Neuroscience

Ocular dominance-dependent binocular combination of monocular neuronal responses in macaque V1

Sheng-Hui Zhang, Xing-Nan Zhao, Shi-Ming Tang, Cong Yu

PKU-Tsinghua Center for Life Sciences • School of Life Sciences • IDG-McGovern Institute for Brain Research • School of Psychological and Cognitive Sciences, Peking University, Beijing, China

Reviewed Preprint

Published from the original preprint after peer review and assessment by eLife.

About eLife's process

Reviewed preprint version 1
February 2, 2024 (this version)

Posted to preprint server
October 27, 2023

Sent for peer review
October 25, 2023

Abstract

Primates rely on two eyes to see depth, while keeping a stable vision when one eye is closed. Although psychophysical and modeling studies have investigated how monocular signals are combined to form binocular vision, the corresponding neuronal mechanisms, especially in V1 where most neurons become binocular but with different eye preferences, are not well understood. Here we used two-photon calcium imaging to compare monocular and binocular responses of thousands of V1 superficial-layer neurons in three awake macaques. Under monocular stimulation, neurons preferring the stimulated eye responded substantially stronger than those preferring both eyes. However, under binocular stimulation, the responses of neurons preferring either eye were suppressed, and those preferring both eyes were enhanced, so that neuronal responses became similar regardless of eye preferences. A neuronally realistic model of binocular combination, which includes ocular dominance-dependent divisive interocular inhibition, and binocular summation, is proposed to account for these observations.

eLife assessment

Overall, the reviewers found the significance of the work **valuable** to the field of visual neuroscience, particularly given the large data set and strength of the method used that allowed for spatial analysis of neuronal responses in macaque V1. The evidence was deemed **compelling**, owing in part to the consistency of responses across animals and the fitness of modeling. Ways to improve the manuscript as outlined include an expanded discussion of similar prior literature and limitations of the method used to read out neuronal responses.

Introduction

Human and non-human primates often use binocular disparity, or the differences between the retinal images of two eyes, to perceive depth (stereopsis). In the meantime, the brain maintains a stable perception of the visual world when one eye closes and the incoming light is halved. Much has been known regarding the neural foundations of stereopsis (Barlow, Blakemore, & Pettigrew,

1967; Henriksen, Tanabe, & Cumming, 2016; Parker, Smith, & Krug, 2016; Welchman, 2016; Read, 2021). Nevertheless, whether and how the neurons respond differently to monocular and binocular stimulations is less studied (Prince, Pointon, Cumming, & Parker, 2002; Dougherty, Cox, Westerberg, & Maier, 2019; Mitchell et al., 2022). Adding to the complexity is the fact that many V1 neurons, although responding to stimulations in both eyes, have various degrees of eye preferences (Hubel & Wiesel, 1962, 1968). Thus, a more complete picture of binocular combination of monocular responses would illustrate monocular and binocular responses of neurons at different eye preferences, which is the purpose of the current study.

It has been reported that overall the binocular responses of macaque V1 neurons are lower than the sum of monocular responses to each eye (Prince et al., 2002; Mitchell et al., 2022), but are stronger than monocular responses either when one eye was stimulated (Prince et al., 2002; Mitchell et al., 2022), or when neurons' preferred eye was stimulated (Mitchell et al., 2022). Further, Dougherty et al. (2019) reported that the responses of monocular neurons are found to be more likely suppressed than facilitated by binocular stimulation. As our results will suggest later, this conclusion is valid when the monocular baseline is either the sum of monocular responses, or the monocular responses of either eye. Dougherty et al. (2019) also reported similar responses of neurons with monocular and binocular preferences, as well as similar responses of binocular neurons to monocular and binocular stimulations, which are not supported by our data.

A big and more complete picture of the binocular combination of monocular responses, as well as the impacts of eye preferences, would come from large samples of neurons, so that the data can be quantitatively described with sufficient statistical power, and analyzed with computational modeling. Two-photon calcium imaging is capable of recording thousands of neurons simultaneously at the single-neuron resolution, which is especially suitable for this task. In this study, we used a two-photon imaging setup, which had been adapted for recording with awake macaques (Li, Liu, Jiang, Lee, & Tang, 2017), to measure responses of V1 superficial-layer neurons to monocular and binocular stimulations. Furthermore, a neuronally realistic binocular combination model, which considers ocular dominance-dependent interocular divisive inhibition, as well as binocular summation, is proposed to explain the data.

Results

We recorded responses of V1 superficial-layer neurons to monocular (contralateral and ipsilateral) and binocular stimulations in three awake, fixating macaques. The stimulus was a high-contrast (0.9) Gabor grating at various orientations and spatial frequencies (SFs). The recording was performed at two cortical depths of the same response field of view (FOV) in first two monkeys (MA & MB), and at one depth in the third monkey (MC) as the first two monkeys displayed similar results at two depths (**Fig. 1A**). Imaging processing identified a total of 10,168 neurons. Among them, 9,390 (92.4%) were tuned to orientation, SF, or both. Results from these orientation- and/or SF-tuned neurons were used in the following data analysis.

V1 superficial-layer neurons exhibited various degrees of eye preferences, consistent with the classical findings of Hubel and Wiesel (1962, 1968). An ocular dominance index (ODI) was calculated to characterize each neuron's eye preference: ODI = $(R_i - R_c)/(R_i + R_c)$, in which R_i and R_c were the neuron's respective peak responses to ipsilateral and contralateral stimulations on the basis of data fitting (see Methods). Here ODI = −1 and 1 would indicate complete contralateral and ipsilateral eye preferences, respectively, and ODI = 0 would indicate equal preferences to both eyes. The ocular dominance functional maps at single-neuron resolution, especially those of Monkeys A and C, showed regions of neurons preferring either the contralateral (blue) or the ipsilateral eye (red), and transitional zones where neurons preferring both eyes (white) (**Fig. 1B**). The ocular dominance maps were similar at two cortical depths in Monkeys A and B,

