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

Sheng-Hui Zhang[1*], Xing-Nan Zhao[1*], Shi-Ming Tang[1,2,3#], & Cong Yu[3,4#] Feb. 06, 2024

[1]PKU-Tsinghua Center for Life Sciences, [2]School of Life Sciences, [3]IDG-McGovern Institute for Brain Research, and [4]School of Psychological and Cognitive Sciences, Peking University, Beijing, China

* Equal-contribution first authors.

Correspondence authors: Shi-Ming Tang (tangshm@pku.edu.cn) & Cong Yu (yucong@pku.edu.cn)

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.

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 (Poggio & Fischer, 1977; 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; Shatz & Stryker, 1978). 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.

Previous neurophysiological recording studies have revealed 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, there is evidence that neurons' eye preference may have important functional roles. For example, Dougherty et al. (2019) reported that the responses of monocular neurons are 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. In addition, Mitchell, Carlson, Westerberg, Cox, and Maier (2023) reported that binocular combination of monocular stimuli with different contrasts is also affected by neurons' eye preference.

A big and more complete picture of binocular combination of monocular responses, as well as the impacts of eye preferences, would come from large samples of neurons in neighboring ocular dominance columns, 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 calcium 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 presented 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 had 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, 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).

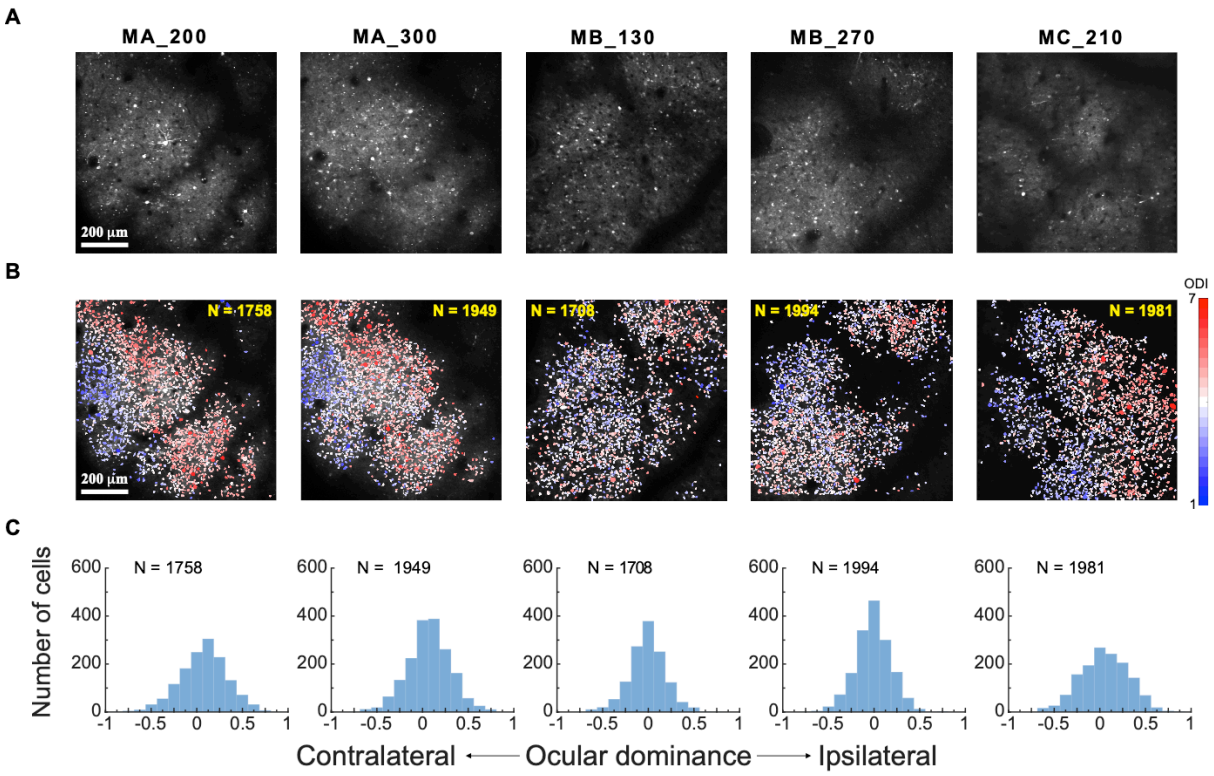


Figure 1. Eye preferences of V1 superficial-layer neurons in three macaques. **A**. Two-photon imaging. Average 2-photon images over a recording session for each response FOV. MA_200: Monkey A at a 200-µm cortical depth. **B**. Ocular dominance functional maps of each FOV/depth at single-neuron resolution. **C**. Frequency distributions of neurons of each FOV/depth as a function of ocular dominance index.

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.

