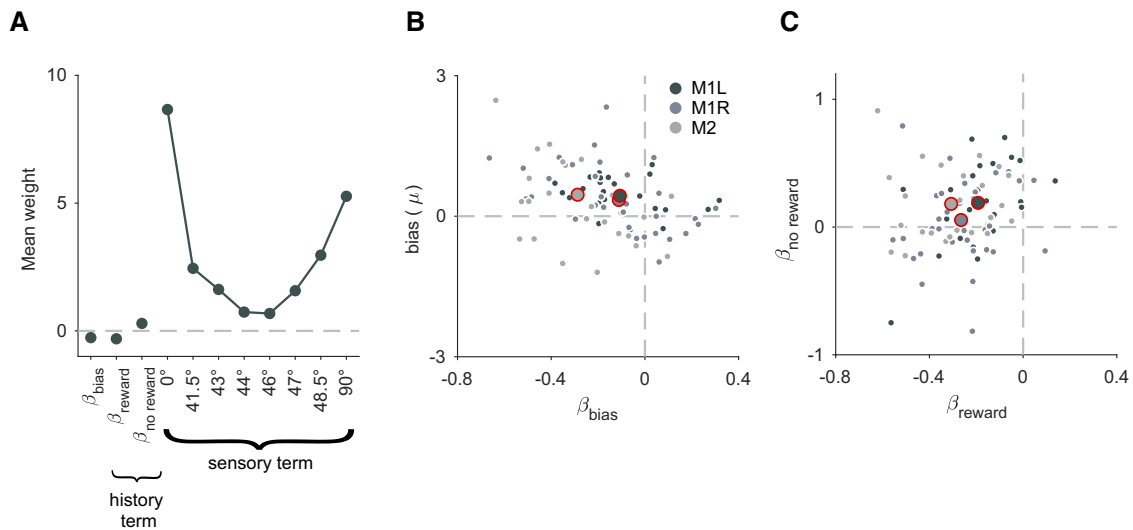


Figure 7. Influence of choice and reward history on the decision. **A**, Example weights for one session in M1L. **B**, Bias estimated from the psychometric function (ordinate) compared with the bias estimated from the choice-history model (abscissa). Small dots are individual sessions; large dots with red circle indicate the mean over sessions for each dataset. **C**, Weights for the history term variables of each session. Small dots represent individual sessions; large dots with red circle indicate the mean over all sessions for each dataset (indicated by shading).

Across sessions, the choice-history model predicted $77.5 \pm 3.1\%$ of choices (M1L: $79.4 \pm 3.0\%$, M1R: $77.0 \pm 3.3\%$, M2: $76.3 \pm 2.2\%$). The estimated bias weight, β_{bias} , agreed well with the bias measured from the psychometric function (Fig. 7B; M1L: $r = -0.29$, M1R: $r = -0.58$, M2: $r = -0.41$), further indicating that the model accurately predicted choices. On average, the weights for choice history were small but consistent across animals (Fig. 7C). Previous rewarded trials (β_{reward}) were weighted negatively -0.26 ± 0.16 ($p < 0.001$, one sample t test; M1L: -0.19 ± 0.16 , M1R: -0.26 ± 0.15 , M2: -0.31 ± 0.17), indicating that the animals had a weak tendency to switch choices after a rewarded trial. Weights for previous unrewarded trials ($\beta_{\text{no reward}}$) were positive 0.13 ± 0.31 ($p < 0.001$, one sample t test; M1L: 0.19 ± 0.16 , M1R: 0.05 ± 0.32 , M2: 0.18 ± 0.28), indicating a tendency to repeat a choice after a previously unrewarded trial. Thus, animals adopted a “win switch, lose stay” strategy.

The strategy adopted by our animals might reflect a small bias in our stimulus sequences, in which it was more common for the correct response to switch between successive trials (54%) than to remain the same. Lueckmann et al. (2018) observed a tendency of their animals to change reports across trials, which they suggested was because of a weak bias for switching in their stimulus sequence (52%). Alternatively, the strategy adopted by our animals might be inherent, as there is a broad range of strategies evident across individuals (Busse et al., 2011; Fründ et al., 2014; Abrahamyan et al., 2016). In any case, to be sure that the stimulus structure was not directly responsible for the choice-history weights we report, we conducted a simple simulation. We used sequences identical to those shown to our animals, but generated “choices” randomly, using the psychometric function. For instance, if the animal chose vertical on 75% of trials with a 48° grating, we generated a vertical choice with this probability for each 48° grating in our sequence. In models fit to these synthetic data, both choice-history weights were indistinguishable from zero (β_{reward} of 0.01 ± 0.14 , $p = 0.45$ for difference from zero; $\beta_{\text{no reward}}$ of -0.01 ± 0.26 , $p = 0.72$). Thus, the weights in the model fit to our data reflect the animal’s behavioral strategy rather than the weak structure of our stimulus sequences.