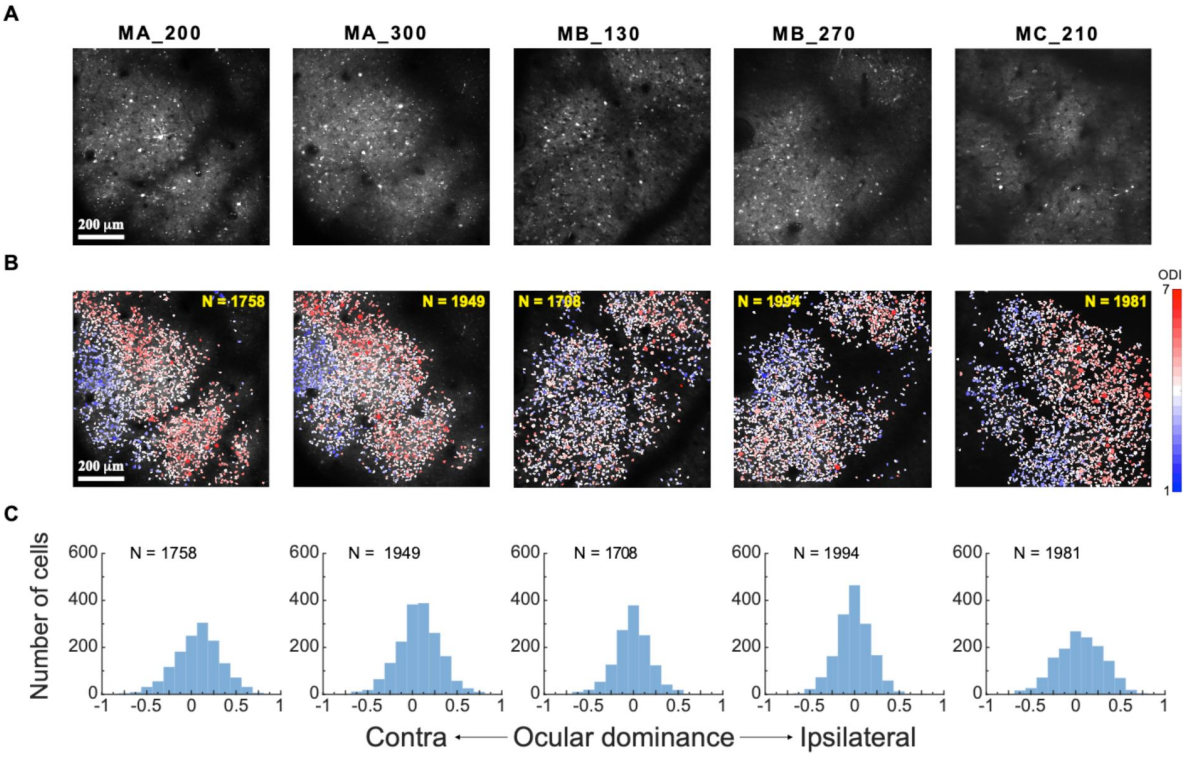


Figure 1.

Eye preferences of V1 superficial-layer neurons in three macaques.

A. Two-photon imaging. Average 2-photon images over a recording session for each response FOV. MA_200: Monkey A at a 200-μm cortical depth. **B**. Ocular dominance functional maps of each FOV/depth at single-neuron resolution. **C**. Frequency distributions of neurons of each FOV/depth as a function of ocular dominance index.

indicating possible ocular dominance columns (**Fig. 1B**). The frequency distributions of ODIs suggest more binocular neurons than monocular neurons in V1 superficial layers (**Fig. 1C**), similar to the normal distribution of ocularity index in Dougherty et al. (2019).

The responses of individual neurons were plotted against the ocular dominance index (ODI) when monocular stimulation was through the contralateral eye (**Fig. 2A**) or the ipsilateral eye (**Fig. 2B**). As expected, neurons with more negative ODIs responded stronger when the contralateral eye was stimulated, and those with more positive ODIs responded stronger when the ipsilateral eye was stimulated. The responses declined as neurons became more binocular. However, when both eyes were stimulated, the response differences among neurons of different eye preferences became not obvious (**Fig. 2C**).

These trends can be better appreciated when the difference between binocular and monocular responses are presented for each FOV/depth in **Fig. 2D**: Under monocular stimulation, if only the neuronal responses to the preferred eye were taken into consideration (i.e., a neuron's higher response to ipsilateral vs. contralateral stimulations), more monocular neurons (ODIs farther from 0) tended to have stronger responses than more binocular neurons (ODIs closer to 0). However, under binocular stimulation, the responses of more monocular neurons were suppressed, and those of more binocular neurons were enhanced, by binocular stimulation. These trends are best appreciated in pooled data over five FOVs/depths (**Fig. 2E**). Further, linear regression confirmed a significant dependence of binocular modulation on absolute ODI ($y = -0.41x + 0.16$, $p < 0.001$) (**Fig. 2E**). That is, responses of neurons with lower absolute ODI (i.e., binocular neurons) tended to be more enhanced, and responses of neurons with higher absolute ODI (i.e., monocular neurons) tended to be more suppressed.

Modeling monocular and binocular responses

To model the monocular and binocular neural responses, monocular and binocular data of each FOV/depth, as well as the pooled data, were first normalized by the respective median of the binocular responses and then divided into 60 bins in the order of the ocular dominance index (**Fig. 3A-C**).

Monocular responses

A neuron's monocular responses to contralateral and ipsilateral stimulations could be respectively described by a divisive gain control model:

$$R_c = \frac{w_c^{m_c} \cdot S_c}{k} \quad \text{and} \quad R_i = \frac{w_i^{m_i} \cdot S_i}{k} \qquad \text{(Eq. 1)}$$

Here S_c and S_i were stimulus contrasts, w_c and w_i were linear transformations of a neuron's ocular dominance index from [-1 1] to [0 1]: $w_c = (ODI + 1)/2$ and $w_i = 1 - w_c$, m_c and m_i represented monocular nonlinearity, and k represented divisive gain control. Because S_c and S_i in our experiments were 0.9, which were about equal to 1 (full contrast) due to neuronal response saturation, the above equations were simplified as

$$R_c = \frac{w_c^{m_c}}{k} \quad \text{and} \quad R_i = \frac{w_i^{m_i}}{k} \qquad \text{(Eq. 2)}$$

Eq. 2 was used to fit the binned median contralateral and ipsilateral data (**Figs. 3A** & B), with the parameter k being equal for contralateral and ipsilateral responses. The fitting revealed that m_c and m_i were negative and close to −1, which resulted in a quick decline of $w_c^{m_c}$ and a quick increase of $w_i^{m_i}$ as a function of the ocular dominance index since w_c and $w_i \in [0, 1]$. The fit quality indices (Busse, Wade, & Carandini, 2009) ranged 0.85-0.95 for the contralateral condition (**Fig. 3A**) and 0.90-0.93 for the ipsilateral condition (**Fig. 3B**), suggesting adequate goodness of fit. The fitting parameters are listed in **Fig. 3D**.