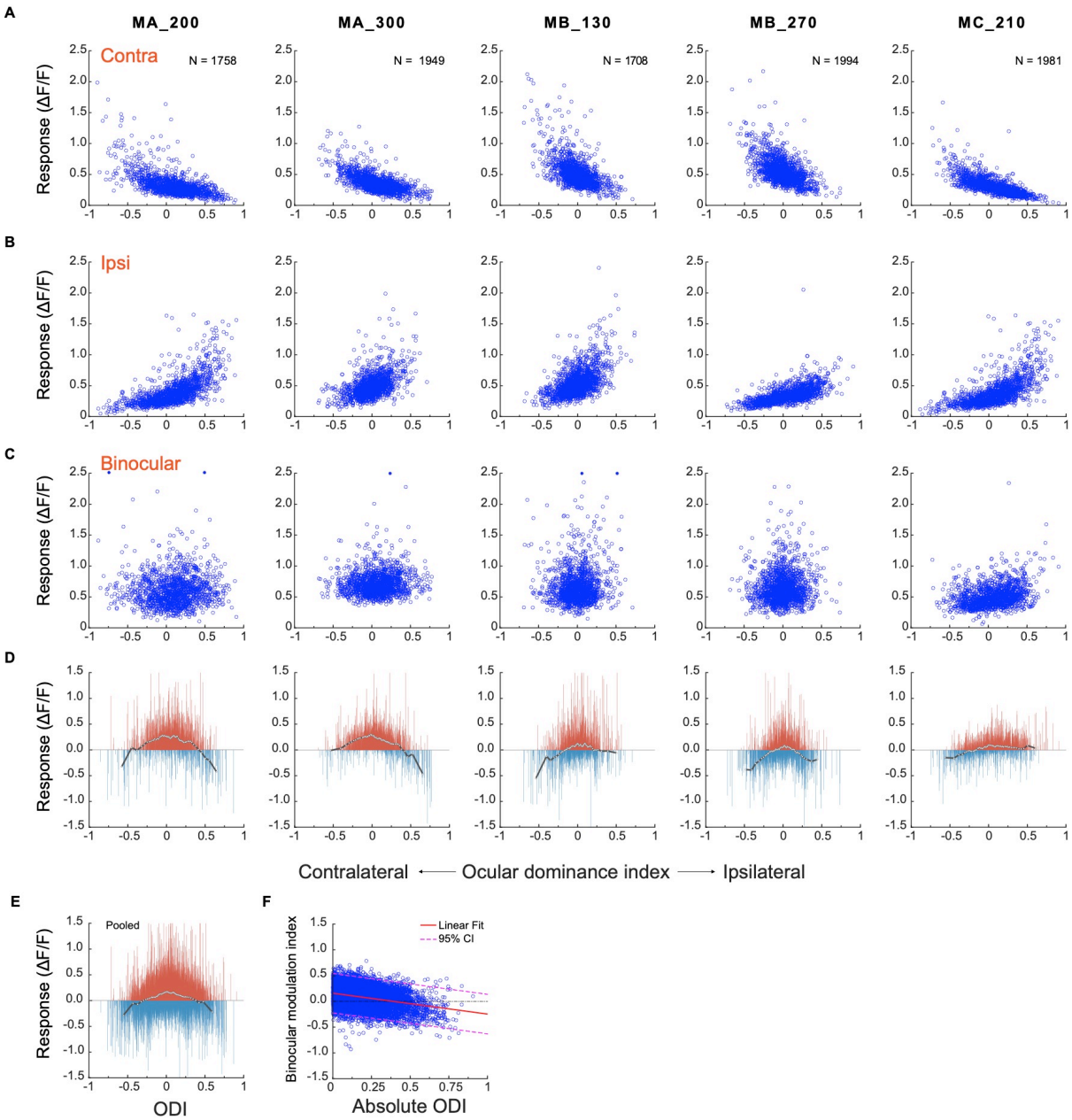


Figure 2. A comparison of neuronal responses to monocular and binocular stimulations. **A**. Responses of individual neurons against their ocular dominance indices with contralateral stimulation. **B**. Responses of the same neurons against their ocular dominance indices with ipsilateral stimulation. **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

125 of respective bins and are connected with a black line. **E**. The differences between binocular
126 and monocular responses of individual neurons pooled over five FOVs/depths. **F**. Binocular
127 modulation index as a function of absolute ODI and the linear fit. The binocular modulation
128 index of each neuron was defined as $(R_b - \max(R_i, R_c))/(R_b + \max(R_i, R_c))$.
129
130 *Modeling monocular and binocular responses*
131 We used the following steps to develop a model that can account for the current
132 monocular responses and their binocular combination. Monocular and binocular data of
133 each FOV/depth, as well as the pooled data, were first normalized by the respective median
134 of the binocular responses of all neurons in the same FOV/depth. The normalized data were
135 then divided into 60 bins in the order of the ocular dominance index, and the median values
136 of 60 bins were used for model fitting (Figs. 3A-C).
137 <u>Monocular responses</u>
138 First, a neuron's monocular responses to contralateral and ipsilateral stimulations