To assess the importance of the choice-history terms, we compared the model described above to a no-history model, consist-

ing only of the sensory terms. The performance of the two models were not significantly different on average (mean no-history model: $77.4 \pm 3.3\%$, $p > 0.2$, for all three datasets and for the pooled data, paired t test). In addition, the history model was significantly better than the no-history model (χ^2 test, $\alpha = 0.05$) in only $3.2 \pm 9.5\%$ of the cases (M1L: $3.2 \pm 11.8\%$, M1R: $3.8 \pm 9.6\%$, M2: $2.4 \pm 6.8\%$). For comparison, Abrahamyan et al. (2016) reported that the choice-history model significantly outperformed a model without history terms for roughly one-half of their human subjects. Together, these analyses suggest that there is little influence of choice history on our animals’ perceptual reports compared with the influence of the sensory stimulus.

Although the influence of choice history was small, we asked whether combining this information with the responses of the measured neuronal population model would lead to improved performance predicting animals’ choices for stimuli near perceptual threshold. For each recorded unit, we thus fit a logistic regression model that used both the measured responses and the choice-history covariates to predict decisions on a trial-by-trial basis. We only considered neuronal responses and perceptual reports from trials involving stimuli near the decision boundary (45° and the two nearest orientations), as in the preceding analyses. However, the choice-history covariates reflected the choice on trials before those involving these stimuli, regardless of whether those previous trials involved stimuli near the discrimination boundary.

Including choice-history information improved single neuron cross-validated performance from 50.4 ± 2.8 to $53.2 \pm 4.5\%$ for V1 units (Fig. 8A, blue; $p < 0.001$, paired t test), and 51.3 ± 3.9 to $54.1 \pm 5.1\%$ for V4 (green; $p < 0.001$). We applied a similar strategy to assess how combining choice-history information with neuronal population responses improved our ability to predict the animals’ decisions. For V1 populations, including choice-history information improved model performance from 54.1 ± 4.5 to $56.0 \pm 4.4\%$ (Fig. 8B, blue; $p = 0.004$, paired t test). However, for V4, adding the choice-history did not improve performance significantly (green; from 58.6 ± 6.5 to $58.4 \pm 5.8\%$, $p = 0.70$, paired t test).

We wondered whether performance for models using V4 populations failed to improve because the choice signals in

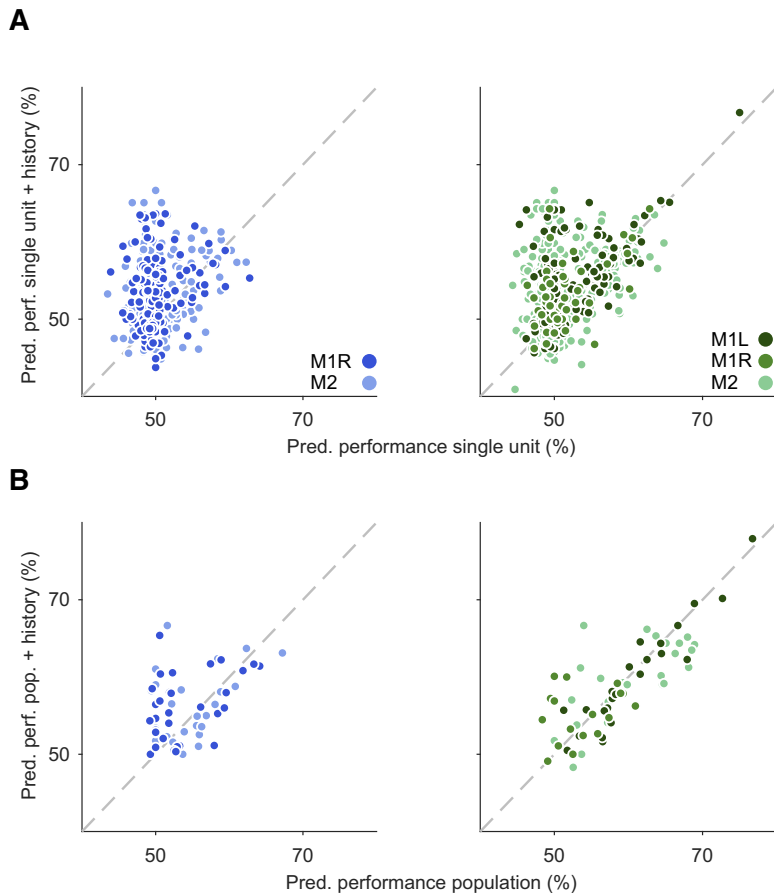


Figure 8. Predicting choices using neuronal responses and choice-history information. **A**, Predictive (Pred) performance (perf) of a model based on single neuron responses alone (abscissa), compared with the performance of a model that also includes history terms for V1 (left) and V4 (right). **B**, Performance of models using the neuronal population (pop) responses alone, compared with the performance of models that include choice-history terms for V1 (left) and V4 (right). Performance is based on responses measured 200–400 ms after stimulus onset.

