The Correlation of Neuronal Signals with Behavior at Different Levels of Visual Cortex and Their Relative Reliability for Behavioral Decisions

Incheol Kang^{1,3} and John H.R. Maunsell^{2,3}

¹Laboratory for Sensorimotor Research, National Eye Institute, National Institutes of Health, Bethesda, Maryland 20892, ²Department of Neurobiology, University of Chicago, Chicago, Illinois 60637, and ³Department of Neurobiology, Harvard Medical School, Howard Hughes Medical Institute, Boston, Massachusetts 02115

Behavior can be guided by neuronal activity in visual, auditory, or somatosensory cerebral cortex, depending on task requirements. In contrast to this flexible access of cortical signals, several observations suggest that behaviors depend more on neurons in later areas of visual cortex than those in earlier areas, although neurons in earlier areas would provide more reliable signals for many tasks. We recorded from neurons in different levels of visual cortex of 2 male rhesus monkeys while the animals did a visual discrimination task and examined trial-to-trial correlations between neuronal and behavioral responses. These correlations became stronger in primary visual cortex as neuronal signals in that area became more reliable relative to the other areas. The results suggest that the mechanisms that read signals from cortex might access any cortical area depending on the relative value of those signals for the task at hand.

Key words: cortical hierarchy; cortical readout; neuron-behavior correlation; perceptual decision; vision

Significance Statement

Information is encoded by the action potentials of neurons in various cortical areas in a hierarchical manner such that increasingly complex stimulus features are encoded in successive stages. The brain must extract information from the response of appropriate neurons to drive optimal behavior. A widely held view of this decoding process is that the brain relies on the output of later cortical areas to make decisions, although neurons in earlier areas can provide more reliable signals. We examined correlations between perceptual decisions and the responses of neurons in different levels of monkey visual cortex. The results suggest that the brain may access signals in any cortical area depending on the relative value of those signals for the task at hand.

Introduction

The idea that the activity of neurons in a given cortical area is weighted differently in different tasks is well accepted. Clinical and experimental studies have provided ample evidence indicating that somatosensory, auditory, or visual discriminations depend primarily on the activity of neurons in the corresponding sensory cortices (e.g., Penfield, 1947). At a finer spatial scale, lesions reveal that visual identification depends on neurons in

areas in the ventral processing pathway, while visual motion analysis and visually guided movements depend on neurons in the dorsal pathway (Ungerleider and Mishkin, 1982; Merigan and Maunsell, 1993; Goodale et al., 1994).

Observations such as these suggest that neurons in any region of cortex might contribute to behaviors in an equally direct way when their signals are well suited to the task at hand (see Histed et al., 2013). However, unlike studies across different sensory modalities, it is less clear whether the mechanisms supporting behaviors can selectively draw signals from areas at earlier and later stages of a cortical pathway for one sensory modality. Sensory signals are processed in a series of stages in cerebral cortex, with simple features being extracted at early stages and increasingly complex response properties elaborated in successive stages (Felleman and Van Essen, 1991; Hong et al., 2016). Because tradeoffs must be made in going from neurons that reliably signal low-level attributes, such as local contrast, to neurons that reliably signal object identity (Zoccolan et al., 2007), some sensory tasks might be better served by signals in earlier rather

Received Oct. 29, 2019; revised Feb. 22, 2020; accepted Mar. 12, 2020.

Author contributions: I.K. and J.H.R.M. designed research; I.K. and J.H.R.M. developed software; I.K. performed research; I.K. analyzed data; I.K. wrote the first draft of the paper; I.K. and J.H.R.M. edited and reviewed the paper; J.H.R.M. supervised research.

The authors declare no competing financial interests.

This work was supported by National Institutes of Health Grant R01EY005911. We thank Marlene R. Cohen, Jackson J. Cone, and Bram E. Verhoef for comments and discussion; and Vivian Imamura and Steven Sleboda for technical assistance.

Correspondence should be addressed to Incheol Kang at incheolkang@gmail.com.

<https://doi.org/10.1523/JNEUROSCI.2587-19.2020>

Copyright © 2020 the authors

than later stages of cortex. Therefore, low-level sensory discriminations might be better served by neurons in early sensory cortex, which likely provide more reliable information for those discriminations than neurons in later sensory cortex.

However, some observations suggest that readout is closely coupled to higher cortical areas. The contributions of individual neurons to particular behaviors are frequently studied using the correlation between trial-to-trial fluctuations of neuronal representations of a stimulus and the subject's perceptual choices based on the same stimulus, which is often quantified by choice probability (CP) (Parker and Newsome, 1998; Nienborg et al., 2012). Studies that compared this correlation across different cortical areas in monkeys have reported that the neuronal responses in higher cortical areas are more tightly coupled with the perceptual reports (Leopold and Logothetis, 1996; Cook and Maunsell, 2002; Williams et al., 2003; de Lafuente and Romo, 2006; Nienborg and Cumming, 2006; Carnevale et al., 2013). These studies suggest that the brain relies more on neuronal signals in higher cortical areas to guide behavior, consistent with the hierarchical processing of sensory information.

Were this the case, however, then neurons in early cortical areas would contribute less to behavior even when they provide more reliable signals. This is unexpected because, within individual cortical areas, a modest correlation has been described between CP of individual neurons and the neuronal performance on the task with which CP was measured (Celebrini and Newsome, 1994; Britten et al., 1996; Uka and DeAngelis, 2004; Gu et al., 2007; Price and Born, 2010; Bosking and Maunsell, 2011), suggesting that neurons carrying more reliable signals are given more weight when the responses are read out to make perceptual decisions (but see Haefner et al., 2013). Moreover, there is abundant evidence that, for some perceptual tasks using elementary stimulus features, behavioral performance is well matched to neuronal sensitivity in early cortical areas (Parker and Hawken, 1985; Bradley et al., 1987; Vogels and Orban, 1990; Geisler and Albrecht, 1997), or even that of thalamic (Kang and Malpeli, 2009) or peripheral (Mountcastle et al., 1972) sensory neurons.

Here, we examine this question by recording from individual neurons in the primary visual cortex (V1), the middle temporal area (MT), and the ventral intraparietal area (VIP) in rhesus monkeys while they performed a direction-change detection task that used stimuli optimized for V1 neurons. Consistent with earlier reports, trial-to-trial correlations between the neuronal response and the monkey's detection performance grew stronger from V1 to MT to VIP, although individual V1 neurons carried the most reliable signals relevant to the behavioral task. However, when the same task was repeated using stimuli that favored V1 neurons to a greater extent, trial-to-trial correlation measures between behavior and the responses of V1 neurons increased relative to those for MT and VIP neurons.

Materials and Methods

Animal preparation and behavioral task

Two adult male rhesus monkeys (*Macaca mulatta*; Monkeys 1 and 2), both weighing 11 kg, were used as subjects. All experimental procedures were approved by the Institutional Animal Care and Use Committee of Harvard Medical School. Before training, the monkeys were surgically implanted with a headpost and a scleral search coil (Robinson, 1963; Judge et al., 1980) for monitoring eye position.

After recovery from the surgery, the monkeys were trained extensively (5–6 months) to perform a direction-change detection task (see Fig. 1A). At the onset of a trial, a small white spot (0.1° diameter) was

presented on a CRT monitor as a fixation point. After the monkey fixated the fixation point for a variable period (375–625 ms), a small Gabor patch flashed multiple times (240 ms per flash) inside the receptive field of the neuron being recorded, with successive presentations separated by a variable period (200–307 ms). On each flash, the Gabor drifted in the same direction (reference stimulus), but on a pseudo-randomly chosen flash, the drifting direction differed (target stimulus). The monkey had to make a saccade to the target stimulus location between 150 and 600 ms after the onset of the target stimulus to get a juice reward. The target stimulus could be any flash, except for the first two flashes, with the probability of the target occurring following a decaying exponential distribution (mean = 1250 ms) that was truncated at 4410 ms. The reason that the earliest target occurred at the third flash was because the response of neurons to the first flash tended to be stronger than those to the subsequent flashes, which could influence the estimation of neuro-metric thresholds. To ensure a flat hazard function of the target occurrence before the limit, occasional catch trials (~7%) were randomly interleaved, on which no target stimulus appeared, and the monkey was rewarded for maintaining fixation for 600 ms after the onset of the final stimulus flash.

Eye positions were sampled at either 200 or 500 Hz, and the monkey's fixation behavior was monitored using a square electronic window (side length between 1.5° and 2°) centered on the fixation point. If the monkey broke fixation prematurely or failed to respond to the target stimulus, the trial ended without reward.

Visual stimuli

Visual stimuli were presented on a γ -corrected CRT monitor (1024 × 768 pixel resolution, 75 Hz vertical refresh rate) located at 48 cm from the monkey's eyes and subtending 43° horizontally and 33° vertically. When recording from neurons having a central receptive field (eccentricity $\leq 3.3^\circ$; mostly V1 neurons), the monitor was positioned at 61 cm from the monkey (35° × 27° visual field) to avoid possible loss of contrast for high spatial frequencies due to the limited pixel resolution (see below for stimulus scaling).

Visual stimuli were achromatic and presented on a uniform gray background with a luminance of 26 cd/m². The Gabor patches used in the direction-change detection task (see Fig. 1A) were of the maximum contrast with the mean luminance same as the background. To explore whether trial-to-trial correlations between behavior and neuronal responses in the three cortical areas interact with manipulations of stimulus parameters, we tested cells with two stimulus settings that differed in their eccentricity scaling factors. At any given retinal eccentricity, the SD of the Gabor function in the smaller stimulus set was two-thirds that of the larger stimulus set and the spatial frequency was 3 times higher (see Fig. 1B). We will refer to the set of larger stimuli as "Set L" and the set of smaller stimuli as "Set S." The SD and spatial frequency of Gabor functions in the two stimulus sets were scaled according to the receptive field eccentricity using the following formulae:

Stimulus Set L: SD (degree) = $0.075 \times \text{eccentricity (degree)}$

Spatial frequency (cycles/degree) = 0.75 cycles/SD .

Stimulus Set S: SD = $0.050 \times \text{eccentricity}$

Spatial frequency = 1.50 cycles/SD .

For example, at an eccentricity of 5°, the spatial frequency and SD of Gabor functions were 2 cycles/degree and 0.375° for stimulus Set L, and 6 cycles/degree and 0.25° for stimulus Set S. We used a fixed temporal frequency of 4.17 cycles/s, such that the Gabor function drifted one full spatial cycle during each stimulus flash on a trial, starting and ending with an odd-symmetric phase.

Electrophysiological recordings

After training was complete for the behavioral tasks described above, a second surgery was done to implant a stainless-steel chamber (19 mm outer diameter, Crist Instruments) on the skull to allow a posterior electrode approach to the recording areas at an angle 30° from horizontal in a parasagittal plane. Placement of the recording chamber, centered 12–14 mm lateral from the midline, was guided by structural MRIs taken before the initial surgery.

Figure 1. Behavioral task and visual stimuli. **A**, The direction-change detection task. After the monkey fixated a small spot (FP) for a randomly selected period (375–625 ms), a small drifting Gabor image flashed (240 ms) multiple times at the center of receptive field of the neuron, separated by blank screens (200–307 ms). The Gabor image drifted in the same direction (arrows), but on one randomly chosen flash the direction changed and the monkey had to make a saccade to the stimulus location within 600 ms. **B**, Stimulus scaling. Example Gabor images centered at 3, 5, and 8 degrees are illustrated for the two stimulus sets (Set L in the upper field, Set S in the lower field). Cross represents the center of the gaze. **C**, Selection of stimulus directions used in the behavioral task. Each dot represents the average response of a representative MT cell to Gabor patches from Set L drifting in the direction indicated on the horizontal axis. Responses were quantified as the mean firing rate during a 150 ms interval starting at 30 ms from the stimulus onset. Gray curve indicates a least-square fit of a wrapped Gaussian to the mean responses. Solid vertical line indicates the reference direction of the stimulus used for this cell in the main task. Dashed line indicates the neuron's preferred direction estimated by fit. Error bars indicate $\pm$ SEM. **D**, Neurometric threshold. The neurometric performance detecting target stimulus (filled circles) of the same MT neuron whose direction tuning is shown in **C** is plotted as a function of direction change of target stimulus from that of the reference stimulus. The neurometric performance for each direction change was defined as the area under ROC curve derived from the neuronal responses to target stimulus and its immediately preceding reference stimulus. A cumulative Weibull function (gray curve) with the lower asymptote set at 0.5 was fit to determine a neurometric threshold (arrow). **E**, Psychometric threshold. The hit rate (filled circles) of the monkey's responses at different target directions during the recording session for the same MT neuron in **C** and **D** is plotted as a function of direction change of the target stimulus. A cumulative Weibull function (gray curve) was fitted to determine a psychometric threshold (arrow), but with the guess rate estimated from false alarm responses set as the lower asymptote (dashed line; see Materials and Methods). **D**, **E**, Error bars indicate 95% CIs estimated from a bootstrap simulation (2000 iterations).