139 could be respectively described by a divisive gain control model:
140 $$R_c = \frac{w_c^{m_c} \cdot S_c}{k} \quad \text{and} \quad R_i = \frac{w_i^{m_i} \cdot S_i}{k} \qquad \text{(Eq. 1)}$$
141 Here S_c and S_i were stimulus contrasts, w_c and w_i were linear transformations of a neuron's
142 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
143 represented monocular nonlinearity, and k represented divisive gain control. Because S_c and
144 S_i in our experiments were 0.9, which were about equal to 1 (full contrast) due to neuronal
145 response saturation, the above equations were simplified as
146 $$R_c = \frac{w_c^{m_c}}{k} \quad \text{and} \quad R_i = \frac{w_i^{m_i}}{k} \qquad \text{(Eq. 2)}$$

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 0.92-0.94 for the contralateral condition (Fig. 3A) and 0.87-0.93 for the ipsilateral condition (Fig. 3B), suggesting adequate goodness of fit. The fitting parameters are listed in Fig. 3D.

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. Therefore, we added interocular response suppression to Eq. 2 by letting monocular responses from each eye be further normalized by an interocular suppression factor w_i^b or w_c^b (Eq. 3). In other words, the strength of interocular response suppression was decided by the linearly transformed ODI with a nonlinear exponent b. Finally, the normalized responses from two eyes were summed to simulate the binocular responses of neurons R_b, completing the model of binocular combination (Eq. 3).