those responses reflected choice history directly. In this case, adding choice-history information would not improve model performance, as that information could already be extracted from the neuronal responses themselves. To test this possibility, we divided each session’s data into two subsets, conditioned on whether the previous trial involved a vertical or horizontal choice. We then compared the performance afforded by the neuronal populations for these two subsets, to the performance of models fit to responses on a subset of data involving trials with a mix of both choices in previous trials. If the choice signals in neuronal responses reflect choice history, then conditioning on the choice made in the previous trial should reduce model performance. However, model performance was not significantly different when fit to choice-conditioned or mixed subsets (V1: $p = 0.07$, $N = 39$ sessions, V4: $p = 0.11$, $N = 40$ sessions, paired t test).

As an additional test of whether neuronal choice signals encoded choice history, we assessed whether V4 units with strong choice signals were observed in sessions with strong choice-history effects. We quantified the strength of choice-history effects as the performance of a model based on choice-history terms alone (in predicting choices), and then compared this performance to a model based on single neuron responses. The correlation coefficient was not significantly different from zero in V1 ($r = -0.03$, Pearson correlation, $p = 0.55$) or V4 ($r = -0.07$, $p =$

0.08), suggesting a dissociation of these two signals. Together, these analyses indicate that choice-related information in neuronal responses does not reflect an encoding of choice history.

In summary, the choice and reward outcome of the previous trial provide information about the animal’s choice on the current trial. Compared with the influence of sensory information, the influence of choice history is small. Nevertheless, this information improves the ability to predict choices from measured neuronal responses. Choice history appears to be an independent predictor of performance, distinct from the choice signals present in the neuronal responses.

Influence of eye movements during fixation

Although we only analyzed trials in which animals maintained fixation within a 1.4° window, small eye movements within that window could potentially affect the observed relationships between neuronal responses and choice (Herrington et al., 2009). We therefore repeated our core analyses, after excluding trials in which microsaccades were detected (see Materials and Methods). Briefly, we defined a microsaccade as an eye movement that remained within the fixation window, involved a velocity of 10°/s or more, and persisted for at least 8 ms. Figure 9A illustrates eye position (left) and velocity (center) on a trial which included a microsaccade (red). The events detected by our method showed a relationship between eye movement velocity and amplitude (Fig. 9A, right), typical for saccades and microsaccades (Zuber et al., 1965; Martinez-Conde et al., 2013). Microsaccades occurred at a rate of 1.7 events/s, consistent with previous studies (Bair and O’Keefe, 1998; Horwitz and Albright, 2003; Bosman et al., 2009; Hafed et al., 2009). Thus, ~61.9% of trials contained a microsaccade.

We found no significant difference in CP calculated from trials containing microsaccades with those calculated from trials that did not (Fig. 9B; $p > 0.32$ for V1 and V4 in all datasets, signed rank test). Similarly, the performance of models fit to neuronal population responses were similar in the two subsets of trials (Fig. 9C), as was the performance of population models which included choice-history terms (Fig. 9D; paired t test, $p > 0.13$ for all comparisons).

We conclude that the reported relationship between choice and neuronal responses in V1 and V4 cannot be attributed to microsaccades.

Discussion

We recorded simultaneously from neuronal populations in V1 and V4 while monkeys performed a fine orientation discrimination task. We found limited choice signal in either area using single neurons, but significant choice information in V4 neuronal populations. Notably, choice information was most evident after stimulus offset, in the brief epoch before the animal indi-

cated its choice. Our ability to predict choice on a trial-by-trial basis was improved by including knowledge about the choice made on the preceding trial, a signal that was distinct from the choice-related information evident in the neuronal responses.

Using small populations of V4 neurons, we were able to correctly predict choices on nearly 60% of trials. This compares favorably with previous measurements of V4 single-neuron CP on a fine disparity discrimination task (mean = 0.55; Shiozaki et al., 2012; Kosai et al., 2014), though differences in method preclude a rigorous comparison (i.e., we measured performance as prediction accuracy on held-out data, whereas CP provides a measure of performance using all data). Neuronal populations were more informative of choice for two reasons. First, our models were able to focus on the neurons that provided reliable choice information in each session, and ignore uninformative neurons. Indeed, the optimally decoded population size usually involved just a few neurons. Second, in some sessions, we were able to effectively combine the weak choice signals present in multiple, individual neurons to yield a more accurate prediction. Our data show that even small neuronal populations can yield qualitatively different conclusions about the presence of choice signals in a given area or animal. Still larger populations than we obtained might provide sufficiently accurate estimates of population noise covariance to attempt to infer read-out weights (Haefner et al., 2013; Pitkow et al., 2015).