We made extracellular recordings from 222 cells in V1 (Monkey 1, 140; Monkey 2, 82), 160 cells in MT (Monkey 1, 79; Monkey 2, 81), and 136 cells in VIP (Monkey 1, 72; Monkey 2, 64). Recordings in V1 were made mostly from the operculum, but 21 V1 cells were recorded from the calcarine sulcus of Monkey 1. The receptive field eccentricities of V1 neurons ranged from 2.5° to 6.1° (4.2° mean, 1.0° SD) for sites on the operculum from 2 monkeys, and from 17.6° to 20° (18.8° mean, 0.7° SD) for sites in the calcarine sulcus from Monkey 1. The ranges of MT and VIP receptive field eccentricities were 2.5°–25° (11.2° mean, 5.2° SD) and 3°–35° (16.8° mean, 7.2° SD), respectively. Although receptive field eccentricities differed between the cortical areas, they did not differ between the two stimulus sets within cortical areas ($p > 0.05$, t test). The geometric means of the preferred spatial frequency (typically measured using at a temporal frequency of 4.17 cycles/s) were 1.94, 0.47, and 0.18 cycles/degree for V1, MT, and VIP, respectively, and those of the preferred temporal frequency (tested at the preferred spatial frequencies) were 4.1, 6.7, and 8.2 cycles/s, respectively. It should be noted that many MT and VIP neurons showed a low-pass tuning for the range of tested spatial frequencies (typically 0.125–8 cycles/degree), and a high-pass tuning for the range of the tested temporal frequencies (0.5–16 cycles/s). Therefore, it is possible that we overestimated the preferred spatial frequency and underestimated the preferred temporal frequency for these neurons.

Recordings were made with custom-made platinum-iridium electrodes (Wolbarsht et al., 1960) (0.5–2.5 M Ω at 1 kHz). Signals from the electrode were amplified (Bak or Plexon Electronics), filtered (bandpass from 0.5–6 kHz) and processed with a time-amplitude window discriminator (Bak Electronics) for isolation of action potentials of a single unit.

The time of each action potential was stored with a precision of 1 ms. The opercular regions of V1 were accessed through small craniotomies (2 mm diameter) made inside the recording chamber. A guide tube and grid system (Crist et al., 1988) was used for recordings from MT, VIP, and V1 in the calcarine sulcus. The recording chamber placement was aimed primarily for recordings from V1 and MT and was not ideal for accessing VIP. Although we accessed all three areas through one recording chamber, we used a grid with holes aligned 10° dorsomedial from the cylinder axis for VIP recordings for Monkey 1. MT and VIP recording sites were confirmed by the emergence of strong motion direction selectivity after typical transitions of gray and white matter assessed by silence or abundance of multiunit activities for which the coordinates of penetrations were guided by the structural MRI scans.

For the most part, recordings were made first in V1, next in MT, and finally in VIP. Because MT and VIP were accessed using guide tubes that penetrated V1, the V1 representation was necessarily compromised during MT and VIP recordings (as is common in electrophysiology experiments). However, because MT and VIP receptive fields were offset from the affected V1 representation and both the V1 representation and the visual stimuli were very small, the results from MT and VIP should be unaffected by the approach. Consistent with this, behavioral thresholds during MT and VIP recording were the same as during V1 recording (see Fig. 2B).

When a single unit was isolated, its receptive field location was determined and tuning for motion direction, spatial, and temporal frequencies were characterized before the main task was introduced. The receptive field location was mapped by manually sweeping bars or Gabor patches on the monitor while the monkey fixated. The spatial and

temporal tuning properties were characterized in separate blocks of trials on which Gabor patches were presented multiple times (240 ms on followed by 250 ms off) in the receptive field. Direction tuning was usually measured first with spatial and temporal frequencies set to values that drove the neuron well. For most neurons, direction tuning was also characterized using Gabor patches with identical size, spatial and temporal frequencies to the stimuli used in the subsequent direction-change detection task. Cells were typically tested for 12 equally spaced directions (an average of 20 repetitions of each direction) and for 6 logarithmically spaced spatial and temporal frequencies (an average of 11 repetitions of each value). The responses measured in direction-tuning blocks were fitted with a wrapped Gaussian. In probing the tuning properties of a neuron, the response magnitude was quantified as spike count during a 210 ms interval from 30 ms after the stimulus onset. Further data collection was pursued for all neurons for which a preferred direction could be determined or that had two peaks that were ~ 180 degrees apart (mostly V1 neurons, see below).

In the direction-change detection task (see Fig. 1A), stimuli from the two stimulus sets were tested in separate blocks of trials. Only a limited number of cells could be tested using both sets of stimuli before isolation of the signal was lost or the animal became sated. The direction of reference stimulus was chosen to lie on one flank of the direction-tuning curve (see, e.g., Fig. 1C). For cells that had two peaks in their direction tuning curves (e.g., most V1 cells, which were frequently orientation-tuned), the direction of reference stimulus was chosen from one of the four flanks (usually flanking the higher peak). The direction of the target stimulus on a given trial was pseudo-randomly chosen from a set of fixed values (typically 6) that spanned behavioral threshold for detecting the change (see Fig. 1E). The target directions were always closer to the neuron's preferred direction than the direction of reference stimulus, and thus elicited stronger neuronal responses. On average, 57 valid trials (i.e., hit or miss trials) per target direction were collected. The number of trials for the smallest one or two direction changes were 1.5 times each of the other target directions so that more trials were available for the calculation of detect probability (DP) (see below).

Visual stimuli were usually presented at the center of the receptive field, but for some VIP cells the receptive field borders could not be defined unambiguously, and stimuli were presented at a location where stronger responses could be invoked by stimuli. The fixation point was usually presented at the center of the screen, but for cells with a receptive field of large eccentricity, it was displaced so that stimuli could be presented without truncation.

Data analysis

The responses of a neuron to each stimulus in a trial were quantified as the number of spikes occurred during a 150 ms interval beginning 30 ms after the stimulus onset. Spike counts during this same response window were used to estimate neurometric thresholds (see Fig. 2), DPs and reaction time correlations (RTCs) (see Fig. 4) described below. The onset of the response window was based on the earliest latency of stimulus-driven responses, and the offset was set to minimize saccadic responses because the earliest saccadic reaction times started at ~ 180 ms from stimulus onset (see Fig. 6). A stimulus deemed valid for analysis when the monkey's response to that stimulus fell into one of four categories. The monkey's response to a target stimulus was counted as a hit if the animal made a saccade to the stimulus within the interval between 150 and 600 ms from the target onset, or as a miss if it failed to respond and continued to fixate. The monkey's response to a reference stimulus was counted as a correct rejection when it maintained fixation within the fixation window, or as a false alarm if it made a saccade to the stimulus location. Only the neuronal responses to stimuli to which the monkey made a valid response were included for further analysis.

Psychometric and neurometric thresholds

To quantify the behavioral performance on the direction-change detection task during a recording session, a psychometric threshold was measured by fitting a modified cumulative Weibull distribution to the proportion of correct trials (i.e., hits/(hits + misses)) using the following equation (see Fig. 1E):

$$pc(\theta) = \lambda + (1 - \lambda) \times (1 - \exp(-(\theta/\alpha)^\beta))$$

where θ is the absolute value of the direction change (i.e., difference in direction between target and reference stimulus) on a log scale and $pc(\theta)$ is the proportion correct at θ . The psychometric function had two free parameters that corresponded to the threshold (α) and slope (β). The lower asymptote (λ ; see Fig. 1E, dashed line), was the guess rate and estimated as follows:

$$\text{guess rate} = \sum_i [FA_i \times p(i)],$$

$$FA_i = \frac{\text{number of false alarms to } i\text{th stimulus}}{\text{number of trials in which target appeared after } i\text{th stimulus}}$$

where FA_i is the false-alarm rate at the i th stimulus flash and $p(i)$ is the probability of target occurrence at the i th flash determined by the exponential distribution (see Animal preparation and behavioral task). Because the monkey's performance at the largest direction change was perfect in most sessions (74%), we set the upper asymptote of the psychometric function 100% (i.e., a lapse rate of zero). The mean lapse rate (defined as the miss rate at the largest direction change) was 0.019, and the lapse rate for 80% of the sessions was smaller than the mean.

Neurometric threshold of a neuron was estimated by comparing the responses to the target stimulus with those to the reference stimulus immediately preceding the target using signal detection theory (Green and Swets, 1966). For each target direction, the detection performance of an ideal observer was defined as the area under the receiver-operating characteristic (ROC) curve derived from the distributions of spike counts for the target and reference stimuli. A neurometric threshold was determined by fitting the same Weibull function used to fit the psychometric function, but with λ fixed at 0.5, to the area under ROC curve at different target directions on a log scale (see Fig. 1D). In fitting both psychometric and neurometric functions, maximum-likelihood estimates for the two free parameters were obtained using an optimization algorithm (*fminsearch*) provided by MATLAB (The MathWorks), and 95% CIs were estimated from bootstrap simulations (2000 iterations).

For a small number of neurons with a narrow direction tuning, the range of stimulus values (i.e., direction changes from the reference direction) that we imposed during data collection extended beyond the neuron's preferred direction resulted in a decrease in the mean response because they were on the opposite flank of the tuning curve. This was to ensure a reliable estimation of psychophysical thresholds. However, in the estimation of neurometric thresholds for those neurons, the responses to the stimulus values on the opposite flank of the direction tuning curve were excluded (mostly the responses to the largest direction change and mostly from V1) to avoid underestimation of neurometric thresholds.

Simulation to estimate population neurometric performance

In the simulation estimating population neurometric performance shown in Figure 3, the mean and variance of the responses of synthetic neurons followed those of the neurons recorded during the main task. To determine the mean response of a synthetic neuron to an arbitrary stimulus direction, we used direction tuning curves fitted to the neuronal responses measured in separate blocks of trials (see above). We first selected neurons with a direction tuning well fit by a wrapped Gaussian ($R^2 > 0.5$; see, e.g., Fig. 3A, gray dots and gray curve; see also Fig. 1C). These neurons constituted the simulation pool from which the responses of synthetic neurons were simulated. The number of neurons included in the simulation pool was 132 from V1 (79 from Monkey 1; 53 from Monkey 2), 106 (56; 50) from MT, and 109 (51; 58) from VIP for Set L. For Set S, it was 109 (67; 42) from V1, 71 (27; 44) from MT, and 55 (17; 38) from VIP.

To estimate population performance based on the responses of neurons during the main task, for each selected neuron, we scaled the direction tuning curve vertically (without changing the parameters specifying the location of the peak response and the width of the Gaussians) to fit the firing rates of the neuronal responses measured during the main task (i.e., spike counts during the 150 ms response window; for an example MT neuron, see Fig. 3A, black dots and curve). If available, we used the

direction tuning curve assessed using Gabor patches of the same size and spatial frequency that were used in the main task.

We also wanted to use the neuronal responses measured during the main task to assign the response variance of synthetic neurons. However, we found that, for most neurons, Fano factor (i.e., ratio of the variance to the mean) of the neuronal responses measured during the main task was not constant, but rather changed systematically as a function of the mean response (see, e.g., Fig. 3C). Therefore, to assign the response variance of a synthetic neuron at an arbitrary stimulus direction, we used a regression of the response variance over the mean response weighted by the square root of the number of trials repeated for each stimulus direction (see Fig. 3B, gray line). If available, the neuronal responses measured with both stimulus sets were included to derive the regression line (e.g., Fig. 3B,C plots the responses of the MT neuron shown in Fig. 3A measured with both stimulus sets).