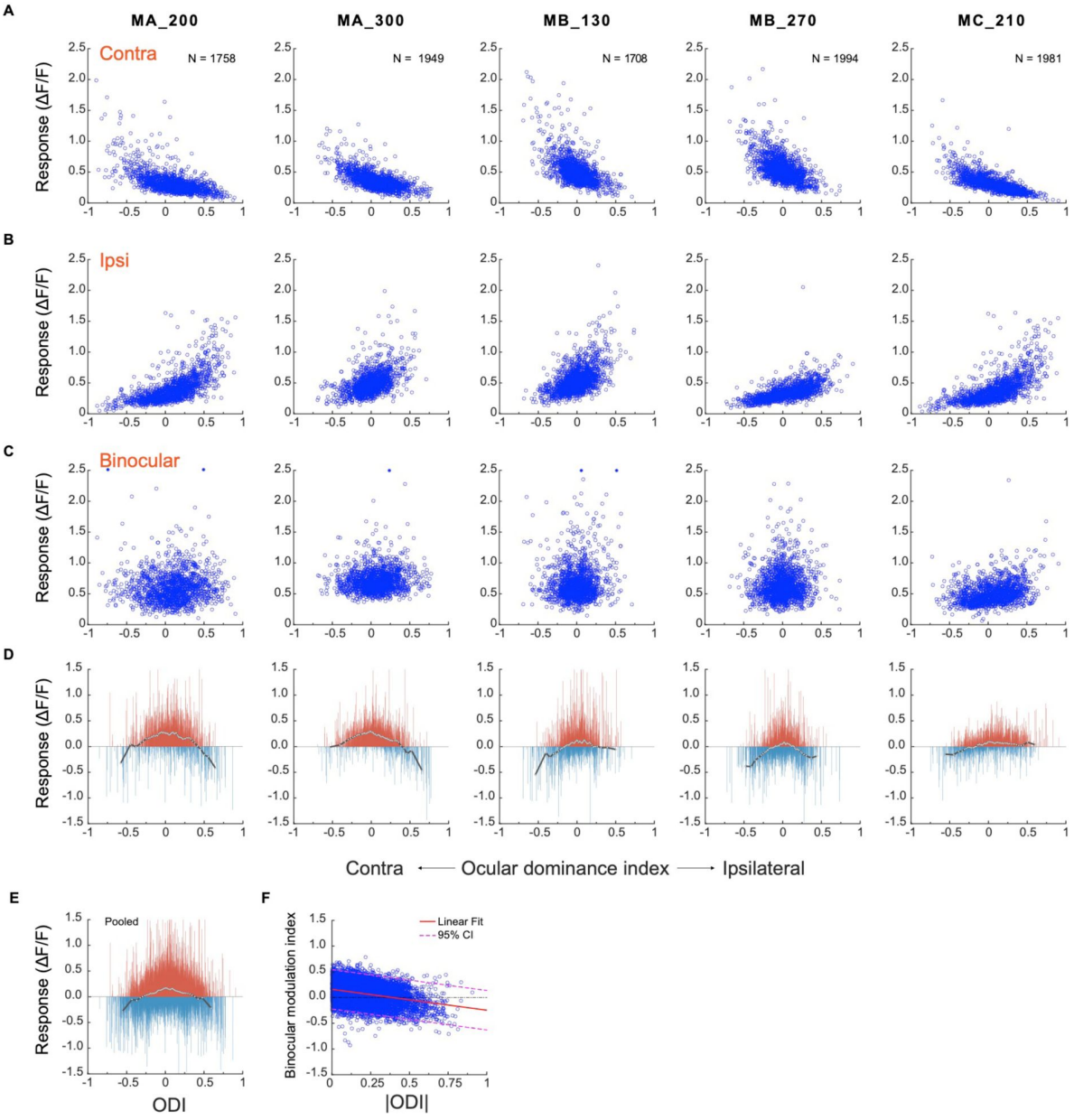
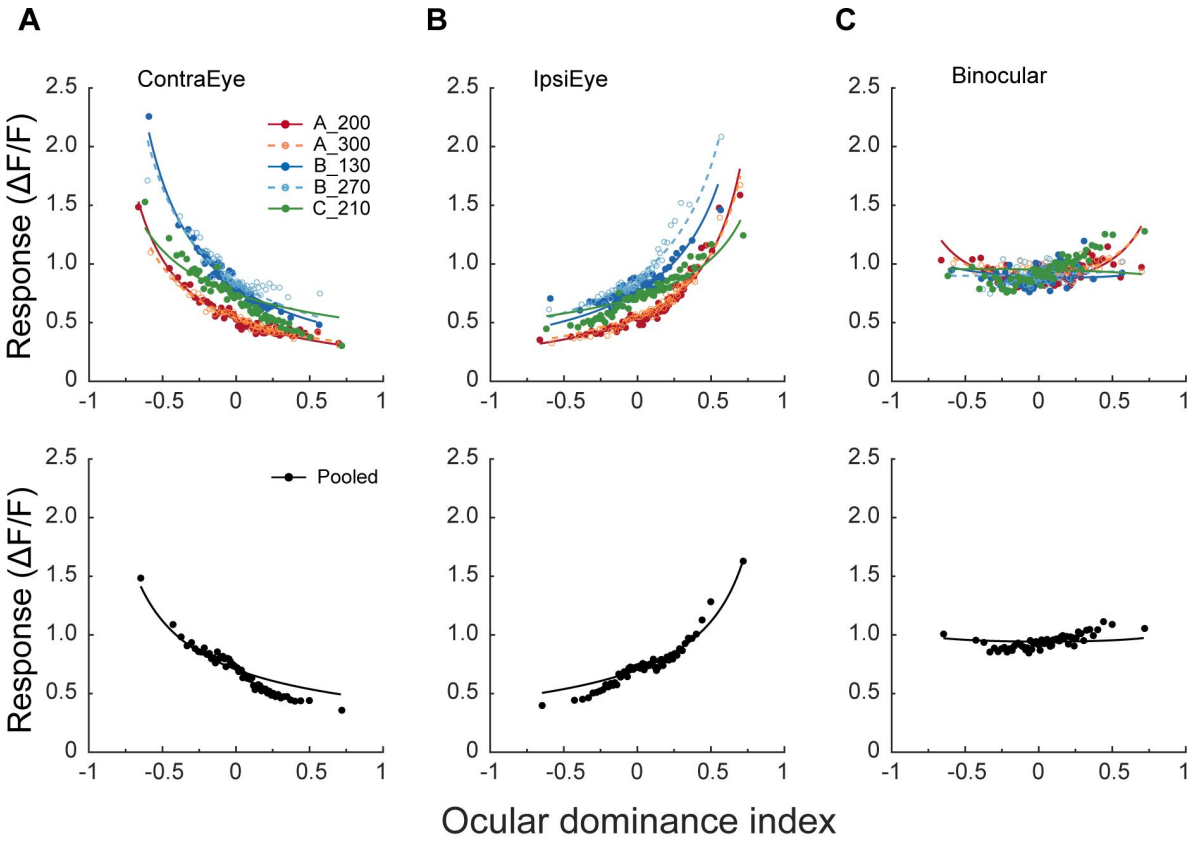


Figure 2.

A comparison of neuronal responses to monocular and binocular stimulations.

A. Responses of individual neurons against their ocular dominance indices with contralateral stimulation. **B**. Responses of the same neurons against their ocular dominance indices with ipsilateral stimulation. **C**. The binocular responses of the same neurons. **D**. The difference between binocular and monocular responses (R_b – max(R_i, R_c)). Each vertical line represents one neuron. To summarize the results, neurons of each FOV/depth are evenly divided into 60 bins in the order of the ocular dominance index. White dots represent the median responses of respective bins and are connected with a black line. **E**. The differences between binocular and monocular responses of individual neurons pooled over five FOVs/depths. **F**. Binocular modulation index as a function of absolute ODI and the linear fit. The binocular modulation index of each neuron was defined as (R_b – max (R_i, R_c))/(R_b + max (R_i, R_c)).



D

	m	k		m	k		b
A_200	-0.981	3.717		-1.008	3.717		0.302
A_300	-0.868	3.380		-0.936	3.380		0.302
B_140	-1.074	2.582		-0.984	2.582		0.825
B_250	-0.967	2.299		-1.036	2.299		0.935
C_210	-0.585	2.005		-0.532	2.005		0.622
Pooled	-0.672	2.270		-0.665	2.270		0.578

Figure 3.

Modeling monocular and binocular responses.

A & B. Median neuronal responses to contralateral and ipsilateral stimulations as a function of ocular dominance index and respective data fitting with **Eq. 2**. Neurons are evenly divided into 60 bins in the order of the ocular dominance index, with each bin containing 28-33 neurons that varied among different FOVs/depths (156 neurons for the pooled data). Each datum represents the median response of a bin. Free parameter k was kept equal during contralateral and ipsilateral data fitting. **C**. Binocular responses as a function of ocular dominance index and data fitting with **Eq. 3** for the same bins of neurons. During binocular data fitting, parameters k, m_i, and m_c were inherited from monocular data fitting, and only b was the free-changing parameter. **D**. The values of free parameters from monocular and binocular data fitting.

Binocular responses

Fig. 2D-F earlier has indicated that neuronal responses to binocular stimulation change from suppression to enhancement as neurons' ocular dominance changes from monocular to binocular, which may reflect the ocular dominance-dependent net effect of interocular suppression and binocular summation. To model interocular response suppression, responses from each eye in **Eq. 2** were further normalized by an interocular suppression factor w_i^b or w_c^b, whose strength was decided by the linearly transformed ODI with a nonlinear exponent b. Then the normalized responses from two eyes were summed to model binocular summation and the final binocular responses of neurons R_b (**Eq. 3**).

$$R_b = \frac{w_i^{m_i}}{k w_i^b} + \frac{w_c^{m_c}}{k w_c^b} \qquad \text{(Eq. 3)}$$

During data fitting, the parameters m_i, m_c, and k were inherited from earlier monocular data fitting and remained fixed. Only b was allowed to change. Data fitting resulted in flat binocular response functions (**Fig. 3C**) with satisfactory goodness-of-fit (fit quality index = 0.86 to 0.94).