In V1, individual neuron CP was negligible and population choice information was only slightly stronger. Some previous studies have also reported that choice signals in V1 are weak or absent (Nienborg and Cumming, 2006; Hass and Horwitz, 2013; Goris et al., 2017), though others have detected measurable choice signals there (Palmer et al., 2007; Nienborg and Cumming, 2014; Goris et al., 2017; Bondy et al., 2018). In addition, fluctuations in V1 population responses, as measured by voltage-sensitive dye optical imaging, can be strongly predictive of stimulus detection (Michelson et al., 2017; Seidemann and Geisler, 2018). Thus, larger populations might have revealed stronger V1 choice signals in our task.

Although our data cannot prove an absence of choice signals in V1 in our task, they do show that, for populations of similar size, V4 has more robust choice signals. This finding is consistent with the view that choice signals are stronger in higher visual cortex, a view that has emerged largely from comparisons made across studies (Nienborg et al., 2012). Several studies have compared choice signals in neurons recorded in different areas within the same animal, albeit in different behavioral sessions (Cook and Maunsell, 2002; Williams et al., 2003; Nienborg and Cumming, 2006; de Lafuente and Romo, 2006). Because choice signals may be strongly influenced by task strategy (Bondy et al., 2018) and fluctuations in attentional state (Cumming and Nienborg, 2016),

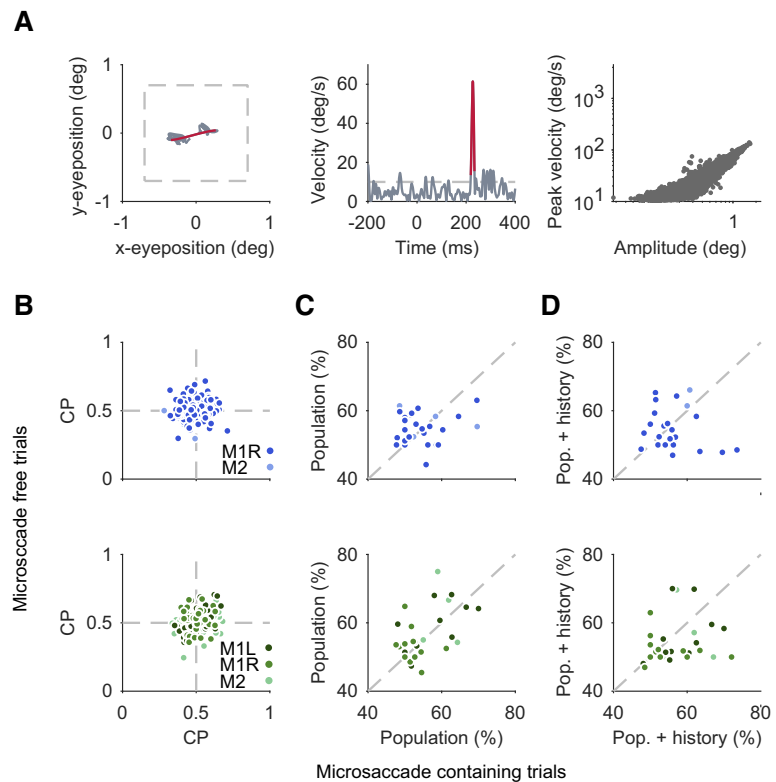


Figure 9. Testing for an influence of microsaccades. **A**, Detecting microsaccades. Left, Eye position during an example trial in M1R, from the moment fixation was established (−200 ms) to the appearance of the choice targets (400 ms). Red line indicates a detected microsaccade; gray square is the fixation window. Middle, Velocity profile for the example trial. Red indicates the detected microsaccade; gray dashed line indicates 10°/s criterion. Right, Peak velocity of detected microsaccades as a function of their amplitude. **B**, CP for trials containing microsaccades compared with CP for trials without microsaccades. **C**, Performance of the neuronal population (pop) on trials with microsaccades compared with trials without. **D**, Performance of models using neuronal population responses and trial history, using trials with microsaccades compared with the performance using trials without. Top (blue), V1; bottom (green), V4. All analyses based on responses measured 200–400 ms after stimulus onset.

both of which can differ across sessions, comparisons across areas should involve simultaneous sampling of their neurons, as we have done.

In our study, choice signals were only weakly related to neuronal sensitivity, a relationship more evident in many previous studies (Purushothaman and Bradley, 2005; for review, see Nienborg et al., 2012) though some others have also found this relationship to be either weak (in V1: Nienborg and Cumming, 2006; in MT: Price and Born, 2010; in V4: Shiozaki et al., 2012) or absent (Palmer et al., 2007; in V1 detection tasks: Hass and Horwitz, 2013). A relationship between neuronal sensitivity and strength of choice signals is expected, if fluctuations in the measured sensory representation drive the perceptual decision and if the sensory information is extracted by an optimal linear readout (Haefner et al., 2013; Pitkow et al., 2015). Thus, the absence of this relationship may indicate suboptimal readout by our animals, though their performance rivalled human subjects in similar fine orientation discrimination tasks (Mäkelä et al., 1993; Schoups et al., 1995; Goris et al., 2017); or that the mapping from sensory representations to decisions is not linear; or that the choice signals reflect rather than drive the decision (Nienborg and Cumming, 2009; Bondy et al., 2018).