In each iteration of the simulation to estimate the population performance of a given cortical area for a given stimulus set, a pool of synthetic neurons was generated by randomly sampling neurons from a subset of the simulation pool (i.e., neurons from the same cortical area tested with the same stimulus set). Then, the scaled direction tuning curve of each synthetic neuron was shifted horizontally by a random amount chosen from $[0^\circ, 360^\circ]$ to distribute preferred directions uniformly. Population responses of synthetic neurons for two stimulus directions differing by 3° were generated by random sampling from multivariate normal distributions where the means were set from the tuning curves and the covariance was a diagonal matrix having the response variances determined from the variance-mean regression as diagonal entries. Therefore, the responses of synthetic neurons were independent. Finally, the population neurometric performance was quantified as d' calculated from the population responses for the two stimulus directions that were projected onto the axis determined by Fisher's linear discriminant (Bishop, 2006).

The number of iterations was inversely proportional to the population size such that the expected frequency of each model neuron to be included in the simulation was the same for different population size. For example, the number of iterations was 10,000 for sample size of 1, 100 for sample size of 100, and 20 for sample size of 500.

Trial-to-trial correlations between neuronal activity and behavior

Trial-by-trial correlation between the neuronal response and the monkey's perceptual behavior was quantified with DP and RTC (Cook and Maunsell, 2002). In measuring DP or RTC (see Fig. 4), it is critical to maintain the same stimulus conditions across trials, so that the correlations are not driven by the variation in the stimulus. Because the behavioral task in this study used Gabor patches with several different directions, these correlations could be estimated from the neuronal responses to each stimulus direction as long as sufficient trials are available for both hits and misses. To estimate single correlation measures for individual neurons with an increased statistical power, the neuronal responses to Gabor patches of different directions were combined after the responses were normalized to remove the stimulus effects on the calculation of DPs and RTCs.

In previous studies in which detect probabilities or choice probabilities were estimated from trials pooled across different stimulus conditions, the neuronal responses were usually normalized within stimulus conditions to remove the stimulus-induced variance (see Kang and Maunsell, 2012). However, we found that, for some neurons, the response to an identical stimulus varied systematically with the stimulus position in the stimulus sequence within trials: the neuronal response to the same target stimulus increased as the target appeared later in a trial. Moreover, in some sessions, the monkey's detection performance improved gradually as target stimuli appeared later in a trial. These concurrent modulations in the neuronal response and behavioral performance can lead to a potentially spurious DP (Kang and Maunsell, 2012). Indeed, we found that the neuronal responses recorded from Monkey 2 to the same stimulus increased noticeably as a function of the order of the stimulus presented on a trial in all three cortical areas for both stimulus conditions. Moreover, both monkeys tended to respond more frequently and faster to target stimuli as they were presented later on a trial,

which was indicated by increasing hit rates and decreasing reaction times with the order of stimulus presentation on a trial. To avoid the confound in the estimation of DP and RTC caused by these concurrent nonstationarities of the neuronal and behavioral responses over stimulus presentation order, the neuronal responses to the same stimulus were normalized within each stimulus position in the stimulus sequence on trials. For example, the neuronal responses to a target stimulus on trials when it was presented at the third flash were normalized separately from those on trials when it was presented at the fourth flash, and so on. These normalized neuronal responses were then pooled across different positions in the stimulus sequence and across different stimulus directions.

If trial-by-trial correlations between the animal's perceptual reports and the neuron's responses to target stimuli in our task were purely driven by the shared noise in the neuronal responses, then one can expect that there would similarly be a correlation between the neuronal response and behavior to reference stimuli (Price and Born, 2010). To increase the statistical power, the neuronal responses to reference stimuli were combined with those to target stimuli in the estimation of DPs and RTCs. That is, taking the animal's responses to reference stimuli as those to zero direction change, the neuronal responses during epochs of correct rejections were combined with those of misses and the neuronal responses during epochs of false alarms were combined with those of hits.

In normalizing the neuronal responses for the estimation of DPs, spike counts during the 150 ms response window were converted to z scores using a modified method (Kang and Maunsell, 2012) to correct for underestimation due to combining samples with uneven numbers of trials for the two behavioral categories across different stimulus conditions. For each stimulus position in the sequence on a trial, neuronal responses were included in the calculation of DPs only when at least 5 samples were available for each of the two behavioral response categories (i.e., hit and miss for a target stimulus; false alarm and correct rejection for a reference stimulus). The normalized neuronal responses pooled across different stimulus directions were then compiled into two distributions: one combining the neuronal responses to reference stimuli followed by correct rejections and those to target stimuli followed by misses; the other combining the neuronal responses to reference stimuli followed by false alarms and those to target stimuli followed by hits. DP was defined as the area under the ROC curve derived from these two distributions. The p value of a DP was estimated from a permutation test with 2000 iterations.

To measure RTCs of individual neurons, we combined hit trials across different target directions, in which the monkey responded to target stimuli, as well as false alarm trials in which the monkey responded to reference stimuli before a target appeared. Neuronal responses were quantified as spike counts during a window spanning 150 ms starting 30 ms after stimulus onset (the same window used for estimation of neurometric threshold and DP). Because concurrent modulations in the neuronal response and behavioral performance within trials can also introduce a confound in RTCs (Kang and Maunsell, 2012), the neuronal responses and reaction times were converted into z scores within each position in the stimulus sequence on a trial before they were pooled across different positions and target directions. Then the Pearson correlation coefficient was calculated from the combined z scored responses. Trials within a stimulus position had to have at least 5 samples to be included in the calculation of RTCs. Reaction time was defined as the time at which the speed of the eye position signals exceeded 20 degrees/s (or 30 degrees/s for some V1 cells from Monkey 2 due to higher noise level in the eye position signals). For both DP and RTC, if the mean spike count during the 150 ms response intervals at a given stimulus direction was <1 (equivalent to a mean firing rate of 6.7 spikes/s), then those intervals were excluded from further analyses.

The grand DP and RTC shown in Figure 4B, D were calculated from the z -scored spike counts described above that were pooled over cells separately for the two behavioral categories (i.e., correct rejections + misses vs false alarms + hits). The significance of the difference in the mean DP and RTC between the stimulus sets in each cortical area was evaluated by a permutation test with 10,000 iterations. More specifically, under the null hypothesis that the neuronal responses for the two

behavioral categories were not different between the two stimulus sets, for each cortical area we first combined the pooled *z*-scored spike counts for the two behavioral categories over the two stimulus sets. In each iteration, for each cortical area separately, we permuted the pooled *z* scores within a behavioral category and arbitrarily divided them into the two stimulus sets in proportion to the number of samples of the original data. We then calculated DP or RTC from the permuted data and obtained the difference between the two stimulus sets. After we repeated this 10,000 times, we took the *p* values of the variables of interest as the proportion of the differences from the synthetic data that were equal to or greater than the observed difference.

To evaluate the interaction effects between cortical area and stimulus condition on DP and RTC, we fit linear mixed effect regression models predicting DP and RTC using cortical area and stimulus set as fixed effects and cell as a random effect where the stimulus set is a repeated variable for cells within cortical areas. The significance of the interaction by cortical area and stimulus set was tested by the likelihood ratio test between the null model without the interaction term and the full model including the interaction by cortical area and stimulus set.

To estimate the time courses of the correlations between the neuronal response and behavior (see Fig. 6), DP and RTC were calculated with 1 ms resolution after spike times on individual trials were smoothed with an exponential filter (time constant of 25 ms). To minimize the carryover effect of smoothing, the tail of the filter was truncated at 1% of the peak of the kernel (the total length of the kernel = 231 ms). Although smoothing the neuronal responses was necessary for the estimation of the time courses, it was also crucial to use a shortest filter as possible with its tail truncated. Otherwise, the time courses will be prolonged and exaggerated after the stimulus offset during which no significant neuronal responses are present because both correlation measures are sensitive to the rank of the variables.

To remove the variance of the neuronal response due to the stimulus order in time during a trial (Kang and Maunsell, 2012) as well as the stimulus-dependent variance, DPs and RTCs were calculated at every millisecond from the filtered neuronal responses separately for trials in which the same stimulus change occurred at the same stimulus position in time during a trial, and then averaged across the stimulus strength and position in time. The time courses of RTC in Figure 6D–F were estimated from the neuronal responses and reaction times on both hit and false-alarm trials (see Fig. 6A–C), and DPs estimated from the neuronal responses to reference stimuli (false alarm vs correct rejection) as well as those to target stimuli (hit vs miss) were included in the time courses shown in Figure 6G–I. To estimate the peristimulus time histograms of the neuronal responses conditioned on the animal's perceptual response (see Fig. 6J–L), the filtered neuronal responses on hit (or false alarm) and miss (or correct rejection) trials were averaged separately for each stimulus position in time during trials on which the stimulus direction changed by the same amount, and then these mean responses conditioned on behavior were averaged across stimulus conditions.

At least 5 trials were required for each behavioral category (hit + false alarm and miss + correct rejection) for the estimation of the time course for a given stimulus strength (i.e., the amount of direction change) and stimulus position in time. Cells for which the average firing rate during 150 ms after stimulus onset was lower than 1 spike/s were excluded. For both DP and RTC, trials in which the animal made a response earlier than 150 ms after the stimulus onset were excluded. To pool data from 2 monkeys, the time courses were estimated for each monkey first and then averaged together. The number of trials per stimulus strength and stimulus position in time during a trial was 13 on average (minimum = 5, maximum = 136, SD = 9) for RTC and 38 for DP (62 misses + correct rejections (minimum = 5, maximum = 692, SD = 89), 13 hits + false alarms (minimum = 5, maximum = 136, SD = 11)). The 95% CIs of the time courses of DP and RTC were estimated using bootstrap simulations with 2000 iterations.

To examine whether the neuron-behavior correlations measured in our study reflected fluctuations in neuronal responses arising from extraretinal factors, such as attention or overall arousal level, we calculated DPs and RTCs from the neuronal responses to stimuli immediately preceding the stimuli from which DPs and RTCs in Figure 4 in the same way described above, assigning the same behavioral categories.

Effects of fixational eye movements and the mean firing rate on neuron-behavior correlations

To see whether DPs of V1 neurons measured with stimulus Set S were confounded by the number of small eye movements during fixation that might have been different depending on perceptual decision, we measured small eye movements during the spike counting interval. Eye position signals were convolved with a Gaussian kernel (SD range, 1.6–7.5 ms), and fixational eye movements were detected using a velocity criterion (2 deg/s with a minimum peak velocity of 6 deg/s). Eye movements could be detected for 91 sessions (62 from Monkey 1 and 31 from Monkey 2) of 105 sessions shown in Figure 4A. The quality of eye position signals from the remaining sessions did not allow reliable detections of small eye movements due to high-frequency noise. To compare the neuronal responses during epochs with no fixational eye movements with those with fixational eye movements (see Fig. 5C), for a given neuron, spike counts during the spike counting interval were converted into *z* scores within a stimulus direction. To control for the eye movement effect on DPs of V1 neurons measured for stimulus Set S (see Fig. 5D), spike counts during the spike counting interval were converted into *z* scores in the same way as DPs of the individual neurons in Figure 4A were estimated, but separately for those containing eye movements and for those containing none. This procedure further reduced the number of trials available for the calculation of DPs; hence, the number of neurons, such that sessions for 85 V1 neurons (55 for Monkey 1 and 30 for Monkey 2) recorded with stimulus Set S could be included.

To examine whether neuron-behavior correlations were affected by the response magnitude driven by stimuli, we compared the mean firing rates of cells across three cortical areas that were included to estimate DPs and RTCs shown in Figure 4. We found that, within a stimulus condition, the mean firing rates of cells from three cortical areas did not differ significantly, except for those included in RTC analyses measured with stimulus Set S, where the mean firing rate of V1 cells were higher than those of MT and VIP neurons (see Results). To see whether the increase in RTC of V1 cells for stimulus Set S was due to the higher mean firing rate of V1 cells, we calculated RTC for subsamples of cells for which the mean firing rate was matched across the three cortical areas. To match the mean firing rate, we first built histograms of the mean firing rates of cells in each cortical area with a bin size of 6.7 spikes/s. For each bin, we randomly sampled the same number of cells from all three areas that belonged to the same bin such that the sample size for that bin was determined by the minimum number of cells belonged to the bin from the three areas. In this way, we randomly subsampled 57 cells from each area and calculated the mean and SE of RTCs of resampled cells. We repeated this resampling 2000 times.

All data analyses were done using MATLAB (The MathWorks), except for the linear mixed effect model analyses used to test the significance of the interaction effects between cortical area and stimulus condition on DP and RTC, for which we used R (www.cran.r-project.org).