The behaviors of interocular suppression and binocular summation in the model, as well as their contributions to the binocular response, may be better appreciated in **Fig. 4**. **Fig. 4A** uses **Eqs. 2** and 3 and the parameters from fitting of pooled data (**Fig. 3D**) to simulate the contralateral (blue curve with label R_c), ipsilateral (red curve with label R_i), and binocular (black curve) response functions against the ocular dominance index. The arithmetic sum of contralateral and ipsilateral response functions was also simulated (grey dashed curve labeled $R_i + R_c$). In addition, neuronal responses to preferred eye stimulation would consist of the higher branches of contralateral and ipsilateral response functions. It is apparent that binocular responses can be described by neither the sum of monocular responses, nor the responses to the preferred eye. Instead, the median of the binocular response function (black arrow) in each data set is close to but still more or less higher than the median of the contralateral (blue arrow) or ipsilateral response function (red arrow), which is consistent with previous reports (Prince et al., 2002; Mitchell et al., 2022).

Fig. 4B plots the interocular suppression factor w_c^b (dotted blue curve) for the contralateral response R_c (solid blue curve from 4A), and w_i^b (dotted red curve) for the ipsilateral response R_i (solid red curve from 4A). The suppression factors w_c^b and w_i^b are larger with neurons that are more monocular than with neurons that are more binocular. The R_c and R_i are divided by the respective interocular suppression factor, producing the normalized contralateral responses (dashed blue curve labeled R_c/w_c^b) and ipsilateral responses (dashed red curve labeled R_i/w_i^b). These normalized curves are lower than original monocular responses, especially for neurons that are more monocular (ODIs farther from 0), showing interocular suppression.

Then the normalized monocular curves are summed up in **Fig. 4C**, which represents binocular summation and produces the final binocular responses (R_b). Therefore, as a result of combined interocular suppression and binocular summation, the final curve for binocular responses becomes nearly flat within the data range.

Discussion

The current study compared the responses of large samples of V1 superficial-layer neurons to monocular and binocular stimulations in three macaques. The monocular response functions show steep changes as a function of the ocular dominance index, but binocular response functions are largely flat, regardless of neurons' eye preferences. Modeling efforts indicated that when signals from two eyes are combined, interocular divisive suppression, which is more prominent

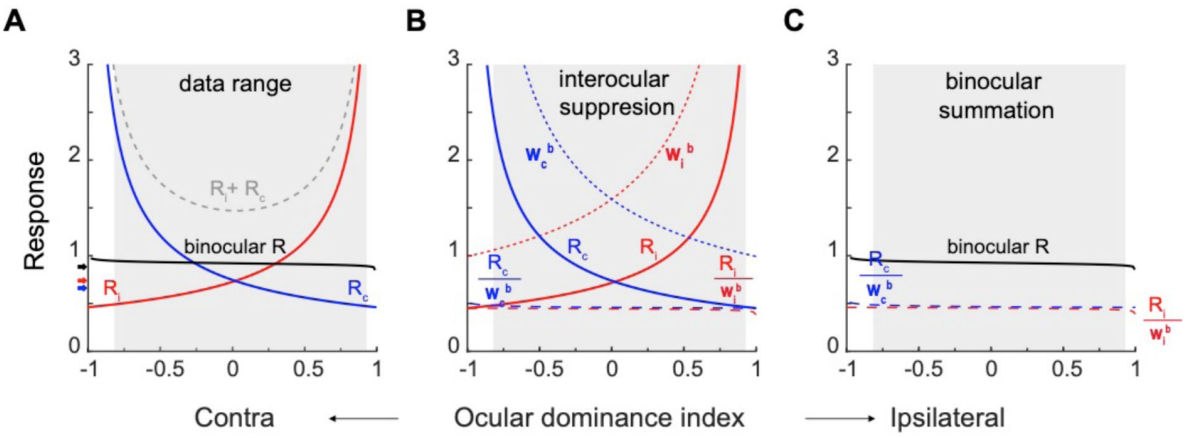


Figure 4.

Monocular and binocular responses modeled with binocular suppression and binocular summation.

A. The solid blue, red, and black curves are simulations of the contralateral, ipsilateral, and binocular responses on the basis of fitting of pooled data (**Fig. 3D**). The grey dashed curve simulates binocular responses as the arithmetic sum of contralateral and ipsilateral responses. The higher branches of contralateral and ipsilateral response functions represent monocular responses with preferred eye stimulation. The black, blue, and red arrows indicate the median binocular, contralateral, and ipsilateral responses, respectively, from pooled data. The shadowed area indicates the region where actual neurons existed on the basis of the ocular dominance index. **B.** Interocular suppression. The contralateral and ipsilateral responses (Rc & Ri) are divided by respective interocular suppression factors Wc^b and W^b to produce interocular-suppressed responses R_c/w_c^b and R_i/w_i^b. **C.** Binocular summation. R_c/w_c^b and R_i/w_i^b are summed to produce the final binocular responses Rb.

with neurons preferring one eye, and binocular summation, which is more prominent with binocular neurons, together produce the nearly flat binocular response function within a broad range of ocular dominance indices.

As shown in **Fig. 4A**, the median of binocular neuronal responses is slightly higher than that of monocular responses of either eye (black vs. red/blue arrows), consistent with earlier reports (Prince et al., 2002; Mitchell et al., 2022). However, individually, compared to monocular responses, responses of monocular neurons more preferring the stimulated eye are actually suppressed, and only responses of binocular neurons are increased by binocular stimulation. Furthermore, overall stronger binocular than monocular responses in V1, even with a small amount, would suggest that the binocular neuronal responses must be further modulated by additional mechanisms to achieve stable vision, regardless of whether the world is viewed with one or two eyes.

In Dougherty et al. (2019), the responses of monocular and binocular neurons are similar under monocular stimulation, and only monocular neurons, but not binocular neurons, are significantly suppressed by binocular stimulation. In contrast, our results show that monocular neurons respond to stimulation in the preferred eye much more strongly than binocular neurons do to the same monocular stimulation (**Fig. 2C**). Moreover, although confirming interocular suppression of monocular neurons responding to the preferred eye, our results also indicate enhanced responses of binocular neurons to binocular stimulation. Considering the large diversity of binocular responses of neurons with similar eye preferences (**Fig. 2C**), it is very possible to obtain inaccurate statistical estimates with small, and possibly biased, samples of neurons in electrode recording experiments.

Binocular combination of monocular signals has long been recognized to involve both interocular suppression and binocular summation (DeSilva & Bartley, 1930; Cohn & Lasley, 1976; Li & Atick, 1994). Most more recent models use divisive normalization to explain interocular suppression (e.g., Cogan, 1987; Anderson & Movshon, 1989; Ding & Sperling, 2006; Moradi & Heeger, 2009; Huang, Zhou, Zhou, & Lu, 2010; Ding, Klein, & Levi, 2013). Our modeling effort shows that a similar divisive interocular suppression and binocular summation model can account for neuronal response changes under monocular and binocular stimulations, with the distinction that the divisive interocular suppression is additionally controlled by a neuron's ocular dominance. The latter also controls the responses of individual neurons to monocular stimulation, as shown in **Figs. 3A-B**. The critical roles of ocular dominance have been largely overlooked by extant binocular vision models to our knowledge, except that Anderson and Movshon (1989) demonstrated that a model consisting of multiple ocular dominance channels can better explain their psychophysical adaptation data. We hope that our two-photon imaging results can be incorporated into future neuronally plausible models of binocular vision.