The weaker choice signal in V1 than V4 is notable given two factors that would predict the opposite outcome. First, most V1 neurons are selective for stimulus orientation (Hubel and Wiesel, 1962; Ringach et al., 2002), the relevant sensory variable in our discrimination task. V4 encoding of orientation is slightly less

selective than in V1 (Desimone and Schein, 1987; McAdams and Maunsell, 1999; Yang and Maunsell, 2004), although we found little difference in selectivity between these areas based on responses to a narrow range of orientations $\sim 45^\circ$ in a blindly-selected sample of neurons. Because higher neuronal sensitivity is often associated with stronger choice-related signals, one would expect that choice signals should be more robust in V1 than V4, opposite to our findings. Second, V1 neurons with similar orientation preferences are spatially-clustered (Hubel and Wiesel, 1962; Bonhoeffer and Grinvald, 1991). In V4, the spatial organization is less ordered (Tanigawa et al., 2010). It has been suggested that strong spatial clustering of selectivity is required for generating strong choice signals (Nienborg and Cumming, 2014). This would also suggest that choice signals should be stronger in V1 than V4, opposite to our findings. Our data indicate instead that choice signals may be more strongly determined by the proximity of the sensory representation to “decision areas” (but see Liu et al., 2013), than by the sensitivity and clustering of neurons within an area. We note, however, that the difference in choice signal strength in V1 and V4 may depend on stimulus parameters (e.g., size, or spatial or temporal frequency), which we did not vary systematically.

The dynamics of choice signals in V4 are also more consistent with these signals reflecting, rather than driving, the decision (a “feedback” rather than “feedforward” origin; Shadlen et al., 1996; Nienborg and Cumming, 2009; Nienborg et al., 2012; Cumming and Nienborg, 2016). In our data, choice signals were strongest after the sensory-evoked response had ended, making it unlikely that the fluctuations in sensory responses drove the decisions. In addition, saccades were initiated shortly after the appearance of choice targets (M1L: 189 ± 41 ms; M1R: 82 ± 61 ms; M2: 113 ± 73 ms), and thus not long after the appearance of choice signals in V4 activity. This leaves little time for V4 population fluctuations to drive the decision and for the corresponding saccade to be planned. By this reasoning, the absence of choice-related signals in V1 might indicate a failure of decision-related information to propagate to the earliest stage of processing, either because this propagation requires additional time or because it is not required in our task (Haefner et al., 2016).

The dynamics of choice information in our data differ from those of most previous studies, which typically find choice information appears shortly after stimulus onset and remains evident throughout the trial (Britten et al., 1996; Nienborg and Cumming, 2009; Shiozaki et al., 2012; Wimmer et al., 2015). This difference may arise from our use of brief, suprathreshold stimuli, whereas most previous measurements of dynamics involved tasks in which animals integrated “noisy” stimuli, containing weak sensory evidence. Under these conditions, it can be difficult to determine the temporal relationship between the arrival of sensory information, the formation of the decision, and the evidence for choice signals in the sensory representation (Cohen and Newsome, 2009; Nienborg and Cumming, 2009). Our use of briefly presented stimuli constrain the time at which sensory information can be used to guide decisions (Price and Born, 2010; Smith et al., 2011).

The close temporal proximity of V4 choice signals to the initiation of the decision saccade might indicate that our choice signals are affected by pre-saccadic modulation of V4 responses (Tolias et al., 2001; Ibbotson and Kregelberg, 2011). However, several factors argue against this interpretation. First, pre-saccadic modulation is most evident within ~ 100 ms of saccade onset (Tolias et al., 2001; Han et al., 2009), but choice information was evident in our data 200–400 ms before saccade initia-

tion. Second, Tolias et al. (2001) reported that pre-saccadic modulation was only evident when neurons were visually driven. Our choice signals were strongest after the stimulus-driven response had ended. Finally, pre-saccadic modulation is greatest for saccades toward the neuronal RF and substantially weaker for saccades away from the RF (Supèr et al., 2004; Steinmetz and Moore, 2010). In our task, saccades were made to choice targets offset from the neuronal spatial RFs.