Results

We recorded well-isolated spikes from individual neurons in V1, MT, and VIP in 2 macaque monkeys while they performed a direction-change detection task (Fig. 1A). In this task, a small Gabor patch flashed multiple times in the neuron's receptive field, drifting in the same selected direction, and the monkey had to make a saccade to the stimulus location when it appeared drifting in a different direction on one randomly chosen flash.

Our primary goal was to compare the responses of neurons in different visual cortical areas to a given stimulus. In particular, we wanted to compare neuronal responses using a stimulus that was effective for MT and VIP, but which might nevertheless evoke responses in individual V1 neurons that were more informative for executing the task than those of individual neurons in MT or VIP. To this end, we chose a drifting Gabor patch as the task stimulus because drifting gratings are one of the most effective stimuli for driving responses in V1 neurons (De Valois et al., 1982; Ringach et al., 2002), but are also very effective in

Figure 2. Psychometric and neurometric thresholds. **A**, Distributions of neurometric thresholds of individually recorded neurons (left) and simultaneously obtained psychometric thresholds (right). Responses were collected from V1 (top row; 130 cells, 79 from Monkey 1, 51 from Monkey 2), MT (middle row; 104 cells, 54 from Monkey 1, 50 from Monkey 2), and VIP (bottom row; 102 cells, 47 from Monkey 1, 55 from Monkey 2) using stimuli from stimulus Set L. **B**, Average psychometric and neuronal thresholds. Geometric means of psychometric (gray) and neurometric (black) thresholds in **A** are plotted as a function of cortical area. Error bars indicate $\pm$ SEM of log threshold. **C**, **D**, Data obtained using stimuli from stimulus Set S. The same conventions as in **A** and **B**. The number of cells included in **C** and **D** are 114 for V1 (71 from Monkey 1, 43 from Monkey 2), 71 for MT (27 from Monkey 1, 44 from Monkey 2), and 57 for VIP (16 from Monkey 1, 41 from Monkey 2). Cells included in the figure had at least 10 trials available for each target direction (average 58 trials). Cells for which a neurometric threshold could not be determined because of a poor fit or those with a threshold >180 degrees are excluded from the analysis.

driving responses in MT (Movshon et al., 1985; Priebe et al., 2003) and VIP (Avillac et al., 2007).

Each neuron recorded was tested with the Gabor centered in its receptive field. Unlike previous studies, we did not adjust the size of the Gabor to fill the receptive field of each neuron recorded. Rather, the stimulus size and its spatial frequency were solely determined by the eccentricity of the receptive field (Fig. 1B). We did not use the same stimulus for all neurons because size and spatial frequency preferences scale with receptive field eccentricity, and the receptive field locations of sampled neurons might differ between the cortical areas. Setting the size and its spatial frequency based on eccentricity provided more consistent comparisons across recording sites at different eccentricities than using a single Gabor size for all measurements. Another purpose of scaling the stimulus with retinal eccentricity was to obtain comparable perceptual performance across the visual field (Rovamo et al., 1978). Thus, the width of Gabor increased and the spatial frequency decreased with receptive field eccentricity so that the stimulus contained the same number of cycles of the carrier sinusoid (Fig. 1B). The stimulus scaling was the same during recordings from all three visual areas. We adopted two scaling factors (Fig. 1B; see Materials and Methods). However, in both cases, the stimulus size was small, such that the stimulus

covered the entire receptive field of V1 neurons, but only a small portion of the receptive fields of MT and VIP neurons.

To assess how neuronal activity was related to behavioral performance, we wanted to examine the responses of neurons that were best equipped to represent the stimulus. We therefore set the drift direction to maximize differences in spiking from each neuron recorded. The drift direction of the Gabor before the change (the reference stimulus) was fixed throughout the recordings for each neuron, chosen to lie on a flank of the direction-tuning curve of the neuron. This maximized the change in neuronal response for small direction changes (Fig. 1C). The direction change was always toward the neuron's preferred direction so that the neuronal response to the target stimulus would be stronger than the response to the reference stimulus. An example of a reference direction chosen for an MT neuron is shown in Figure 1C (solid vertical line), together with the neuron's preferred direction (dashed line), which was derived from a fitted tuning curve (thick gray line).

To explore whether trial-to-trial correlations between behavior and neuronal responses in the three cortical areas interact with manipulations of stimulus parameters, we tested cells with two stimulus settings that differed in their eccentricity scaling factors. At any given retinal eccentricity, the SD of Gabor function in the smaller stimulus set (Set S) was two-thirds that of the larger stimulus set (Set L) and the spatial frequency was 3 times higher (Fig. 1B; see Materials and Methods).

For both stimulus sets, we expected that among the three cortical areas V1 neurons would carry the most reliable signals for the task because the spatial frequency was set closer to those preferred by V1 neurons than those preferred by MT or VIP neurons: for eccentricities from 3° to 20° , spatial frequency ranged from 0.5–3.3 cycles/degree in Set L, and from 1.5–10.0 cycles/degree in Set S. We also expected that the stimulus manipulations in Set S (i.e., decrease in size and increase in spatial frequency) would reduce response strength and therefore the statistical reliability of the neuronal responses (Tolhurst et al., 1983) in all three areas. However, because spatial frequency in Set S was further away from those preferred by MT and VIP neurons than those preferred by V1 neurons, we expected that the reduction in statistical reliability of the neuronal responses might be more pronounced for MT and VIP than for V1, making the response of MT and VIP neurons even less informative for the task relative to V1 when stimuli from Set S were used.

V1 neurons carried the most reliable signals for the task

To quantify the degree to which the animal could rely on the response of a neuron to perform the task, a neurometric threshold was estimated by comparing the responses to the target stimulus with those to the reference stimuli immediately preceding the

target stimuli. For each unit recorded, the monkey was typically tested on different trials with 6 different target directions, which were randomly interleaved and spanned the threshold for behavioral detection (Fig. 1E). To measure a neurometric threshold, the neuronal responses to the target and preceding reference stimulus on each trial were quantified as the number of spikes that occurred during a 150 ms interval starting 30 ms after stimulus onset. Then, for each target direction, the performance of an ideal observer detecting the direction change based on the neuronal responses was estimated as the area under the ROC curve (Green and Swets, 1966) derived from spike-count distributions for the target and reference stimuli. A neurometric threshold was defined as the direction change corresponding to an 82% detection rate determined from a cumulative Weibull function fitted to the performance of the ideal observer estimated for each target direction. Figure 1D plots the performance of the ideal observer (filled circles) estimated from the responses of the MT neuron in Figure 1C to stimuli from stimulus Set L, and the fitted neurometric function (gray line). The neuronal threshold for detection was 14.3° (arrow, 11.1° – 18.1° , 95% CI marked by the bracket).

From behavioral responses that were simultaneously obtained with the neuronal data, a psychophysical threshold was also measured by fitting a cumulative Weibull function to the monkey's performance at different direction changes (Fig. 1E). We estimated the monkey's chance detection rate (Fig. 1E, dashed line) from the false alarm rate within each session (see Materials and Methods) and used it as the lower asymptote of the fitting function (0.05 for the session in Fig. 1E, median 0.08 for 578 sessions in Fig. 2), and the threshold was defined as the direction change corresponding to a 63% detection performance of the fitted psychometric function (i.e., the actual hit rate at threshold was slightly higher owing to the false alarm rate). The behavioral threshold in this example was 7.6° (arrow, 6.9° – 8.4° , 95% CI marked by the bracket). As was typical for most cells, the behavioral performance was superior to the neurometric performance of this MT neuron.

Figure 2 shows distributions of neurometric thresholds of individual neurons recorded from the three cortical areas in 2 monkeys and those of simultaneously measured behavioral thresholds. Overall, neurometric thresholds (Fig. 2A,C, left columns) were higher and more variable than behavioral thresholds (Fig. 2A,C, right columns) for stimuli both from Set L (Fig. 2A) and Set S (Fig. 2C), although some neurons had a threshold comparable with the behavioral threshold. A few V1 neurons had thresholds superior to the behavior. The distributions of the neuronal and behavioral thresholds are summarized by their geometric means in Figure 2B and D, which compare the difference across three cortical areas. Superior behavioral performance is consistent with previous studies that compared the sensitivity of individual neurons with the behavior measured in reaction time tasks similar to ours (Cook and Maunsell, 2002; Cohen and Newsome, 2009; but see Palmer et al., 2007).

Although the distributions of neurometric thresholds largely overlapped between the three cortical areas, for both stimulus sets individual V1 neurons were on average most sensitive and individual neurons in VIP were least sensitive. Average neuronal thresholds were higher in all areas for stimuli from Set S (Fig. 2D) compared with Set L (Fig. 2B). However, the increase was significant only for MT and VIP neurons ($p > 0.48$ for V1, $p < 10^{-4}$ for MT, and $p < 10^{-5}$ for VIP, t test carried on log thresholds, Bonferroni-corrected for three simultaneous comparisons), and was greatest for VIP.

In contrast to neurometric thresholds that differed systematically between the areas, behavioral thresholds were relatively uniform across the three cortical areas, despite the stimuli being presented at different visual field locations due to the diverse receptive field locations of the neurons sampled from the three cortical areas (for distributions of the receptive field locations, see Materials and Methods). This suggests that the scaling of stimuli with eccentricity was successful in achieving similar behavioral performance across recordings from the three cortical areas, excluding the possibility that the differences in the neurometric threshold between the cortical areas arose from the difference in the receptive field eccentricities of sampled neurons. Somewhat surprisingly, behavioral thresholds were little affected by using the smaller stimuli (but see the following section). Behavior thresholds were elevated only for V1 neurons with stimulus Set S. For practical reasons, the stimuli from the two stimulus sets were not randomly interleaved within a block of trials, and it is likely that the animals compensated for the more difficult stimuli by increasing their effort (Boudreau et al., 2006).

Overall, among the three cortical areas, the responses of individual V1 neurons carried statistically the most reliable signals that the monkey could use to perform the task in both stimulus conditions. Additionally, the thresholds of MT and VIP neurons increased markedly when measured with stimulus Set S, while the corresponding behavioral thresholds were similar in both stimulus conditions.

Differences between cortical areas in neurometric sensitivity measured for single neurons are relevant for the population level
In measuring neurometric thresholds of individual neurons shown in Figure 2, stimulus directions were optimized for each neuron to probe the best discrimination performance it could offer, which depends on the slope of tuning curve and the response variance. Therefore, the superior neurometric performance of V1 neurons probably reflects a narrower tuning bandwidth and/or a higher response gain of V1 neurons than MT or VIP neurons. However, at population level, where some cells preferring directions nonoptimal for the task are presumably pooled as well, the differences in tuning bandwidth and response gain might be less important. Although we do not have data to test this possibility directly, we performed a simple simulation to gain insight about the population performance of the three cortical areas.

We estimated population neurometric performances from the pooled responses of synthetic neurons whose response properties were modeled on those of the recorded neurons. In this simulation, the mean firing rate of a synthetic neuron to an arbitrary stimulus direction was determined from the direction tuning curve fitted to responses from a real neuron (Fig. 3A, gray curve) scaled to fit the responses measured during the main task for that neuron (Fig. 3A, black dots and curve). The response variance of the synthetic neuron was determined from a regression line relating the variance of spike counts to the mean spike counts measured from the same real neuron at different stimulus directions, and with different stimulus sets if available, during the main task (Fig. 3B).

In each iteration of the simulation, a pool of synthetic neurons was generated from randomly selected real neurons, but the tuning curve of each synthetic neuron was shifted horizontally by a random amount chosen from $[0^\circ, 360^\circ]$. Then population responses for two stimulus directions differing by 3° were generated by random sampling from multivariate normal distributions with the means and covariance specified by the direction tuning

Figure 3. Simulation of population neurometric performance. **A**, Direction tuning curve of a synthetic neuron. Gray dots represent the mean firing rates of an MT neuron measured using Gabor patches with a size and spatial frequency set to the scaling of stimulus Set L. Gray curve indicates the fitted wrapped Gaussian. Black dots represent the mean firing rates based on spike counts during the 150 ms response window measured in the main direction-change detection task using Gabor patches from stimulus Set L. To determine the mean response of a synthetic neuron to an arbitrary stimulus direction, the tuning curve of the model neuron was scaled vertically to fit the responses measured during the main task (black curve). **B**, Variance of mean spike counts. The variance of the mean spike counts of the same MT neuron measured during the main task using both stimulus sets are plotted against the mean spike counts. Gray line indicates a weighted regression of the variance over the mean spike counts, which was used to determine the response variance of a synthetic neuron. **C**, Fano factor of mean spike counts. Fano factor of the mean spike counts for the same MT neuron (i.e., the ratio of the variance to the mean for spike count data shown in **B**) is plotted against the mean spike count. **D**, Population neurometric performance for Set L. Population neurometric performance discriminating two stimulus directions (178.5° vs 181.5°) quantified by d' is plotted as a function of population size for three cortical areas. **E**, Population neurometric performance for Set S. Same format as in **D**. Gray horizontal lines at a d' of 1 are drawn to facilitate the comparison of the population performance of three cortical areas within and between stimulus conditions. SEMs are not shown, but constant on a log scale over population size and smaller than the symbol size in **D** and **E**.

curves and the variance-mean regressions, respectively. Finally, the population neurometric performance was quantified as d' calculated from the population responses for the two stimuli directions that were projected onto the axis determined by Fisher's linear discriminant (Bishop, 2006).