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

Although not shown in Eq. 3, we also assumed that the nonlinear exponent b also depends on the contrast of the stimulus presented to the other eye (i.e., S_c or S_i). Consequently, when S_c or $S_i = 0$ under monocular stimulation, R_c or $R_i = 0$ (Eq. 1), and interocular suppression w_i^b or $w_c^b = 1$, so Eq. 3 changes back to Eq. 2. It is only when S_c and S_i

170 are equal and close to 1, as in the current study, that interocular suppression and binocular

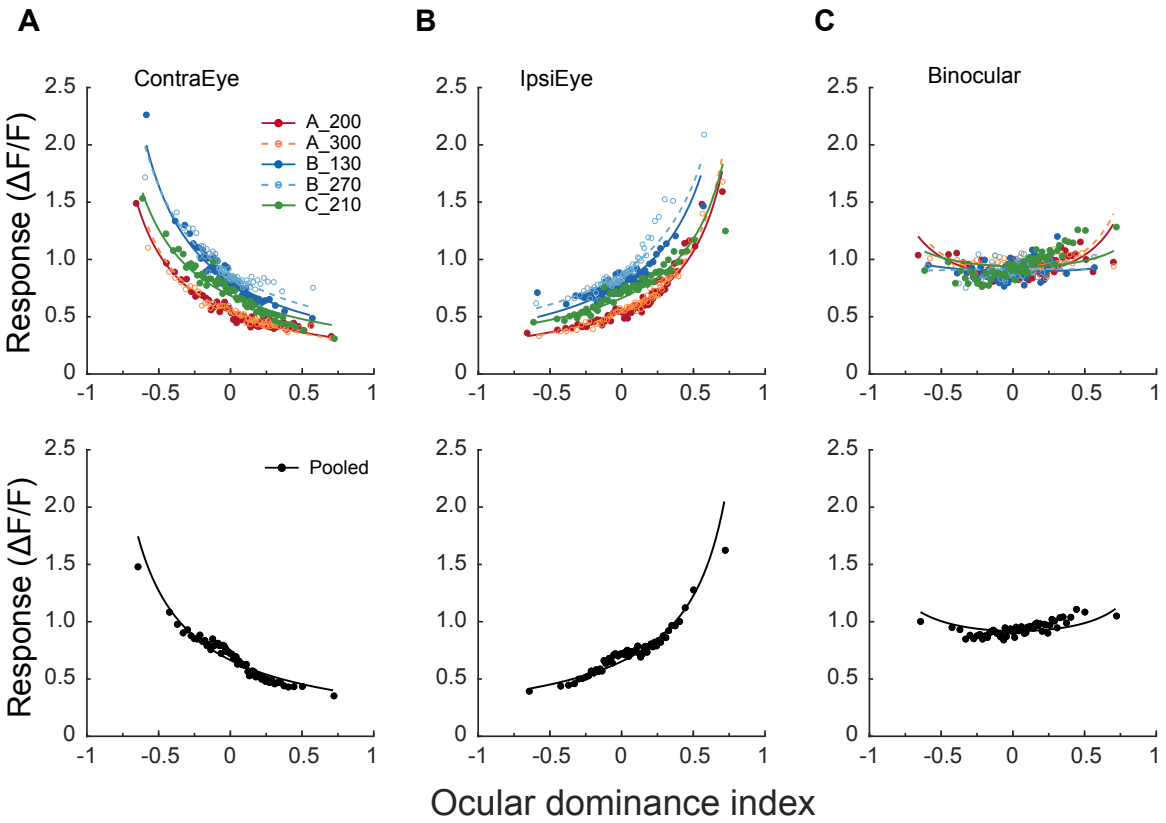
171 combination would be in the current Eq. 3 format.

172 When fitting binocular responses, the parameters m_i, m_c, and k were inherited from

173 earlier monocular data fitting and remained fixed. Only b was allowed to change. Data fitting

174 resulted in flat binocular response functions (Fig. 3C) with satisfactory goodness-of-fit (fit

175 quality index = 0.88 to 0.92).