Our ability to predict choice from neuronal responses was improved by considering choice and reward history, at least for V1 responses for which performance was weak otherwise. However, this choice-history information does not appear to be encoded explicitly in V1 or V4 responses: our ability to predict choices from neuronal responses was unaffected by conditioning on the choice of the previous trial. Several recent studies, using distinct analyses, have reached similar conclusions, finding that choice history has a measurable effect on behavior but that this information is not represented in V1 (Lueckmann et al., 2018), or in MT or lateral intraparietal (LIP) cortex (Gold et al., 2008). The absence of a neural signature of choice-history information in early and midlevel sensory areas suggests the internal representation (memory) of choice and reward history is not relayed to those areas. This is in contrast to the information about the decision on the current trial, which our data and those of others (Nienborg and Cumming, 2009; Bondy et al., 2018) suggest contributes strongly to the choice signals in sensory cortex. It thus appears that not all decision-related information is incorporated in the sensory representation (Haefner et al., 2016), but instead a selective component of it, namely the decision at the current time.

References

- Abrahamyan A, Silva LL, Dakin SC, Carandini M, Gardner JL (2016) Adaptable history biases in human perceptual decisions. *Proc Natl Acad Sci U S A* 113:E3548–E3557.
- Akrami A, Kopec CD, Diamond ME, Brody CD (2018) Posterior parietal cortex represents sensory history and mediates its effects on behaviour. *Nature* 554:368–372.
- Astrand E, Enel P, Ibos G, Dominey PF, Baraduc P, Ben Hamed S (2014) Comparison of classifiers for decoding sensory and cognitive information from prefrontal neuronal populations. *PLoS One* 9:e86314.
- Averbeck BB, Latham PE, Pouget A (2006) Neural correlations, population coding and computation. *Nat Rev Neurosci* 7:358–366.
- Bair W, O’Keefe LP (1998) The influence of fixational eye movements on the response of neurons in area MT of the macaque. *Vis Neurosci* 15:779–786.
- Bair W, Zohary E, Newsome WT (2001) Correlated firing in macaque visual area MT: time scales and relationship to behavior. *J Neurosci* 21:1676–1697.
- Bondy AG, Haefner RM, Cumming BG (2018) Feedback determines the structure of correlated variability in primary visual cortex. *Nat Neurosci* 21:598–606.
- Bonhoeffer T, Grinvald A (1991) Iso-orientation domains in cat visual cortex are arranged in pinwheel-like patterns. *Nature* 353:429–431.
- Bosman CA, Womelsdorf T, Desimone R, Fries P (2009) A microsaccadic rhythm modulates gamma-band synchronization and behavior. *J Neurosci* 29:9471–9480.
- Britten KH, Shadlen MN, Newsome WT, Movshon JA (1992) The analysis of visual motion: a comparison of neuronal and psychophysical performance. *J Neurosci* 12:4745–4765.
- Britten KH, Newsome WT, Shadlen MN, Celebrini S, Movshon JA (1996) A relationship between behavioral choice and the visual responses of neurons in macaque MT. *Vis Neurosci* 13:87–100.
- Busse L, Ayaz A, Dhruv NT, Katzner S, Saleem AB, Schölvinck ML, Zaharia AD, Carandini M (2011) The detection of visual contrast in the behaving mouse. *J Neurosci* 31:11351–11361.
- Chang CC, Lin CJ (2011) LIBSVM: a library for support vector machines. *ACM Trans Intell Syst Technol* 2:27:1–27:27.