On the basis of current findings, future two-photon imaging work shall compare neural responses to monocular and binocular stimulations with uneven effective stimulus contrasts due to physical contrast differences (Anderson & Movshon, 1989), monocular adaptation (Anderson & Movshon, 1989), and short-term monocular deprivation (Lunghi, Burr, & Morrone, 2011), as well as the relevant roles of ocular dominance of individual neurons. These results would help understand abnormal binocular vision in patients with strabismus and amblyopia. In addition, under current zero-disparity stimulus conditions, neuronal responses to binocular stimulation appear no longer affected by ocular dominance of individual neurons (**Fig. 2D**). Therefore, another interesting question is whether the same is true when disparity is present in the stimuli. These future results would enrich and refine our current simplified but neuronally realistic model.

Materials and Methods

Monkey preparation

Monkey preparations were identical to those reported in a previous study (Guan, Ju, Tao, Tang, & Yu, 2021; Ju, Guan, Tao, Tang, & Yu, 2021). Three rhesus monkeys (*Macaca mulatta*) aged 4 to 6 years, respectively, were each prepared with two sequential surgeries under general anesthesia and strictly sterile condition. In the first surgery, a 20-mm diameter craniotomy was performed on the skull over V1. The dura was opened and multiple tracks of 100-150 nL AAV1.hSynap.GCaMP5G.WPRE.SV40 (AV-1-PV2478, titer 2.37e13 (GC/ml), Penn Vector Core) were pressure-injected at a depth of ~350 μm. Then the dura was sutured, the skull cap was re-attached with three titanium lugs and six screws, and the scalp was sewn up. After the surgery, the animal was returned to the cage, treated with injectable antibiotics (Ceftriaxone sodium, Youcare Pharmaceutical Group, China) for one week. Postop analgesia was also administered. The second surgery was performed 45 days later. A T-shaped steel frame was installed for head stabilization, and an optical window was inserted onto the cortical surface. Data collection could start as early as one week later. More details of the preparation and surgical procedures can be found in Li et al. (2017). The procedures were approved by the Institutional Animal Care and Use Committee, Peking University.

Behavioral task

After a ten-day recovery from the second surgery, monkeys were seated in primate chairs with head restraint. They were trained to hold fixation on a small white spot (0.1°) with eye positions monitored by an ISCAN ETL-200 infrared eye-tracking system (ISCAN Inc.) at a 120-Hz sampling rate (Monkeys A & B) or an Eyelink-1000 (SR Research) at a 1000-Hz sampling rate (Monkey C). During the experiment, trials with the eye position deviated 1.5° or more from the fixation before stimulus offset were discarded as ones with saccades and repeated. For the remaining trials, the eye positions were mostly concentrated around the fixation point, with eye positions in over 95% of trials within 0.5° from the fixation point. The viewing was binocular.

Visual stimuli

For Monkeys A and B, visual stimuli were generated by the ViSaGe system (Cambridge Research Systems) and presented on a 21'' Sony G520 CRT monitor (refresh rate = 80 Hz, resolution = 1280 pixel × 960 pixel, pixel size = 0.31 mm × 0.31 mm). Because of the space limit, the viewing distance and the monitor position varied depending on the stimulus spatial frequency (30 cm at 0.25, 0.5, and 1 cpd, 60 cm at 2 cpd, and 120 cm at 4 and 8 cpd). For Monkey C, visual stimuli were generated by Psychotoolbox 3 (Pelli & Zhang, 1991) and presented on a 27'' Acer XB271HU LCD monitor (refresh rate = 80 Hz native, resolution = 2560 pixel × 1440 pixel native, pixel size = 0.23 mm × 0.23 mm). The viewing distance was 50 cm for lower frequencies (0.25 – 1 cpd) and 100 cm for higher frequencies (2 - 8 cpd). For both monitors, the screen luminance was linearized by an 8-bit look-up table, and the mean luminance was ~47 cd/m^2.

A drifting square-wave grating (SF = 4 cpd, contrast = full, speed = 3 cycles/s, starting phase = 0°, and size = 0.4° in diameter) was first used to determine the location, eccentricity (3.4° for Monkey A, 1.7° for Monkey B, and 1.1° for Monkey C) and size (0.8 - 1°) of the population receptive field associated with a recording field of view (FOV), as well as ocular dominance columns when monocularly presented to confirm the V1 location. This fast process used a 4 × objective lens mounted on the two-photon microscope and revealed no cell-specific information.

Neuronal responses were then measured with a high-contrast (0.9) Gabor grating (Gaussian-windowed sinusoidal grating) drifting at 2 cycles/s in opposite directions perpendicular to the Gabor orientation. The starting phase of the drafting Gabors was always 0°. The Gabor grating varied at 12 orientations from 0° to 165° in 15° steps, and 6 spatial frequencies from 0.25 to 8 cpd

in 1-octave steps. In addition, three stimulus sizes (with constant stimulus centers) were used at each spatial frequency, for two purposes. First, our pilot measurements suggested very strong surround suppression with larger stimuli. Therefore, comparing the responses to different stimulus sizes helped approximate the RF size of each neuron that produced maximal response and least surround suppression. Second, for neurons whose RF centers and the stimulus center were misaligned, the RFs of some misaligned neurons might be better covered by larger stimuli. Still the RFs of additional neurons might have less overlap with the stimuli. These neurons would have weaker and less orientation-tuned responses because of the Gaussian-blurred stimulus edge, and would most likely be filtered out during our multiple steps of selection of orientation tuned neurons (see below). Specifically, the α of the Gaussian envelope of the Gabor were 0.64λ and 0.85λ at all spatial frequencies, and was additionally smaller at 0.42λ when spatial frequencies were 0.25 - 1 cpd, and larger at 1.06λ when spatial frequencies were 2 - 8 cpd (λ: wavelength; Gabors with the same 0 in wavelength unit had the same number of cycles). Here at the smallest 0 (0.42λ), the Gabors still had sufficient number of cycles (frequency bandwidths = 1 octave) (Graham, 1989), so that the actual stimulus spatial frequencies were precise at nominal values. In terms of visual angle, 0 = 1.68°, 2.56°, and 3.36° at 0.25 cpd; 0.84°, 1.28°, and 1.68° at 0.5 cpd; 0.42°, 0.64°, and 0.85° at 1 cpd; 0.34°, 0.42°, and 0.53° at 2 cpd; 0.17°, 0.21°, and 0.26° at 4 cpd, and 0.08°, 0.11°, and 0.13° at 8 cpd, respectively.

Each stimulus was presented for 1000 ms, with an inter-stimulus interval (ISI) of 1500 ms that is sufficient to allow the calcium signals back to the baseline level (Guan, Zhang, Zhang, Tang, & Yu, 2020). Each stimulus condition was repeated 12 times with six repeats for each opposite direction. When a stimulus was presented monocularly to one eye, the other eye was covered with a translucent eye patch to reduce the impacts of short-term monocular deprivation. For Monkey A, binocular recording preceded monocular recording in separate days. During monocular recording contralateral and ipsilateral stimulations alternated in blocks of trials with at least 10-minute breaks, during which the eye mask was taken off. Recording at a specific viewing distance was completed with all trials at relevant SFs pseudo-randomly presented before proceeding to the next distance. For Monkeys B and C, binocular and monocular recordings were mixed and completed in two daily sessions. At a specific viewing distance, all binocular trials at relevant SFs were completed first, then contralateral and ipsilateral trials were completed in alternating blocks of trials with at least 10-minute eye-patch-off breaks. Again, recordings at a specific viewing distance were completed before proceeding to a different distance. Each block of trials typically lasted 20-25 minutes, but for Monkeys A & B some blocks involving three SFs could last up to 45 minutes. The strength of fluorescent signals (mean luminance of a small area) was monitored and adjusted if necessary for drift of fluorescent signals. We compared the response ratios of the last 2 trials over the first 2 trials for each stimulus condition in these extended blocks with ipsilateral and contralateral stimulations. The respective mean ratios were 0.94 and 0.86, suggesting that the recorded neuronal responses remained largely stable over extended blocks of trials.