D

	m	k		m	k		b
A_200	-0.965	3.655		-0.986	3.655		0.301
A_300	-1.018	3.801		-1.042	3.801		0.289
B_140	-1.013	2.513		-0.981	2.513		0.828
B_250	-0.905	2.152		-0.923	2.152		0.940
C_210	-0.867	2.677		-0.829	2.677		0.519
Pooled	-0.932	2.870		-0.932	2.870		0.521

176

177 Figure 3. Modeling monocular and binocular responses. **A & B**. Median neuronal responses

178 to contralateral and ipsilateral stimulations as a function of ocular dominance index and

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

187

188 The effects of interocular suppression and binocular summation in the model, as well
189 as their contributions to the binocular response, may be better appreciated in Fig. 4. Fig. 4A
190 uses Eqs. 2 and 3 and the parameters from fitting of pooled data (Fig. 3D) to simulate the
191 contralateral (blue curve with label R_c), ipsilateral (red curve with label R_i), and binocular
192 (black curve) response functions against the ocular dominance index. The arithmetic sum of
193 contralateral and ipsilateral response functions was also simulated (grey dashed curve
194 labeled $R_i + R_c$). In addition, neuronal responses to preferred eye stimulation would consist of
195 the higher branches of contralateral and ipsilateral response functions. It is apparent that
196 binocular responses cannot be explained by the sum of monocular responses, as binocular
197 responses are substantially lower than the summed monocular responses for both
198 monocular and binocular neurons. Nor can binocular responses be explained by the
199 responses to the preferred eye, as binocular responses are also lower than those to the
200 preferred eye (the larger of the two monocular responses) for monocular neurons. Instead,
201 the median of the binocular response function (black arrow) in each data set is close to but

202 still more or less higher than the median of the contralateral (blue arrow) or ipsilateral

203 response function (red arrow), which is consistent with previous reports (Prince et al., 2002;

204 Mitchell et al., 2022).

205 Fig. 4B plots the interocular suppression factor w_c^b (dotted blue curve) for the

206 contralateral response R_c (solid blue curve from 4A), and w_i^b (dotted red curve) for the

207 ipsilateral response R_i (solid red curve from 4A). The interocular suppression factors w_c^b and

208 w_i^b are larger with neurons that are more monocular than with neurons that are more

209 binocular. The R_c and R_i are divided by the respective interocular suppression factor,

210 producing the normalized contralateral responses (dashed blue curve labeled R_c/w_c^b) and

211 ipsilateral responses (dashed red curve labeled R_i/w_i^b). These normalized curves are lower

212 than original monocular responses, especially for neurons that are more monocular (ODIs

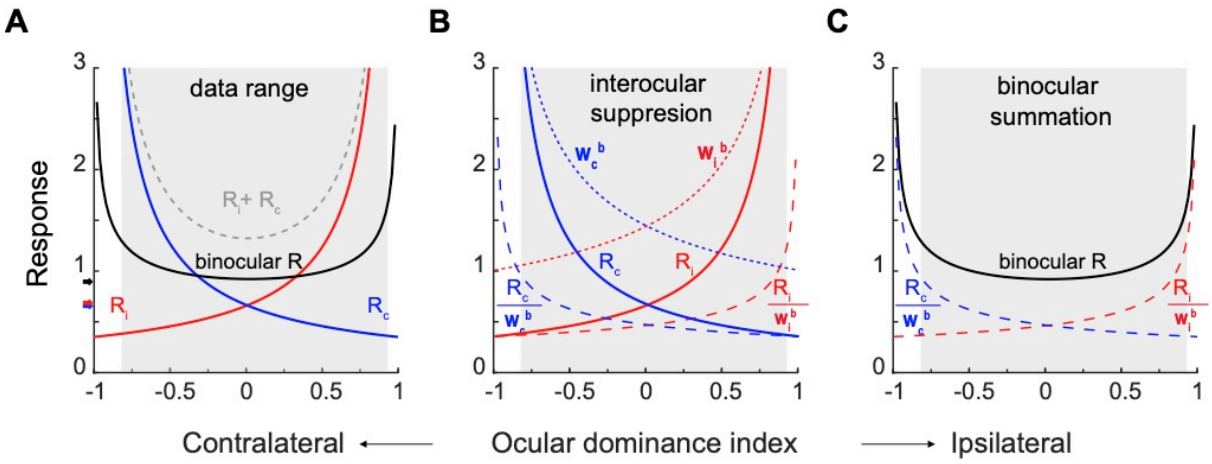
213 farther from 0), showing interocular suppression.