- Chen A, Gu Y, Liu S, DeAngelis GC, Angelaki DE (2016) Evidence for a causal contribution of macaque vestibular, but not intraparietal, cortex to heading perception. *J Neurosci* 36:3789–3798.
- Cohen MR, Kohn A (2011) Measuring and interpreting neuronal correlations. *Nat Neurosci* 14:811–819.
- Cohen MR, Newsome WT (2009) Estimates of the contribution of single neurons to perception depend on timescale and noise correlation. *J Neurosci* 29:6635–6648.
- Cook EP, Maunsell JH (2002) Dynamics of neuronal responses in macaque MT and VIP during motion detection. *Nat Neurosci* 5:985–994.
- Crapse TB, Basso MA (2015) Insights into decision making using choice probability. *J Neurophysiol* 114:3039–3049.
- Cumming BG, Nienborg H (2016) Feedforward and feedback sources of choice probability in neural population responses. *Curr Opin Neurobiol* 37:126–132.
- de Lafuente V, Romo R (2006) Neural correlate of subjective sensory experience gradually builds up across cortical areas. *Proc Natl Acad Sci U S A* 103:14266–14271.
- Desimone R, Schein SJ (1987) Visual properties of neurons in area V4 of the macaque: sensitivity to stimulus form. *J Neurophysiol* 57:835–868.
- Dodd JV, Krug K, Cumming BG, Parker AJ (2001) Perceptually bistable three-dimensional figures evoke high choice probabilities in cortical area MT. *J Neurosci* 21:4809–4821.
- Fründ I, Wichmann FA, Macke JH (2014) Quantifying the effect of intertrial dependence on perceptual decisions. *J Vis* 14(7):9 1–16.
- Gattass R, Sousa AP, Gross CG (1988) Visuotopic organization and extent of V3 and V4 of the macaque. *J Neurosci* 8:1831–1845.
- Gold JL, Law CT, Connolly P, Benuer S (2008) The relative influences of priors and sensory evidence on an oculomotor decision variable during perceptual learning. *J Neurophysiol* 100:2653–2668.
- Goris RLT, Ziemba CM, Stine GM, Simoncelli EP, Movshon JA (2017) Dissociation of choice formation and choice-correlated activity in macaque visual cortex. *J Neurosci* 37:5195–5203.
- Haefner RM, Gerwinn S, Macke JH, Bethge M (2013) Inferring decoding strategies from choice probabilities in the presence of correlated variability. *Nat Neurosci* 16:235–242.
- Haefner RM, Berkes P, Fiser J (2016) Perceptual decision-making as probabilistic inference by neural sampling. *Neuron* 90:649–660.
- Hafed ZM, Goffart L, Krauzlis RJ (2009) A neural mechanism for microsaccade generation in the primate superior colliculus. *Science* 323:940–943.
- Han X, Xian SX, Moore T (2009) Dynamic sensitivity of area V4 neurons during saccade preparation. *Proc Natl Acad Sci U S A* 106:13046–13051.
- Hass CA, Horwitz GD (2013) V1 mechanisms underlying chromatic contrast detection. *J Neurophysiol* 109:2483–2494.
- Herrington TM, Masse NY, Hachmeh KJ, Smith JE, Assad JA, Cook EP (2009) The effect of microsaccades on the correlation between neural activity and behavior in middle temporal, ventral intraparietal, and lateral intraparietal areas. *J Neurosci* 29:5793–5805.
- Horwitz GD, Albright TD (2003) Short-latency fixational saccades induced by luminance increments. *J Neurophysiol* 90:1333–1339.
- Hubel DH, Wiesel TN (1962) Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J Physiol* 160:106–154.
- Ibbotson M, Krekelberg B (2011) Visual perception and saccadic eye movements. *Curr Opin Neurobiol* 21:553–558.
- Katz LN, Yates JL, Pillow JW, Huk AC (2016) Dissociated functional significance of decision-related activity in the primate dorsal stream. *Nature* 535:285–288.
- Kelly RC, Smith MA, Samonds JM, Kohn A, Bonds AB, Movshon JA, Lee TS (2007) Comparison of recordings from microelectrode arrays and single electrodes in the visual cortex. *J Neurosci* 27:261–264.
- Kimmel DL, Mammo D, Newsome WT (2012) Tracking the eye non-invasively: simultaneous comparison of the scleral search coil and optical tracking techniques in the macaque monkey. *Front Behav Neurosci* 6:49.
- Kohn A, Smith MA (2005) Stimulus dependence of neuronal correlation in primary visual cortex of the macaque. *J Neurosci* 25:3661–3673.
- Kohn A, Coen-Cagli R, Kanitscheider I, Pouget A (2016) Correlations and neuronal population information. *Annu Rev Neurosci* 39:237–256.
- Kosai Y, El-Shamayleh Y, Fyall AM, Pasupathy A (2014) The role of visual area V4 in the discrimination of partially occluded shapes. *J Neurosci* 34:8570–8584.
- Lakshminarasimhan KJ, Pouget A, DeAngelis GC, Angelaki DE, Pitkow X (2018) Inferring decoding strategies for multiple correlated neural populations. *PLoS Comput Biol* 14:e1006371.
- Liu S, Gu Y, DeAngelis GC, Angelaki DE (2013) Choice-related activity and correlated noise in subcortical vestibular neurons. *Nat Neurosci* 16:89–97.
- Lueckmann JM, Macke JH, Nienborg H (2018) Can serial dependencies in choices and neural activity explain choice probabilities? *J Neurosci* 38:3495–3506.
- Mäkelä P, Whitaker D, Rovamo J (1993) Modelling of orientation discrimination across the visual field. *Vision Res* 33:723–730.
- Martinez-Conde S, Otero-Millan J, Macknik SL (2013) The impact of microsaccades on vision: towards a unified theory of saccadic function. *Nat Rev Neurosci* 14:83–96.
- McAdams CJ, Maunsell JH (1999) Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4. *J Neurosci* 19:431–441.
- Michelson C, Pillow J, Seidemann E (2017) Majority of choice-related variability in perceptual decisions is present in early sensory cortex. *bioRxiv* 207357.
- Nienborg H, Cumming BG (2006) Macaque V2 neurons, but not V1 neurons, show choice-related activity. *J Neurosci* 26:9567–9578.
- Nienborg H, Cumming BG (2009) Decision-related activity in sensory neurons reflects more than a neuron's causal effect. *Nature* 459:89–92.
- Nienborg H, Cumming BG (2014) Decision-related activity in sensory neurons may depend on the columnar architecture of cerebral cortex. *J Neurosci* 34:3579–3585.
- Nienborg H, Cohen MR, Cumming BG (2012) Decision-related activity in sensory neurons: correlations among neurons and with behavior. *Annu Rev Neurosci* 35:463–483.
- Palmer C, Cheng SY, Seidemann E (2007) Linking neuronal and behavioral performance in a reaction-time visual detection task. *J Neurosci* 27:8122–8137.
- Parker AJ, Newsome WT (1998) Sense and the single neuron: probing the physiology of perception. *Annu Rev Neurosci* 21:227–277.
- Pitkow X, Liu S, Angelaki DE, DeAngelis GC, Pouget A (2015) How can single sensory neurons predict behavior? *Neuron* 87:411–423.
- Price NS, Born RT (2010) Timescales of sensory- and decision-related activity in the middle temporal and medial superior temporal areas. *J Neurosci* 30:14036–14045.
- Purushothaman G, Bradley DC (2005) Neural population code for fine perceptual decisions in area MT. *Nat Neurosci* 8:99–106.
- Qian J, Hastie T, Friedman J, Tibshirani R, Simon N (2013) Glmnet for MATLAB. Available at: http://www.stanford.edu/~hastie/glmnet_matlab/.
- Ringach DL, Shapley RM, Hawken MJ (2002) Orientation selectivity in macaque V1: diversity and laminar dependence. *J Neurosci* 22:5639–5651.
- Schoups AA, Vogels R, Orban GA (1995) Human perceptual learning in identifying the oblique orientation: retinotopy, orientation specificity and monocularity. *J Physiol* 483:797–810.
- Seidemann E (1998) Neuronal mechanisms mediating conversion of visual signals into perceptual decisions in a direction discrimination task. PhD dissertation, Stanford University.
- Seidemann E, Geisler WS (2018) Linking V1 activity to behavior. *Annu Rev Vis Sci* 4:287–310.
- Shadlen MN, Britten KH, Newsome WT, Movshon JA (1996) A computational analysis of the relationship between neuronal and behavioral responses to visual motion. *J Neurosci* 16:1486–1510.
- Shiozaki HM, Tanabe S, Doi T, Fujita I (2012) Neural activity in cortical area V4 underlies fine disparity discrimination. *J Neurosci* 32:3830–3841.
- Smith JE, Zhan CA, Cook EP (2011) The functional link between area MT neural fluctuations and detection of a brief motion stimulus. *J Neurosci* 31:13458–13468.
- Smolyanskaya A, Haefner RM, Lomber SG, Born RT (2015) A modality-specific feedforward component of choice-related activity in MT. *Neuron* 87:208–219.
- Steinmetz NA, Moore T (2010) Changes in the response rate and response variability of area V4 neurons during the preparation of saccadic eye movements. *J Neurophysiol* 103:1171–1178.
- Supèr H, van der Togt C, Spekrijse H, Lamme VA (2004) Correspondence of presaccadic activity in the monkey primary visual cortex with saccadic eye movements. *Proc Natl Acad Sci U S A* 101:3230–3235.