Population neurometric performances simulated from the responses of neurons recorded from three cortical areas were plotted as a function of population size separately for Set L (Fig. 3D) and Set S (Fig. 3E). In all cases, the d' increased linearly with the square root of sample size, which is expected because the responses of synthetic neurons were independent. However, although the population performance was estimated from the responses of synthetic neurons with uniformly distributed preferred directions, the simulation replicated the observations made in Figure 2: the order of three cortical areas in neurometric sensitivity and the difference between the two stimulus conditions, which is manifested by vertical offsets of the lines in Figure 3D, E. Therefore, V1 neurons carry the most reliable signals for the task at population level unless one assumes a correlation

structure that would selectively degrade the population performance of V1 neurons, which is highly unlikely given that both MT and VIP neurons are likely to receive considerable amount of input from V1 either directly or indirectly. In other words, to explain behavioral performance from the responses of MT or VIP neurons, one would have to assume a decorrelation mechanism that removes noise correlations aligned with stimulus tunings (i.e., differential correlations) (Moreno-Bote et al., 2014) present in V1 population. One interesting finding from the simulation is that the population performance of V1 neurons for stimulus Set S was slightly better than that for Set L, which might explain why psychophysical thresholds differed little between the two stimulus conditions (Fig. 2), although this could be due to the difference in the animal's effort (see Discussion).

Trial-to-trial correlations between neuronal response and behavior across cortical areas interact with stimulus condition

We next examined the degree to which the neuronal responses in these cortical areas were correlated with the monkey's detection behavior on trial-by-trial basis. For this, we measured correlations between the neuronal response and two aspects of the monkey's behavioral response: perceptual report and reaction time.

Trial-by-trial correlation between the neuronal response and the monkey's perceptual report was quantified with DP (Cook and Maunsell, 2002; Bosking and Maunsell, 2011), which reflects the probability that one can correctly predict the monkey's detection behavior on single trials from the

responses of the neuron being recorded. DP is analogous to CP in a discrimination task, which has been often used to assess the causal relation of the neuronal response to a perceptual behavior (Britten et al., 1996; Nienborg et al., 2012). The trial-to-trial correlation between the neuronal response and the animal's reaction time (RTC) has also been used as a measure to probe a linkage between a sensory cortical area and behavior (Cook and Maunsell, 2002; Masse and Cook, 2008; Cohen and Newsome, 2009; Price and Born, 2010; Bosking and Maunsell, 2011). This correlation is expected if the neuronal response contributes to the perceptual behavior and the variation in the sensory process is a source for the variation in the motor response, including reaction time (Osborne et al., 2005).

Figure 4 shows distributions of DP (Fig. 4A) and RTC (Fig. 4C) for individual neurons recorded from the three cortical areas (columns) in the two stimulus conditions (rows). To calculate DP and RTC, the neuronal response to each stimulus on a trial

Figure 4. Correlations between the neuronal responses and behavior. **A**, Distributions of DPs estimated from the responses of individual neurons in three cortical areas (columns) and two stimulus conditions (rows). Cells included in the figure had at least 10 samples for each of the two response categories (average 636; 523 correct rejections or misses, 112 false alarms or hits). Numbers above each distribution are mean DPs, and the numbers of cells in each distribution are denoted by *n*. Darker histograms represent cells having DP significantly different from 0.5 ($p < 0.05$, permutation test), and their numbers are denoted by *n**. The statistics testing the null hypothesis that the mean DP is 0.5 for stimulus Set L were as follows: V1-SD = 0.057, $t_{(118)} = 1.53$, $p > 0.1$; MT-SD = 0.054, $t_{(106)} = 4.37$, $p < 10^{-4}$; VIP-SD = 0.078, $t_{(100)} = 6.17$, $p < 10^{-7}$; and for stimulus Set S: V1-SD = 0.077, $t_{(104)} = 5.07$, $p < 10^{-5}$; MT-SD = 0.082, $t_{(72)} = 1.94$, $p > 0.05$; VIP-SD = 0.073, $t_{(67)} = 5.61$, $p < 10^{-6}$. The statistics testing whether the mean DP within a cortical area differed between the two stimulus conditions were as follows: V1- $t_{(222)} = 3.37$, $p < 0.001$; MT- $t_{(178)} = -0.40$, $p > 0.68$; VIP- $t_{(167)} = 0.19$, $p > 0.84$. **B**, Grand DPs estimated for populations of neurons in three cortical areas are plotted separately for stimulus Set L (filled circles) and Set S (open circles). An average of 60,822 samples of the neuronal responses (50,039 for correct rejections or misses, 10,782 for false alarms or hits) collected from 100 recording sessions went into the grand DP for a given stimulus set and cortical area. Error bars indicate 95% CIs determined by bootstrap. **C**, Distributions of RTCs are shown in the same format as in **A**. RTCs of the neurons included in the distributions were estimated from an average of 245 neuronal response and reaction time pairs, with a minimum of 10. The statistics testing the null hypothesis that the mean RTC is zero for stimulus Set L were as follows: V1-SD = 0.117, $t_{(138)} = -2.31$, $p = 0.023$; MT-SD = 0.104, $t_{(113)} = -7.43$, $p < 10^{-10}$; VIP-SD = 0.155, $t_{(109)} = -7.02$, $p < 10^{-9}$; and for stimulus Set S: V1-SD = 0.122, $t_{(117)} = -7.20$, $p < 10^{-10}$; MT-SD = 0.119, $t_{(79)} = -3.72$, $p < 0.001$; VIP-SD = 0.162, $t_{(76)} = -4.97$, $p < 10^{-5}$. The statistics testing whether the mean RTC within a cortical area differed between the two stimulus conditions were as follows: V1- $t_{(255)} = -4.04$, $p < 10^{-4}$; MT- $t_{(192)} = 1.41$, $p > 0.1$; VIP- $t_{(185)} = -0.51$, $p > 0.6$. **D**, Grand RTCs plotted in the same format as in **B**. The grand RTC for a given stimulus set and cortical area was calculated from an average of 26,036 neuronal responses and reaction time pairs collected from 107 recording sessions. Error bars indicate 95% CIs determined using the Fisher transformation.

was quantified as the number of spikes occurred during a 150 ms interval starting 30 ms after stimulus onset (the same interval for which the neurometric performance was measured in Fig. 2). Because each neuron was tested with several target directions, DP and RTC could be calculated for each target direction for which sufficient numbers of hit and miss trials were both available. To obtain single estimates of DP and RTC for each neuron and to increase the statistical power of these estimates, we combined the neuronal responses to target stimuli across different target directions, as well as those to reference stimuli. That is, we combined the neuronal responses to reference stimuli when the

animal made a false alarm response with those to target stimuli when the animal made a hit response, and the neuronal responses to reference stimuli when the animal made a correct rejection with those to target stimuli when the animal made a miss (for details of combining the neuronal responses across different stimuli and behavioral categories, see Materials and Methods).

When the monkeys performed the detection task with stimuli from stimulus Set L, the correlation between neuronal responses and detection behavior grew progressively stronger from V1 to MT and VIP (Fig. 4A, top row). The mean DP of VIP neurons

was 0.55, whereas those of the MT and V1 populations were 0.52 and 0.51, respectively. Among these, the mean DPs of VIP and MT neurons differed statistically from 0.50 ($p < 0.001$; t test, Bonferroni-corrected for 6 simultaneous comparisons; p values of t tests reported below are all Bonferroni-corrected for multiple comparisons unless otherwise stated; for individual statistics and uncorrected p values, see Fig. 4 legend). Consistent with DP, the mean RTC magnitude was largest in VIP (-0.103), smallest in V1 (-0.022), and intermediate in MT (-0.072 ; Fig. 4C, top row). The mean RTC were significantly different from zero for MT and VIP ($p < 10^{-8}$ for both MT and VIP; t test), but not for V1. These results are consistent with previous studies that observed tighter correlations between the neuronal responses and perceptual reports in higher cortical areas than earlier ones (Leopold and Logothetis, 1996; Cook and Maunsell, 2002; Williams et al., 2003; de Lafuente and Romo, 2006; Nienborg and Cumming, 2006; Carnevale et al., 2013).

Greater trial-by-trial correlations with behavior for the responses of neurons in later stages of cortical processing despite the poorer neurometric performance estimated for the same populations of neurons (Fig. 2A) suggest that a tighter correlation between the neuronal response and behavior in higher cortical areas might be a general feature regardless of the statistical reliability of the neuronal response with which the brain can extract information relevant to the perceptual task from individual neurons. However, the results obtained with stimulus Set S differed. The mean DPs of MT and VIP neurons measured with stimulus Set S (Fig. 4A, bottom row) were similar to those measured with the Set L, although the mean DP of MT did not statistically differ from 0.5. However, unlike MT and VIP neurons, DPs of V1 neurons increased appreciably when measured with stimulus Set S: the mean DP was 0.54 and significantly different from 0.50 ($p < 10^{-4}$).

The same interaction between cortical area and stimulus condition was also observed for the correlations between the neuronal response and the monkey's reaction time (Fig. 4C). The mean RTC of V1 neurons became more negative (-0.081) when measured with stimulus Set S, indicating that the responses of V1 neurons became more strongly correlated with the monkey's reaction time, whereas the mean RTCs of MT and VIP neurons changed little: -0.050 and -0.092 , respectively. The mean RTCs of all three areas differed statistically from zero when measured with stimulus Set S ($p < 10^{-9}$ for V1, $p = 0.002$ for MT, and $p < 10^{-4}$ for VIP).

The comparisons of DP and RTC in Figure 4A and C indicated that the correlation between the neuronal responses and perceptual decisions in three cortical areas interacted with the stimulus condition. This interaction was statistically significant when tested using a linear mixed effect model with the stimulus set as a repeated variable ($\chi^2_{(2)} = 7.68$, $p = 0.022$ for DP; $\chi^2_{(2)} = 13.84$, $p < 0.001$ for RTC; see Materials and Methods for details), although it was not always significant for each monkey individually (significant for DP in Monkey 1 and RTC in Monkey 2). Multiple comparisons revealed that only V1 neurons had statistically different mean DPs and RTCs between the two stimulus conditions (t test, $p = 0.003$ for DP and $p < 0.001$ for RTC), and this underlies the significant interaction between the cortical area and stimulus condition.

To quantify the population statistics in a different way, we calculated a grand DP from the z-scored spike counts combined across all cells from the same cortical area tested with the same stimulus set (Fig. 4B; for details, see Materials and Methods). Similarly, a grand RTC was calculated between the z-scored spike

counts and reaction times that were combined across cells within a given cortical area and stimulus set (Fig. 4D). Population grand DPs and RTCs closely followed the corresponding population mean values in Figure 4A and C, recapitulating the interaction effects between cortical area and stimulus condition on the neuron-behavior correlations. When tested with a permutation test, the difference in the grand correlation measures between the two stimulus conditions was statistically significant for V1 ($p < 10^{-4}$ for both grand DP and grand RTC) but not for MT or VIP ($p > 0.05$ for both grand DP and grand RTC). Consistent changes were seen in both animals individually, although the DPs and RTCs for Monkey 1 were further from 0.5 and 0.0, respectively.

It is somewhat surprising to find only a modest DP (0.52) for MT neurons (Fig. 4A,B), although we used stimuli with strong motion energy. One possible explanation for this small magnitude is that we used a short interval (150 ms) to measure the neuronal response, whereas many of previous studies measured CP from longer intervals (typically 1 or 2 s; e.g., Britten et al., 1996; Dodd et al., 2001; Chen et al., 2013). Another possibility is that trial-to-trial correlation between the neuronal response and perceptual decision might be more pronounced in MT for less-structured stimuli such as random dot kinematograms.