Two-photon imaging

Two-photon imaging was performed with a Prairie Ultima IV (In Vivo) two-photon microscope (Prairie Technologies) on Monkeys A and B, or a FENTOSmart two-photon microscope (Femtonics) on Monkey C, and a Ti:sapphire laser (Mai Tai eHP, Spectra Physics). One FOV of 850 x 850 μm^2 was selected in each animal and imaged using a 1000-nm femtosecond laser under a 16 × objective lens (0.8 N.A., Nikon) at a resolution of 1.6 μm/pixel. Fast resonant scanning mode (32 frames per second) was chosen to obtain continuous images of neuronal activity (8 frames per second after averaging every 4 frames). Recording was first performed at a shallower depth, and some neurons with high brightness or unique dendrite patterns were selected as landmarks. In the next daily session, the same FOV at the same depth was located before recording with the help of landmarks, and the depth plane was lowered if recording was performed at a deeper depth (Monkeys A & B).

Because of the time limit, recordings at a specific FOV/depth with monocular and binocular stimulations were completed in 2-3 consecutive daily sessions, but the same neurons could be precisely tracked over multiple recording sessions with the use of multiple landmark cues.

Imaging data analysis: Initial screening of ROIs

Data were analyzed with customized MATLAB codes. A normalized cross-correlation based translation algorithm was used to reduce motion artifacts (Li et al., 2017). Then fluorescence changes were associated with corresponding visual stimuli through the time sequence information recorded by Neural Signal Processor (Cerebus system, Blackrock Microsystem). By subtracting the mean of the 4 frames before stimuli onset (*F0*) from the average of the 6th-9th frames after stimuli onset (*F*) across 5 or 6 repeated trials for the same stimulus condition (same orientation, spatial frequency, size, and drifting direction), the differential image ($\Delta F = F - F0$) was obtained. For a specific FOV at a specific recording depth, the regions of interest (ROIs) or possible cell bodies were decided through sequential analysis of 216 differential images in the order of spatial frequency (6), size (3), and orientation (12) (6 x 3 x 12 = 216). The first differential image was filtered with a band-pass Gaussian filter (size = 2-10 pixels), and connected subsets of pixels (>25 pixels, which would exclude smaller vertical neuropils) with average pixel value > 3 standard deviations of the mean brightness were selected as ROIs. Then the areas of these ROIs were set to mean brightness in the next differential image before the bandpass filtering and thresholding were performed (This measure gradually reduced the SDs of differential images and facilitated detection of neurons with relatively low fluorescence responses). If a new ROI and an existing ROI from the previous differential image overlapped, the new ROI would be on its own if the overlapping area OA < 1/4 ROI_{new}, discarded if 1/4 ROI_{new} < OA < 3/4 ROI_{new}, and merged with the existing ROI if OA > 3/4 ROI_{new}. The merges would help smooth the contours of the final ROIs. This process went on through all 512 differential images twice to select ROIs. Finally, the roundness for each ROI was calculated as:

$$Roundness = \frac{\sqrt{4\pi \times A}}{P}$$

where *A* was the ROI's area, and *P* was the perimeter. Only ROIs with roundness larger than 0.9, which would exclude horizontal neuropils, were selected for further analysis.

Imaging data analysis: Orientation tuning, SF tuning, and ocular dominance

The ratio of fluorescence change ($\Delta F/F0$) was calculated as a neuron's response to a specific stimulus condition. For a specific cell's response to a specific stimulus condition, the $F0_n$ of the n-th trial was the average of 4 frames before stimulus onset, and F_n was the average of 5th-8th frames after stimulus onset. $F0_n$ was then averaged across 12 trials to obtain the baseline F0 for all trials (for the purpose of reducing noises in the calculations of responses), and $\Delta F_n/F0 = (F_n\text{-}F0)/F0$ was taken as the neuron's response to this stimulus at this trial. For a small portion of neurons (e.g., ~3% in Monkeys A, ~8% in monkey B and ~2% in Monkey C) showing direction selectivity as their responses to two opposite directions differed significantly ($p < 0.05$, Friedman test), the 6 trials at the preferred direction was considered for calculations of $\Delta F_n/F0$ as the cell's responses to a particular stimulus. F0 was still averaged over 12 trials at two opposite directions.

Several steps were then taken to decide whether a neuron was tuned to orientation and/or spatial frequency, and if so, its ocular dominance index. For each monocular condition, first the orientation, SF, and size (0) producing the maximal response among all conditions were selected. Then responses to other 11 orientations and 5 SFs were decided at the selected SF and size. Second, to select orientation and/or SF tuned neurons, a non-parametric Friedman test was performed to

test whether a neuron's responses at 12 orientations or 6 SFs were significantly different from each other at least under one monocular stimulation. To reduce Type-I errors, the significance level was set at α = 0.01.

Third, for those showing significant orientation difference, the trial-based orientation responses of each neuron were fitted with a Gaussian model with a MATLAB nonlinear least-squares function: lsqnonlin:

$$R(\theta) = a_1 2^{-(\frac{\theta-\theta_0}{\sigma})^2} + b$$

where R(8) was the response at orientation 8, free parameters a_1, 8_0, 0, and b were the amplitude, peak orientation, standard deviation of the Gaussian function, and minimal response of the neuron, respectively. Only neurons with goodness of fit $R^2 > 0.5$ at least under one stimulation condition were finally selected as orientation tuned neurons. Fourth, for those showing significant SF difference, the trial-based SF responses of each neuron were further fitted with a Difference-of-Gaussian model.

$$R(sf) = a_1 e^{-(\frac{sf}{\sigma_1})^2} - a_2 e^{-(\frac{sf}{\sigma_2})^2} + b$$

where R(sf) was a neuron's response at spatial frequency sf, free parameters a_1, 0_1, a_2, and 0_2 were amplitudes and standard deviations of two Gaussians, respectively, and b was the minimal response among 6 spatial frequencies. Only those with goodness of fit $R^2 > 0.5$at least under one monocular stimulation were selected as SF tuned neurons.

The ocular dominance index (ODI) was calculated to characterize each orientation and/or SF tuned neuron's eye preference: $ODI = (R_i - R_c)/(R_i + R_c)$, in which R_i and R_c were the neuron's respective peak responses at the best orientation and SF to ipsilateral and contralateral stimulations on the basis of data fitting. Here ODI = −1 and 1 would indicate complete contralateral and ipsilateral eye preferences, respectively, and ODI = 0 would indicate equal preferences to both eyes.

Model fitting

Monocular and binocular data in **Fig. 4** were fitted by **Eqs. 2** and 3, respectively. The goodness-of-fit was indicated by a fit quality index q with a range of 0-1, which is the root mean square deviation between the observed responses and the model normalized by the observed response mean (Busse et al., 2009):

$$q = 1 - \frac{\sqrt{\frac{\sum_{i=1}^{n}(r_i - m_i)^2}{n}}}{\bar{r}}$$

where *i* indicates the i_{th} bin, *r* indicates median response of a specific bin, and m indicates the corresponding model predictions.