- Tanigawa H, Lu HD, Roe AW (2010) Functional organization for color and orientation in macaque V4. *Nat Neurosci* 13:1542–1548.
- Tolias AS, Moore T, Smirnakis SM, Tehovnik EJ, Siapas AG, Schiller PH (2001) Eye movements modulate visual receptive fields of V4 neurons. *Neuron* 29:757–767.
- Van Essen DC, Van Newsome WT, Maunsell JH (1984) The visual field representation in striate cortex of the macaque monkey: asymmetries, anisotropies, and individual variability. *Vision Res* 24:429–448.
- Williams ZM, Elfar JC, Eskandar EN, Toth LJ, Assad JA (2003) Parietal activity and the perceived direction of ambiguous apparent motion. *Nat Neurosci* 6:616–623.
- Wimmer K, Compte A, Roxin A, Peixoto D, Renart A, de la Rocha J (2015) Sensory integration dynamics in a hierarchical network explains choice probabilities in cortical area MT. *Nat Commun* 6:6177.
- Wissig SC, Kohn A (2012) The influence of surround suppression on adaptation effects in primary visual cortex. *J Neurophysiol* 107:3370–3384.
- Yang T, Maunsell JH (2004) The effect of perceptual learning on neuronal responses in monkey visual area V4. *J Neurosci* 24:1617–1626.
- Yu X, Gu Y (2018) Probing sensory readout via combined choice-correlation measures and microstimulation perturbation. *Neuron* 100:715–727.e5.
- Zohary E, Shadlen MN, Newsome WT (1994) Correlated neuronal discharge rate and its implications for psychophysical performance. *Nature* 370:140–143.
- Zuber BL, Stark L, Cook G (1965) Microsaccades and the velocity-amplitude relationship for saccadic eye movements. *Science* 150:1459–1460.