Within a cortical area, previous studies often found a modest correlation between CP (or DP) of individual neurons and their neurometric performance on the behavioral task with which the neuron-behavior correlations were measured (Celebrini and Newsome, 1994; Britten et al., 1996; Uka and DeAngelis, 2004; Gu et al., 2007; Price and Born, 2010; Bosking and Maunsell, 2011). We examined whether this correlation existed for our data. Although many of the correlations between DP and neurometric threshold of individual neurons had the expected (negative) sign ($r = -0.24$ to -0.03 , but $r = 0.23$ for V1 neurons measured with Set S), they reached statistical significance in only a few cases (t test, uncorrected for multiple comparisons). The observation that DP and RTC changed in the same way with stimulus condition across the cortical areas suggests that DP and RTC of individual neurons are correlated, which is expected if these two measurements reflect the same source of the variation between the neuronal response and behavior. Indeed, we found a negative correlation between DP and RTC for most of cortical areas in both stimulus condition ($r = -0.67$ to -0.09). This correlation was strongest in VIP ($r = -0.67$ for Set L, -0.62 for Set S).

Other variables that might affect the neuron-behavior correlations: eye movement, mean firing rate, motivation, and perceptual learning

Any variables modulating the neuronal responses can affect DP and RTC if they covary with perceptual decisions. Eye movement is one such variable (Uka and DeAngelis, 2004; Nienborg and Cumming, 2006, 2014; Herrington et al., 2009; Martinez-Conde et al., 2013). Given the response modulations associated with microsaccades (Martinez-Conde et al., 2013), a systematic difference in the number of microsaccades between hit and miss trials could affect the estimation of DP (Herrington et al., 2009). Also, a systematic difference in gaze angle between the two perceptual decisions could influence the estimation of CP and DP. These are of particular concern for V1 neurons because of their small receptive fields relative to the tolerance of gaze deviations allowed in most studies (1.5° – 2° square in our study), and DP and RTC changed significantly only for V1 neurons between the two stimulus conditions in our study (Fig. 4).

To see whether the increase in DP of V1 neurons in the small stimulus condition was confounded with the difference in the eye position during fixation (i.e., the difference in stimulus position relative to the neuron's receptive field), we compared the mean fixational eye position between the two behavioral response categories (i.e., hit + false alarm vs miss + correct rejection). For the sessions in which V1 neurons shown in Figure 4A were recorded with the small stimulus set, the distributions of the mean eye position during fixation were nearly identical for the two behavioral response categories (Fig. 5A,B). The difference in the mean eye position was 0.001° for Monkey 1 and 0.018° for Monkey 2.

We also examined whether small eye movements during fixation affected DP of V1 neurons while the monkeys performed the task with stimulus Set S. For a subset of the V1 neurons shown in Figure 4A that were recorded with the stimulus Set S (61 and 31 cells from Monkeys 1 and 2, respectively), we found that the neuronal response during the spike counting interval was slightly but significantly lower when small eye movements were present than when they were absent for Monkey 1 (a mean difference of -0.08 in z score, t test, $p < 10^{-20}$), but not different for Monkey 2 (a mean difference of 0.009 in z score, t test, $p = 0.58$; see Fig. 5C). Although statistically significant in 1 monkey, these differences are quite small, corresponding to a difference of 0.16 spikes in Monkey 1, and 0.02 spikes in Monkey 2, for which the mean spike counts were 2.80 and 4.04 spikes, respectively. Moreover, the number of fixational eye movements was quite similar for the two behavioral categories, although the difference was statistically significant in 1 animal because of the large number of samples: the means were 0.46 (hits and false alarms; $n = 41,327$) versus 0.45 (misses and correct rejections; $n = 10,747$) in Monkey 1 (t test, $p = 0.002$), and 0.33 ($n = 20,246$) versus 0.35 ($n = 3057$) in Monkey 2 (t test, $p = 0.17$).

To further check that our estimation of DP was not affected by these small differences in the number of fixational eye movements, we recalculated DPs by normalizing the neuronal responses separately for epochs with and without fixational eye movements. For 85 cells (55 from Monkey 1, 30 from Monkey 2) that had enough samples to allow calculation of a DP, the recalculated DPs were not different from the original ones shown in Figure 4A (Fig. 5D). The mean of the recalculated and the original DPs was 0.54 and 0.55 for Monkey 1 (t test, $p = 0.64$), and 0.52 and 0.52 for Monkey 2 (t test, $p = 0.88$; see Fig. 5D). Based on these analyses, we conclude that any differences in fixation eye position or fixational eye movements did not affect our estimation of DPs from the response of V1 neurons measured with stimulus Set S.

Figure 5. Effects of fixation behavior on DPs of V1 neurons measured for stimulus Set S. **A**, For stimulus epochs that were included in the estimation of DPs of V1 neurons for stimulus Set S, the distribution of mean eye positions during the spike counting interval was summarized by the mean (crosses) and contours plotting Mahalanobis distance of 1 (inner ellipses) and 2.4 (outer ellipses) from the mean separately for epochs in which Monkey 1 maintained fixation (i.e., correct rejections to reference stimuli or misses to targets; blue) and those in which it made a response (false alarms or hits; red). The two contours in each distribution would include 39% and 95% of the data points under a bivariate normal distribution. The numbers of epochs in the two distributions are shown in the top left corner in corresponding colors. **B**, Same as in **A**, but for Monkey 2. **C**, Effect of fixational eye movements on the neuronal response. For stimulus epochs that were included in the estimation of DPs of V1 neurons for stimulus Set S, the neuronal responses during epochs in which fixation eye movements were absent (EM absent) are compared with those during which one or more fixational eye movements occurred (EM present). Small fixational eye movements were detected using a velocity criterion (2 deg/s with a minimum peak velocity of 6 deg/s) after eye position signals were convolved with a Gaussian kernel (SD range, 1.6–7.5 ms). Data are shown for 91 (62 from Monkey 1 and 31 from Monkey 2) of 105 cells shown in Figure 4A. **D**, DPs corrected for the effects of fixational eye movements (vertical axis) were plotted against the original DPs (horizontal axis) for 85 V1 neurons (55 for Monkey 1 and 30 for Monkey 2) that were recorded with stimulus Set S. The mean DPs are marked with $\times$ for Monkey 1 and $+$ for Monkey 2. Gray line indicates unity.

Mean firing rate of the neuronal response is another factor that could influence the magnitude of DP (Kang and Maunsell, 2012). We compared the mean firing rates during the spike counting interval between stimulus conditions and across three cortical areas. For the cells included in DP analyses, the mean firing rates measured with stimulus Set L were 35.5, 34.3, and 36.0 spikes/s for V1, MT, and VIP, respectively, and those measured with Set S were 27.9, 23.2, and 27.2 spikes/s. Although the mean firing rate was greater for Set L than Set S within a cortical area ($p < 0.05$, t test, p values Bonferroni-corrected for multiple comparisons), they were not different across the cortical areas within a stimulus condition (one-way ANOVA, $F_{(2,324)} = 0.12$, $p = 0.88$ for Set L, $F_{(2,243)} = 1.47$, $p = 0.23$ for Set S). Also, the variance of the mean firing rate of the cells included in the DP comparison in Figure 4 did not differ across the three cortical areas within a stimulus condition (Levene's test, $F_{(2,324)} = 0.78$, $p = 0.46$ for Set L, $F_{(2,243)} = 1.29$, $p = 0.28$ for Set S).

We did the same analyses for the cells included in RTC comparison in Figure 4. We found that the mean firing rates

measured with stimulus Set L (49.8, 44.0, and 46.3 spikes/s for V1, MT, and VIP, respectively) were also greater than those measured with Set S (40.2, 30.5, and 30.5 spikes/s; $p < 0.05$) within a stimulus condition (t test, p values Bonferroni-corrected for multiple comparisons). Across the cortical areas, the mean and variance of the mean firing rates did not differ significantly for Set L (one-way ANOVA: $F_{(2,360)} = 1.02$, $p = 0.36$; Levene's test: $F_{(2,360)} = 1.27$, $p = 0.28$). For Set S, they differed (one-way ANOVA testing the mean: $F_{(2,272)} = 5.29$, $p = 0.006$; Levene's test for the variance: $F_{(2,272)} = 4.04$, $p = 0.018$). For example, the mean firing rate of V1 cells measured with Set S was significantly higher than MT or VIP ($p < 0.05$, t test, Bonferroni-corrected), whereas those of MT and VIP cells were not different each other ($p = 0.99$, t test). However, the mean RTC values of random resamples of a subset of cells measured with Set S for which the mean firing rate was matched for three cortical areas (57 cells from each area) did not differ from the population values (see Materials and Methods). Therefore, we conclude that the significant change in DP and RTC we observed for V1 neurons between the two stimulus sets is unlikely to be explained by the difference in the mean firing rate across the cortical areas.

We quantified trial-to-trial correlations using DP and RTC. Unlike CP, DP and RTC are susceptible to fluctuations in neuronal responses arising from extraretinal factors, such as attention or overall arousal level (Nienborg et al., 2012). For example, a significant DP (or RTC) would be seen if attention varied from trial to trial and greater attention made neuronal responses stronger and made subjects more likely (or faster) to detect a change in the stimulus. If this were the case, then one could expect that the neuronal response to stimuli before the target would also show a correlation with behavior, although perhaps a weaker one than the response to the target (Cohen and Maunsell, 2009). We examined DP and RTC of the neuronal responses to stimuli presented immediately before the target stimuli. We found in all three areas that the correlations of these neuronal responses with perceptual decision were either absent or very weak. The mean values of DPs measured for preceding stimuli ranged from 0.495 to 0.507 for all cortical areas and stimulus conditions, which did not differ from 0.5 (t test, $p > 0.05$), except for VIP neurons measured with stimulus Set S (0.521, $p < 10^{-3}$). The mean values of RTC measured for preceding stimuli differed from zero only for VIP neurons (-0.020 , $p = 0.02$ for Set L; -0.032 , $p = 0.03$ for Set S), but not for V1 or MT neurons ($p > 0.05$, the mean RTC ranged from -0.018 – 0.003 for both stimulus sets). Moreover, for individual cells, the DP and RTC measured for the two stimuli (i.e., target and the stimulus before target) were mostly uncorrelated ($r = -0.02$ – 0.19 , $p > 0.05$, Pearson correlation coefficient), except for a few cases: DP of MT ($r = 0.36$, $p < 0.01$) and VIP neurons ($r = 0.51$, $p < 10^{-4}$) measured with Set S, and RTC of VIP neurons measured with Set L ($r = 0.29$, $p < 0.01$). These results suggest that it is unlikely that the correlations shown in Figure 4 were driven by trial-by-trial fluctuations in the animal's attentional state or arousal level.

Before data collection, the monkeys were trained for the main behavioral task extensively for stimuli from both stimulus sets at various locations until behavioral performance stabilized, anticipating diversity in receptive field locations between the three cortical areas. Nevertheless, it is possible that data from the three cortical areas and for the two stimulus conditions might have been collected at different time points in perceptual learning given that recordings were made sequentially from V1, MT, and VIP. This is of particular concern because a previous study

reported that CP of MT neurons increases with perceptual learning even while neuronal sensitivity to stimuli is unchanged (Law and Gold, 2008). Therefore, it is conceivable that MT neurons had smaller DP and RTC values than V1 neurons in Set S because the monkeys were less trained for stimuli presented at different locations than the V1 receptive fields. However, for the neurons shown in Figure 4, the geometric mean of psychometric thresholds measured with Set S was higher for V1 sessions than MT sessions in Monkey 1 (14.7° vs 6.0° , $p < 0.01$, t test), and not different from MT sessions in Monkey 2 (6.8° vs 6.8° , $p > 0.05$). Also, those measured with Set L did not differ between V1 and MT sessions for both monkeys (8.1° vs 7.6° , for Monkey 1, 6.9° vs 6.7° for Monkey 2, $p > 0.05$). On the other hand, DP and RTC of V1 neurons might have been greater for Set S than Set L, a critical observation in our study, because the monkeys were more fully trained for Set S. However, the geometric mean of psychometric thresholds from V1 sessions was higher for Set S than Set L in Monkey 1 (8.1° vs 14.7° , $p < 0.01$) while they were similar in Monkey 2 (6.9° vs 6.8° , $p > 0.05$). Therefore, psychometric thresholds were not consistent with the differences in DP and RTC measured between V1 and MT neurons arising from the degree to which perceptual learning progressed.