Acknowledgements

This study was supported by the National Science and Technology Innovation 2030 Major Program (2022ZD0204600), Natural Science Foundation of China grants 31230030 and 31730109, and funds from Peking-Tsinghua Center for Life Sciences, Peking University.

Competing interests

None.

References

Anderson P. A., Movshon J. A (1989) **Binocular combination of contrast signals** *Vision Res* **29**:1115–1132

Barlow H. B., Blakemore C., Pettigrew J. D (1967) **The neural mechanism of binocular depth discrimination** *J Physiol* **193**:327–342

Busse L., Wade A. R., Carandini M (2009) **Representation of concurrent stimuli by population activity in visual cortex** *Neuron* **64**:931–942

Cogan A. I (1987) **Human binocular interaction: towards a neural model** *Vision Res* **27**:2125–2139

Cohn T. E., Lasley D. J (1976) **Binocular vision: two possible central interactions between signals from two eyes** *Science* **192**:561–563

DeSilva H. R., Bartley S. H (1930) **Summation and subtraction of brightness in binocular perception** *British Journal of Psychology* **20**:242–252

Ding J., Klein S. A., Levi D. M (2013) **Binocular combination of phase and contrast explained by a gain-control and gain-enhancement model** *J Vis* **13**

Ding J., Sperling G (2006) **A gain-control theory of binocular combination** *Proc Natl Acad Sci U S A* **103**:1141–1146

Dougherty K., Cox M. A., Westerberg J. A., Maier A (2019) **Binocular Modulation of Monocular V1 Neurons** *Curr Biol* **29**:381–391

Graham N (1989) **Visual Pattern Analyzers (Oxford Psychology Series, No. 16)**

Guan S. C., Ju N., Tao L., Tang S. M., Yu C (2021) **Functional organization of spatial frequency tuning in macaque V1 revealed with two-photon calcium imaging** *Progress in Neurobiology* **205**

Guan S. C., Zhang S. H., Zhang Y. C., Tang S., Yu C (2020) **Plaid detectors in macaque V1 revealed by two-photon calcium imaging** *Current Biology* **30**:934–940

Henriksen S., Tanabe S., Cumming B (2016) **Disparity processing in primary visual cortex** *Philos Trans R Soc Lond B Biol Sci* **371**

Huang C. B., Zhou J., Zhou Y., Lu Z. L (2010) **Contrast and phase combination in binocular vision** *PLoS One* **5**

Hubel D. H., Wiesel T. N (1962) **Receptive fields, binocular interaction and functional architecture in the cat's visual cortex** *J Physiol* **160**:106–154

Hubel D. H., Wiesel T. N (1968) **Receptive fields and functional architecture of monkey striate cortex** *J Physiol* **195**:215–243

Ju N., Guan S. C., Tao L., Tang S. M., Yu C (2021) **Orientation tuning and end-stopping in macaque V1 studied with two-photon calcium imaging** *Cerebral Cortex* **31**:2085–2097

Li M., Liu F., Jiang H., Lee T. S., Tang S (2017) **Long-term two-photon imaging in awake macaque monkey** *Neuron* **93**:1049–1057

Li Z. P., Atick J. J (1994) **Efficient Stereo Coding in the Multiscale Representation** *Network-Computation in Neural Systems* **5**:157–174

Lunghi C., Burr D. C., Morrone C (2011) **Brief periods of monocular deprivation disrupt ocular balance in human adult visual cortex** *Curr Biol* **21**:R538–539

Mitchell B. A., Dougherty K., Westerberg J. A., Carlson B. M., Daumail L., Maier A., Cox M. A (2022) **Stimulating both eyes with matching stimuli enhances V1 responses** *iScience* **25**

Moradi F., Heeger D. J (2009) **Inter-ocular contrast normalization in human visual cortex** *J Vis* **9**:11–22

Parker A. J., Smith J. E., Krug K (2016) **Neural architectures for stereo vision** *Philos Trans R Soc Lond B Biol Sci* **371**

Pelli D. G., Zhang L (1991) **Accurate control of contrast on microcomputer displays** *Vision Res* **31**:1337–1350

Prince S. J., Pointon A. D., Cumming B. G., Parker A. J (2002) **Quantitative analysis of the responses of V1 neurons to horizontal disparity in dynamic random-dot stereograms** *J Neurophysiol* **87**:191–208

Read J. C. A (2021) **Binocular vision and stereopsis across the animal kingdom** *Annu Rev Vis Sci* **7**:389–415

Welchman A. E. (2016) **The Human Brain in Depth: How We See in 3D** *Annu Rev Vis Sci* **2**:345–376

Article and author information

Sheng-Hui Zhang
PKU-Tsinghua Center for Life Sciences

Xing-Nan Zhao
PKU-Tsinghua Center for Life Sciences

Shi-Ming Tang
PKU-Tsinghua Center for Life Sciences, School of Life Sciences, IDG-McGovern Institute for Brain Research
For correspondence: tangshm@pku.edu.cn
ORCID iD: 0000-0003-0294-3259

Cong Yu
IDG-McGovern Institute for Brain Research, School of Psychological and Cognitive Sciences, Peking University, Beijing, China
For correspondence: yucong@pku.edu.cn
ORCID iD: 0000-0002-8453-6974

Copyright

Editors

Reviewing Editor
Eric Gaier
Boston Children's Hospital, Boston, United States of America

Senior Editor
Joshua Gold
University of Pennsylvania, Philadelphia, United States of America

Reviewer #1 (Public Review):

Summary:
Zhang et al., investigated the relationship between monocular and binocular responses of V1 superficial-layer neurons using two-photon calcium imaging. They found a strong relationship in their data: neurons that exhibited a greater preference for one eye or the other (high ocular dominance) were more likely to be suppressed under binocular stimulation, whereas neurons that are more equivalently driven by each other (low ocular dominance) were more likely to be enhanced by binocular stimulation. This result chiefly demonstrates the relationship between ocular dominance and binocular responses in V1, corroborating what has been shown previously using electrophysiological techniques but now with greater spatial resolution (albeit less temporal resolution). The binocular responses were well-fitted by a model that institutes divisive normalization between the eyes that accounts for both the suppression and enhancement phenomena observed in the subpopulation of binocular neurons. In so doing, the authors reify the importance of incorporating ocular dominance in computational models of binocular combination.

The conclusions of this paper are mostly well supported by the data, but there are some limitations of the methodology that need to be clarified, and an expansion of how the results relate to previous work would better contextualize these important findings in the literature.

Strengths:
The two-photon imaging technique used to resolve the activity of individual neurons within intact brain tissue grants a host of advantages. Foremost, two-photon imaging confers considerably high spatial resolution. As a result, the authors were able to sample and analyze the activity from thousands of verified superficial-layer V1 neurons. The animal model used, awake macaques, is also highly relevant for the study of binocular combination. Macaques, like humans, are binocular animals, meaning they have forward-facing eyes that confer overlapping visual fields. Importantly, macaque V1 is organized into cortical columns that process specific visual features from the separate eyes just like in humans. In combination

with a powerful imaging technique, this allowed the authors to evaluate the monocular and binocular response profiles of V1 neurons that are situated within neighboring ocular dominance columns, a novel feat. To this aim, the approach was well-executed and should instill further confidence in the notion that V1 neurons combine monocular information in a manner that is dependent on the strength of their ocular dominance.