To further examine whether a significant perceptual learning occurred during each stimulus condition within a cortical area, for the cells included in Figure 4, we compared psychometric thresholds from the early, middle, and late one-third of sessions. For both monkeys, the geometric means of psychometric thresholds were not significantly different across the three groups of sessions (one-way ANOVA, $p > 0.05$), except for 3 cases described below. The geometric means of the late one-third of sessions of V1 neurons in Monkey 1 differed significantly from those of the early and middle one-third of sessions for both stimulus conditions. The mean thresholds for Set L were 7.4° , 6.6° , and 11.3° for the early, middle, and late sessions, respectively ($p < 10^{-6}$ for early vs late sessions; $p < 10^{-8}$ for middle vs late sessions, Tukey-Kramer multiple comparison), and those for Set S were 13.9° , 16.4° , and 8.6° ($p < 0.01$ for early vs late sessions; $p < 10^{-5}$ for middle vs late sessions). In Monkey 2, the mean psychometric threshold of the early MT sessions for Set S (7.8°) was different from the middle (6.2° , $p < 0.01$) but not from the late one-third of sessions (6.6° , $p > 0.05$). Overall, stable psychometric thresholds across the three temporally grouped sessions suggest that no significant perceptual learning progressed over recording sessions within cortical areas, although Monkey 1 might not have been fully trained for Set S. We also found that the mean DP and RTC values were not significantly different across the three temporally grouped sessions (one-way ANOVA, $p > 0.05$), except for DP of V1 neurons for Set S in Monkey 1 in which the mean DP from the early sessions (0.60) was significantly different from the late (0.51, $p < 0.001$), but not from the middle one-third of sessions (0.55, $p > 0.05$). These results suggest that the differences in the neuron-behavior correlations we observed across different stimulus conditions and cortical areas are unlikely explained by the differences in the degree of perceptual learning at the time the data were collected.

Time courses of correlations between the neuronal response and behavior reveal that only V1 neurons show a difference between stimulus conditions

For the analyses above, we measured the neuronal responses during a short interval after stimulus onset before the animal made a response (see reaction time distributions in Fig. 6A–C). That interval was selected to minimize neuronal activity related to motor

Figure 6 Time courses of correlations between the neuronal response and behavior. Time courses of RTC (*D–F*) and DP (*G–I*) of the neuronal responses in V1 (left), MT (middle), and VIP (right) are plotted with the distributions of reaction times on trials that contributed to the estimation of RTC time courses (*A–C*) and the mean firing rates of the neurons included in the estimation of DP time courses (*J–L*). The mean firing rates shown in *J–L* were obtained after spike trains on individual trials were smoothed using an exponential filter. Values for the two stimulus conditions are plotted separately (blue represents Set L; orange represents Set S). *D–I*, Shaded bands represent 95% CIs estimated by bootstrap. *J–L*, The mean firing rates for hit and false alarm responses of the cells included in the estimation of DP time courses are shown in darker color, and those for miss and correct rejection responses are shown in lighter color. *D–L*, Two vertical gray lines indicate the interval during which spikes were counted to measure DPs and RTCs shown in Figure 4.

preparation or response execution, and also feedback from higher centers in our measurement of the neuronal response. Nevertheless, it has been demonstrated that these correlations, especially choice probabilities, tend to increase after the subject's choice has been made, casting a doubt whether these measures indeed reflect the causal effect of the neuronal responses on behavior. (Nienborg and Cumming, 2009; but see Ni et al., 2018). To explore this issue, we examined how these DP and RTC evolved over time in our task (Fig. 6).

We first smoothed spike times on individual trials with an exponential filter, and then calculated DP and RTC at 1 ms resolution based on the filtered neuronal responses. To minimize the carryover effect of filtering on these correlation measures, we used a filter with a short time constant (25 ms) that was truncated at 1% of the peak of the kernel (total kernel length of 231 ms; see Materials and Methods).

In Figure 6, the time courses of RTC (Fig. 6*D–F*) and DP (Fig. 6*G–I*) in the three cortical areas (V1, left; MT, middle; VIP, right column) for the two stimulus conditions (blue for Set L, orange for Set S) are plotted. Figure 6*A–C* shows the distributions

of reaction times on trials that contributed to the estimation of RTC time courses (Fig. 6*D–F*), and Figure 6*J–L* plots the mean firing rates for the two behavioral categories (darker color for hit + false alarm responses, and lighter color for miss + correct rejection responses) that were used to estimate the DP time courses (Fig. 6*G–I*). The neuronal responses on individual trials were included in the estimation of the time courses up to the point when the animal initiated its response. As a result, fewer trials contribute to later times, and the time courses shown in Figure 6 are truncated at 440 ms after stimulus onset (i.e., 200 ms after stimulus offset).

Consistent with the previous reports (Nienborg and Cumming, 2009), in all three cortical areas DP tended to increase after stimulus onset with peak seen only in V1 and VIP at ~350 ms, at which point the animal had made most of its responses. This again raises the question of whether DP in our study reflects the causal effect of the neuronal response on the perceptual decision. Nevertheless, the time courses of DP for the two stimulus sets diverged in V1 during the early period before the animal made a response (Fig. 6*G*), whereas those of MT and

VIP neurons were more or less indistinguishable between the two stimulus conditions (Fig. 6*H,I*), which mirrors the results shown in Figure 4*A* and *B*.

Unlike DP, however, RTC reached a maximum earlier (Fig. 6*D–F*), during the interval for which the neuronal responses were measured to calculate DP and RTC shown in Figure 4 (demarcated by two gray vertical lines in Fig. 6*D–L*). This can be seen most clearly in the time courses of RTC for stimulus Set S in V1 (Fig. 6*D*, orange curve) and both conditions in VIP (Fig. 6*F*). It is surprising that the two measurements of trial-to-trial correlations between the neuronal response and behavior showed different time courses with the stimulus condition given that they are significantly correlated. However, it also makes sense for the correlation between neuronal responses in a sensory cortical area and reaction time to peak before the time that the animal made responses if the neurons are causally linked to the behavior. Notably, only the responses of V1 neurons show differences in DP and RTC between the two stimulus conditions during the early period, in which the neuronal responses were less likely to be influenced by feedback.

Discussion

Previous studies have found stronger correlations between neuronal responses and behavior in later cortical areas (Leopold and Logothetis, 1996; Cook and Maunsell, 2002; Williams et al., 2003; de Lafuente and Romo, 2006; Nienborg and Cumming, 2006), suggesting that cortical readout is more tightly coupled to later stages of the hierarchical processing. However, only a few studies have measured neuronal sensitivity across multiple cortical areas using the same task for which CP or DP was compared (e.g., Nienborg and Cumming, 2006). To address this, we analyzed trial-to-trial correlation between the neuronal response and behavior and neurometric performance across cortical areas using stimuli optimized for earlier stages of the visual hierarchy. In line with earlier reports, correlations between the neuronal response and perceptual decision measured for stimulus Set L increased monotonically and were strongest in VIP, despite the superior neuronal performance in V1. However, we also found that when the stimulus was scaled such that the responses of V1 neurons carried far more reliable signals than those of MT and VIP neurons (Fig. 2), correlations between the neuronal responses and the monkey's perceptual decisions increased considerably for V1 neurons, exceeding those measured from MT neurons, and approaching those in VIP (Fig. 4).

The strongest neuronal-behavioral correlations were always in VIP, despite the fact that V1 neurons provided more reliable signals. This is consistent with many earlier reports suggesting that greater weight is given to later stages of sensory cortex when making perceptual decisions. Nevertheless, the changes we saw between the two stimulus sets suggest that this weighting is not fixed, and that decisions can be guided to a greater or lesser extent by neurons in earlier or later visual cortex depending on which areas contain the most reliable information for the task at hand. Such flexibility would be consistent with the widely supported observation that modality-specific discriminations are guided primarily by the activity of neurons in the corresponding sensory cortices, and that visual identification and motion analysis depend on activity in the ventral and dorsal pathways, respectively (Ungerleider and Mishkin, 1982; Merigan and Maunsell, 1993; Goodale et al., 1994).

Because we used drifting gratings for the change detection task, the subjects might have adopted a strategy of detecting

orientation changes rather than direction changes, or combined signals from both orientation- and direction-sensitive neurons to detect changes. In humans, sensitivity to changes in the direction of drifting gratings is comparable to sensitivity to changes in the orientation of static gratings (Heeley and Timney, 1988; Heeley and Buchanan-Smith, 1992). Interpretation of the results would be complicated if the subjects relied exclusively on orientation changes and detection of orientation changes were mediated by cortical visual areas other than MT and VIP. That possibility is unlikely for multiple reasons. Although orientation is critical for object recognition, which is predominantly mediated by the ventral visual pathway, assessment of orientation of an object in space (e.g., a grating) involves different computations. Both ablation studies (Gross, 1978) and neurophysiological records (Tsutsui et al., 2005) point to object orientation being supported by the dorsal pathway. The fact that the absolute sizes of DPs and RTCs for MT and VIP were as large as those for V1 (Fig. 4*B, D*) also supports the idea that the dorsal pathway was strongly engaged in the detection task used in our study.

When trial-to-trial correlations between neuronal responses and behavioral reports were first observed in sensory cortex (Celebrini and Newsome, 1994), they were thought to reflect a causal link between neuronal responses and behavior that arose from shared noise in afferent inputs (Zohary et al., 1994; Shadlen et al., 1996). However, subsequent studies recognized that CP (or DP) can arise either because a cell contributes to the behavioral outcome or because the activity of that cell is correlated with the activity of other neurons that contribute to the behavior (Shadlen et al., 1996; Cohen and Newsome, 2009; Nienborg et al., 2012; Goris et al., 2017). Haefner et al. (2013) showed analytically that the CP of a neuron can be little affected by the readout weight compared with effects from spiking that is correlated with other neurons in the population that contribute to the behavioral response. Some aspects of our data, such as stronger DP and RTC in later areas, might reflect differences between areas in the magnitude of their interneuronal correlations, rather than more readout weight assigned to later areas. However, several considerations make it difficult to explain the effects we found with different stimulus sets as arising changes in interneuronal correlations rather than changes in readout weight.

Compelling experimental evidence suggests that top-down signals dominate interneuronal correlation related to CP in early visual cortex (Nienborg et al., 2012; Bondy et al., 2018). One prominent candidate of such top-down signals is attention. There are three ways that attention might have affected the interneuronal correlations we measured, and hence DP and RTC. First, different levels of attention might have been directed toward stimuli in Sets L and S. Although we did not measure the attentional effort of the animals in our experiments, the transition from stimulus Set L to stimulus Set S likely resulted in a higher level of attention because the animals achieved similar behavioral thresholds with both stimulus sets. When animals pay more attention to a visual stimulus, the average pairwise correlation in the spiking of neurons responding that stimulus decreases in tasks like the one used here (Cohen and Maunsell, 2009; Mitchell et al., 2009; Zenon and Krauzlis, 2012; Herrero et al., 2013), but not all tasks (see Ruff and Cohen, 2014). This would be expected to reduce DP and RTC. However, DP and RTC of V1 neurons were stronger with stimulus Set S. Additionally, any changes in interneuronal correlations that arose from attention differences between the two stimulus sets would be expected to be larger in later cortical areas, where attention-related modulations are larger (see Maunsell and Cook, 2002). Instead, changing

stimulus sets had the biggest effects on DP and RTC in V1. Second, fluctuations in attention from trial to trial can introduce interneuronal and neuronal-behavioral correlations. Indeed, the more robust DP and RTC values seen in later stages of cortex might depend on the greater modulation of firing rate by attention that is seen in those areas (see Maunsell and Cook, 2002). It is possible that attention was more variable when animals worked with stimulus Set S, introducing the stronger DP and RTC seen with those stimuli. But again, the difference between the two stimulus sets would also be expected to be larger in later cortex areas where attention-related modulations are larger, not in V1. Third, changes in attention affect average rates of firing, and lower rates of firing can be associated with lower spike correlations (Cohen and Maunsell, 2009; Cohen and Kohn, 2011). Here, too, the changes in DP and RTC are opposite to that expected from greater attention during stimulus Set S. Overall, it is difficult to explain the changes in DP and RTC as arising from changes in interneuronal correlations related to attention.