Weaknesses:
While two-photon imaging provides excellent spatial resolution, its temporal resolution is often lower compared to some other techniques, such as electrophysiology. This limits the ability to study the fast dynamics of neuronal activity, a well-understood trade-off of the method. The issue is more so that the authors draw comparisons to electrophysiological studies without explicit appreciation of the temporal difference between these techniques. In a similar vein, two-photon imaging is limited spatially in terms of cortical depth, preferentially sampling from neurons in layers 2/3. This limitation does not invalidate any of the interpretations but should be considered by readers, especially when making comparisons to previous electrophysiological reports using microelectrode linear arrays that sample from all cortical layers. Indeed, it is likely that a complete picture of early cortical binocular processing will require high spatial resolution (i.e., sampling from neurons in neighboring ocular dominance columns, from pia mater to white matter) at the biophysically relevant timescales (1ms resolution, capturing response dynamics over the full duration of the stimulus presentation, including the transient onset and steady-state periods).

https://doi.org/10.7554/eLife.92839.1.sa2

Reviewer #2 (Public Review):

Summary:
This study examines the pattern of responses produced by the combination of left-eye and right-eye signals in V1. For this, they used calcium imaging of neurons in V1 of awake, fixating monkeys. They take advantage of calcium imaging, which yields large populations of neurons in each field of view. With their data set, they observe how response magnitude relates to ocular dominance across the entire population. They analyze carefully how the relationship changed as the visual stimulus switched from contra-eye only, ipsi-eye only, and binocular. As expected, the contra-eye-dominated neurons responded strongly with a contra-eye-only stimulus. The ipsi-eye-dominated neurons responded strongly with an ipsi-eye-only stimulus. The surprise was responses to a binocular stimulus. The responses were similarly weak across the entire population, regardless of each neuron's ocular dominance. They conclude that this pattern of responses could be explained by interocular divisive normalization, followed by binocular summation.

Strengths:
A major strength of this work is that the model-fitting was done on a large population of simultaneously recorded neurons. This approach is an advancement over previous work, which did model-fitting on individual neurons. The fitted model in the manuscript represents the pattern observed across the large population in V1, and washes out any particular property of individual neurons. Given the large neuronal population from which the conclusion was drawn, the authors provide solid evidence supporting their conclusion. They also observed consistency across 5 fields of view.

The experiments were designed and executed appropriately to test their hypothesis. Their data support their conclusion.

Weaknesses:
One weakness of their study is that calcium signals can exaggerate the nonlinear properties of neurons. Calcium imaging renders poor responses poorer and strong responses stronger,

compared to single-unit recording. In particular, the dramatic change in the population response between monocular stimulation and binocular stimulation could actually be less pronounced when measured with single-unit recording methods. This means their choice of recording method could have accidentally exaggerated the evidence of their finding.

The implication of their finding is that strong ocular dominance is the result of release from interocular suppression by a monocular stimulus, rather than the lack of binocular combination as many traditional studies have assumed. This could significantly advance our understanding of the binocular combination circuitry of V1. The entire population of neurons could be part of a binocular combination circuitry present in V1.

https://doi.org/10.7554/eLife.92839.1.sa1

Reviewer #3 (Public Review):

The authors have made simultaneous recordings of the responses of large numbers of neurons from the primary visual cortex using optical two-photon imaging of calcium signals from the superficial layers of the cortex. Recordings were made to compare the responses of the cortical neurons under normal binocular viewing of a flat screen with both eyes open and monocular viewing of the same screen with one eye's view blocked by a translucent filter. The screen displayed visual stimuli comprising small contrast patches of Gabor function distributions of luminance, a stimulus that is known to excite cortical neurons.

This is an important data set, given the large numbers of neurons recorded. The authors present a simple model to explain the binocular combination of neuronal signals from the right and left eyes.

The limitations of the paper as written are as follows. These points can be addressed with some additional analysis and rewriting of sections of the paper. No new experimental data need to be collected.

1. The authors should acknowledge the fact that these recordings arise from neurons in the superficial layers of the cortex. This limitation arises from the usual constraints on optical imaging in the macaque cortex. This means that the sample of neurons forming this data set is not fully representative of the population of binocular neurons within the visual cortex. This limitation is important in comparing the outcome of these experiments with the results from other studies of binocular combination, which have used single-electrode recording. Electrode recording will result in a sample of neurons that is drawn from many layers of the cortex, rather than just the superficial layers.
2. Single-neuron recording of binocular neurons in the primary visual cortex has shown that these neurons often have some spontaneous activity. Assessment of this spontaneous level of firing is important for accurate model fitting [1]. The paper here should discuss the level of spontaneous neuronal firing and its potential significance.

3. The arrangements for visual stimulation and comparison of binocular and monocular responses mean that the stereoscopic disparity of the binocular stimuli is always at zero or close to zero. The animal's fixation point is in the centre of a single display that is viewed binocularly. The fixation point is, by definition, at zero disparity. The other points on the flat display are also at zero disparity or very close to zero because they lie in the same depth plane. There will be some small deviations from exactly zero because the geometry of the viewing arrangements results in the extremities of the display being at a slightly different distance than the centre. Therefore, the visual stimulation used to test the binocular condition is always at zero disparity, with a slight deviation from zero at the edges of the display, and never changes. [There is a detail that can be ignored. The experimenters tested neurons with visual stimulation at different real distances from the eyes, but this is not relevant here. Provided the animals accurately converged their eyes on the provided binocular fixation point, then the disparity of the visual stimuli will always be at or close to zero, regardless of viewing distance in these circumstances.] However, we already know from earlier work that neurons in the visual cortex exhibit a range of selectivity for binocular disparity. Some neurons have their peak response at non-zero disparities, representing binocular depths nearer than the fixation depth or beyond it. The response of other neurons is maximally suppressed by disparities at the depth of the fixation point (so-called Tuned Inhibitory [TI] neurons). The simple model and analysis presented in the paper for the summation of monocular responses to predict binocular responses will perform adequately for neurons that are tuned to zero disparity, so-called tuned excitatory neurons [TE], but is necessarily compromised when applied to neurons that have other, different tuning profiles. Specifically, when neurons are stimulated binocularly with a non-preferred disparity, the binocular response may be lower than the monocular response[2, 3]. This more realistic view of binocular responses needs to be considered by the authors and integrated into their modelling.

4. The data in the paper show some features that have been reported before but are not captured by the model. Notably for neurons with extreme values of ocular dominance, the binocular response is typically less than the larger of the two monocular responses. This is apparent in the row of plots in Figure 2D from individual animals and in the pooled data in Figure 2E. Responses of this type are characteristic of tuned inhibitory [TI] neurons[2]. It is not immediately clear why this feature of the data does not appear in the summary and analysis in Figure 3. The paper text states that the responses were "first normalized by the median of the binocular responses". This will certainly get rid of this characteristic of the data, but this step needs better justification, or an amendment to the main analysis is needed. In the present form, the model and analysis do not appear to fit the data in Figure 2 as accurately as needed. The authors should address the discrepancy between the data as presented in Figures 2D, E, and Figure 3.

Citations

1. Prince, S.J.D., Pointon, A.D., Cumming, B.G., and Parker, A.J., (2002). Quantitative analysis of the responses of V1 neurons to horizontal disparity in dynamic random-dot stereograms. Journal of Neurophysiology, 87: 191-208.
2. Prince, S.J.D., Cumming, B.G., and Parker, A.J., (2002). Range and mechanism of encoding of horizontal disparity in macaque V1. Journal of Neurophysiology, 87: 209-221.
3. Poggio, G.F. and Fischer, B., (1977). Binocular interaction and depth sensitivity in striate and prestriate cortex of behaving rhesus monkey. Journal of Neurophysiology, 40: 1392-1405 doi 10.1152/jn.1977.40.6.1392.

https://doi.org/10.7554/eLife.92839.1.sa0