Top-down influences other than attention might have affected interneuronal correlations in V1. For example, it has been proposed that V1 plays a role as a cognitive blackboard that higher levels modulate to mediate processes, such as segmentation or working memory (Roelfsema and de Lange, 2016). Conceivably such modulation alters interneuronal modulations in V1 differently depending on which stimuli are present. However, the absence of any measurable DP or RTC for the neuronal responses to stimuli immediately preceding stimuli to which the animal responded (see Results) argues against the idea that these correlation measures were largely driven by fluctuations such top-down factors.

Bottom-up effects that change interneuronal correlations could also affect DP and RTC without any changes in readout weights. Stimuli from Sets L and S produced different rates of firing, with stronger responses associated with Set L (Fig. 6J–L), presumably because more neurons preferred the lower spatial frequencies in that set. Once again, this does not easily explain the stronger DP and RTC seen with stimulus Set S in V1 because lower rates of firing typically produce weaker interneuronal correlations (Cohen and Maunsell, 2009; Cohen and Kohn, 2011) and the difference in firing rates in V1 was small (Fig. 6J).

Although trial-to-trial correlations between the responses of V1 neurons and behavior increased in the small stimulus condition, the neuronal-behavioral correlation was largest in VIP in both stimulus conditions. It might be that greater weight is always given to activity in later stages of visual cortex because the more abstract representations found in later stages are somehow critically important for mediating perceptual decisions. Alternatively, it might indicate that area VIP is not a purely sensory area, but instead, like LIP (Gold and Shadlen, 2007), receives abundant top-down decision-related inputs. Decision-related signals are expected to be strongly correlated with behavioral responses (Goris et al., 2017). In line with this possibility, studies investigating the activity of VIP neurons in relation to heading discrimination suggest that sensory information present in the response of VIP neurons is not decoded in generating perceptual choices (Pitkow et al., 2015), but rather the response of VIP neurons reflect top-down signals reflecting the perceptual choice (Zaidel et al., 2017). Consistent with this, inactivation of neurons in VIP or neighboring LIP does not impair performance in a visual heading discrimination task (Chen et al., 2016), or sensitivity to direction of motion in a direction discrimination task using dynamic random dots (Katz et al., 2016). While we think it is less likely, it is still possible that the increases we saw in DP and RTC in V1 with Set S were unrelated to how much those

neurons contributed to the perceptual reports, but instead were a sign of stronger top-down inputs to V1 that reflected the perceptual choice when that stimulus set was used.

References

- Avillac M, Ben Hamed S, Duhamel J (2007) Multisensory integration in the ventral intraparietal area of the macaque monkey. *J Neurosci* 27:1922–1932.
- Bishop CM (2006) Pattern recognition and machine learning: information. In: Science and statistics. New York: Springer.
- Bondy AG, Haefner RM, Cumming BG (2018) Feedback determines the structure of correlated variability in primary visual cortex. *Nat Neurosci* 21:598–606.
- Bosking WH, Maunsell JHR (2011) Effects of stimulus direction on the correlation between behavior and single units in area MT during a motion detection task. *J Neurosci* 31:8230–8238.
- Boudreau C, Williford T, Maunsell JHR (2006) Effects of task difficulty and target likelihood in area V4 of macaque monkeys. *J Neurophysiol* 96:2377–2387.
- Bradley A, Skottun BC, Ohzawa I, Sclar G, Freeman RD (1987) Visual orientation and spatial frequency discrimination: a comparison of single neurons and behavior. *J Neurophysiol* 57:755–772.
- 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.
- Carnevale F, de Lafuente V, Romo R, Parga N (2013) An optimal decision population code that accounts for correlated variability unambiguously predicts a subject's choice. *Neuron* 80:1532–1543.
- Celebrini S, Newsome WT (1994) Neuronal and psychophysiology sensitivity to motion signals in extrastriate area MST of the macaque monkey. *J Neurosci* 14:4109–4124.
- Chen A, DeAngelis GC, Angelaki DE (2013) Functional specializations of the ventral intraparietal area for multisensory heading discrimination. *J Neurosci* 33:3567–3581.
- 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, Maunsell JHR (2009) Attention improves performance primarily by reducing interneuronal correlations. *Nat Neurosci* 12:1594–1600.
- 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 JHR (2002) Dynamics of neuronal responses in macaque MT and VIP during motion detection. *Nat Neurosci* 5:985–994.
- Crist CF, Yamasaki DS, Komatsu H, Wurtz RH (1988) A grid system and a microsyringe for single cell recording. *J Neurosci Methods* 26:117–122.
- de Lafuente V, Romo R (2006) Neural correlate of subjective sensory experience gradually builds up across cortical areas. *Proc Natl Acad Sci USA* 103:14266–14271.
- De Valois RL, Yund EW, Hepler N (1982) The orientation and direction selectivity of cells in macaque visual cortex. *Vision Res* 22:531–544.
- 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.
- Felleman DJ, Van Essen DC (1991) Distributed hierarchical processing in the primate cerebral cortex. *Cereb Cortex* 1:1–47.
- Geisler WS, Albrecht DG (1997) Visual cortex neurons in monkeys and cats: detection, discrimination and identification. *Vis Neurosci* 14:897–919.
- Gold JI, Shadlen MN (2007) The neural basis of decision making. *Annu Rev Neurosci* 30:535–574.
- Goodale MA, Meenan JP, Bühlhoff HH, Nicolle DA, Murphy KJ, Racicot CI (1994) Separate neural pathways for the visual analysis of object shape in perception and prehension. *Curr Biol* 4:604–610.
- Goris RL, 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.
- Green DM, Swets JA (1966) Signal detection theory and psychophysics. New York: Wiley.

- Gross CG (1978) Inferior temporal lesions do not impair discrimination of rotated patterns in monkeys. *J Comp Physiol Psychol* 92:1095–1109.
- Gu Y, DeAngelis G, Angelaki D (2007) A functional link between area MSTd and heading perception based on vestibular signals. *Nat Neurosci* 10:1038–1047.
- 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.
- Heeley DW, Timney B (1988) Meridional anisotropies of orientation discrimination for sine wave gratings. *Vision Res* 28:337–344.
- Heeley DW, Buchanan-Smith HM (1992) Directional acuity for drifting plaids. *Vision Res* 32:97–104.
- Herrero JL, Gieselmann MA, Sanayei M, Thiele A (2013) Attention-induced variance and noise correlation reduction in macaque V1 is mediated by NMDA receptors. *Neuron* 78:729–739.
- Herrington T, Masse N, Hachmeh K, Smith J, Assad J, Cook E (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.
- Histed MH, Ni AM, Maunsell JHR (2013) Insights into cortical mechanisms of behavior from microstimulation experiments. *Prog Neurobiol* 103:115–130.
- Hong H, Yamins DL, Majaj NJ, DiCarlo JJ (2016) Explicit information for category-orthogonal object properties increases along the ventral stream. *Nat Neurosci* 19:613–622.
- Judge SJ, Richmond BJ, Chu FC (1980) Implantation of magnetic search coils for measurement of eye position: an improved method. *Vision Res* 20:535–538.
- Kang I, Malpeli JG (2009) Dim-light sensitivity of cells in the awake cat's lateral geniculate and medial interlaminar nuclei: a correlation with behavior. *J Neurophysiol* 102:841–852.
- Kang I, Maunsell JHR (2012) Potential confounds in estimating trial-to-trial correlations between neuronal response and behavior using choice probabilities. *J Neurophysiol* 108:3403–3415.
- 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.
- Law C-T, Gold JI (2008) Neural correlates of perceptual learning in a sensory-motor, but not a sensory, cortical area. *Nat Neurosci* 11:505–513.
- Leopold DA, Logothetis NK (1996) Activity changes in early visual cortex reflect monkeys' percepts during binocular rivalry. *Nature* 379:549–553.
- 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.
- Masse N, Cook E (2008) The effect of middle temporal spike phase on sensory encoding and correlates with behavior during a motion-detection task. *J Neurosci* 28:1343–1355.
- Maunsell JHR, Cook EP (2002) The role of attention in visual processing. *Philos Trans R Soc Lond B Biol Sci* 357:1063–1072.
- Merigan WH, Maunsell JHR (1993) How parallel are the primate visual pathways? *Annu Rev Neurosci* 16:369–402.
- Mitchell JF, Sundberg KA, Reynolds JH (2009) Spatial attention decorrelates intrinsic activity fluctuations in macaque area V4. *Neuron* 63:879–888.
- Moreno-Bote R, Beck J, Kanitscheider I, Pitkow X, Latham P, Pouget A (2014) Information-limiting correlations. *Nat Neurosci* 17:1410–1417.
- Mountcastle VB, LaMotte RH, Carli G (1972) Detection thresholds for stimuli in humans and monkeys: comparison with threshold events in mechanoreceptive afferent nerve fibers innervating the monkey hand. *J Neurophysiol* 35:122–136.
- Movshon JA, Adelson EH, Gizzi MS, Newsome WT (1985) The analysis of moving visual patterns. *Exp Brain Res* 11:117–151.
- Ni AM, Ruff DA, Alberts JJ, Symmonds J, Cohen MR (2018) Learning and attention reveal a general relationship between population activity and behavior. *Science* 359:463–465.
- 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.
- Osborne LC, Lisberger SG, Bialek W (2005) A sensory source for motor variation. *Nature* 437:412–416.
- 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, Hawken MJ (1985) Capabilities of monkey cortical cells in spatial-resolution tasks. *J Opt Soc Am A* 2:1101–1114.
- Parker AJ, Newsome WT (1998) Sense and the single neuron: probing the physiology of perception. *Annu Rev Neurosci* 21:227–277.
- Penfield W (1947) Some observations on the cerebral cortex of man. *Proc R Soc Lond B Biol Sci* 134:329–347.
- 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.
- Priebe NJ, Cassanello CR, Lisberger SG (2003) The neural representation of speed in macaque area MT/V5. *J Neurosci* 23:5650–5661.
- Ringach D, Shapley R, Hawken M (2002) Orientation selectivity in macaque V1: diversity and laminar dependence. *J Neurosci* 22:5639–5651.
- Robinson DA (1963) A method of measuring eye movements using a scleral search coil in a magnetic field. *IEEE Trans Biomed Eng* 101:131–145.
- Roelfsema PR, de Lange FP (2016) Early visual cortex as a multiscale cognitive blackboard. *Annu Rev Vis Sci* 2:131–151.
- Rovamo J, Virsu V, Nasanen R (1978) Cortical magnification factor predicts the photopic contrast sensitivity of peripheral vision. *Nature* 271:54–56.
- Ruff DA, Cohen MR (2014) Attention can either increase or decrease spike count correlations in visual cortex. *Nat Neurosci* 17:1591–1597.
- 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.
- Tolhurst DJ, Movshon JA, Dean AF (1983) The statistical reliability of signals in single neurons in cat and monkey visual cortex. *Vision Res* 23:775–785.
- Tsutsui K, Taira M, Sakata H (2005) Neural mechanisms of three-dimensional vision. *Neurosci Res* 51:221–229.
- Uka T, DeAngelis G (2004) Contribution of area MT to stereoscopic depth perception: choice-related response modulations reflect task strategy. *Neuron* 42:297–310.
- Ungerleider LG, Mishkin M (1982) Two cortical visual systems. In: *Analysis of visual behavior* (Ingle DJ, Goodale MA, Mansfield RJ, eds), pp 549–585. Cambridge, MA: Massachusetts Institute of Technology.
- Vogels R, Orban GA (1990) How well do response changes of striate neurons signal differences in orientation: a study of the discriminating monkey. *J Neurosci* 10:3543–3558.
- 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.
- Wolbarsht MI, MacNichol EF, Wagner HG (1960) Glass insulated platinum microelectrode. *Science* 132:1309–1310.
- Zaidel A, DeAngelis GC, Angelaki DE (2017) Decoupled choice-driven and stimulus-related activity in parietal neurons may be misrepresented by choice probabilities. *Nat Commun* 8:715.
- Zenon A, Krauzlis RJ (2012) Attention deficits without cortical neuronal deficits. *Nature* 489:434–437.
- Zoccolan D, Kouh M, Poggio T, DiCarlo JJ (2007) Trade-off between object selectivity and tolerance in monkey inferotemporal cortex. *J Neurosci* 27:12292–12307.
- Zohary E, Celebriani S, Britten KH, Newsome WT (1994) Neuronal plasticity that underlies improvement in perceptual performance. *Science* 263:1289–1292.