214 Then the normalized monocular curves are summed up in Fig. 4C, which represents

215 binocular summation and produces the final binocular responses (R_b). Therefore, as a result

216 of combined interocular suppression and binocular summation, the final curve for binocular

217 responses becomes nearly flat within the data range.



218

219 Figure 4. Monocular and binocular responses modeled with binocular suppression and

220 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 (R_c & R_i) are divided by respective interocular suppression factors W_c^b and W_i^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 R_b.

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 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. These findings implicate that at least for neurons in superficial layers of V1, significant ocular dominance may result from a release of interocular suppression during monocular stimulation, an unusual viewing condition as our vision is typically binocular, rather than a lack of binocular combination of inputs from upstream monocular neurons.

We introduced this paper by citing the stable vision with monocular or binocular viewing. Relevant to this issue, Fig. 4A shows that the median of binocular neuronal responses is only 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). The overall similar monocular and binocular responses may account for a large proportion of the stable vision under both viewing conditions. What is intriguing is our finding that the very small response changes actually comprise of dramatic binocular suppression with monocular neurons and facilitation with binocular neurons. Furthermore, slightly higher binocular responses suggest that they might be further modulated by additional mechanisms to achieve stable vision.

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 the observation of interocular suppression of monocular neurons responding to the preferred eye in Dougherty et al. (2019), 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. In addition, it is unclear whether the discrepancies are caused by different temporal resolutions of electrode recording and calcium imaging. The results of Dougherty et al. (2019) represent changes of neuronal spike activities over a period of approximately 50-200 ms after the stimulus onset, which may reflect the sustained

neuronal responses to the stimulus and possible feedback signals. Calcium signals are much slower and indicative of the aggregated neuronal responses over a longer period (up to 1000 ms in the current study. They should have smeared, rather than exaggerated, the differences between monocular and binocular responses, though we cannot exclude the possibility that some neuronal response changes beyond 200 ms are responsible for the discrepancies.

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; Mitchell et al., 2023). 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, and that Mitchell et al. (2023) revealed that binocular combination of different contrasts presented to different eyes are affected by neurons' ocularity preference. 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

294 contrasts due to physical contrast differences (Anderson & Movshon, 1989; Mitchell et al.,
295 2023), monocular adaptation (Anderson & Movshon, 1989), and short-term monocular
296 deprivation (Lunghi, Burr, & Morrone, 2011), as well as the relevant roles of ocular
297 dominance of individual neurons. These results would help understand abnormal binocular
298 vision in patients with strabismus and amblyopia. In addition, in our experiments, binocular
299 stimuli were presented with zero disparity, which best triggered the responses of neurons
300 with zero-disparity tuning. A more realistic model of binocular combination also requires the
301 consideration of neurons with other disparity-tuning profiles.
302

303 Limitations of the current study

304 Although capable of sampling a large number of neurons at cellular resolution and
305 with low sampling bias, two-photon calcium imaging has its known limitations that may
306 better make it a complementary research tool to electrophysiological recordings. For
307 example, two-photon imaging can only sample neurons from superficial-layers, while
308 binocular neurons also exist in deeper layers, and even neurons in the input layer are
309 affected by feedback from downstream binocular neurons to exhibit binocular response
310 properties (Dougherty et al., 2019). Furthermore, calcium signals are relatively slow and
311 cannot reveal the fast dynamics of neuronal responses. Due to these spatial and temporal
312 limitations, a more complete picture of the neuronal mechanisms underlying binocular
313 combination of monocular responses may come from studies using both technologies.

314 In addition, calcium signals may exaggerate the nonlinear properties of neurons.
315 Although calcium signals indicated by GCaMP5, our favored choice of calcium indicator,
316 displays a linear relationship to neuronal spike rates within a range of 10-150 Hz (Li et al.,
317 2017), weak and strong signals out of this range are more nonlinear, and may appear poorer

318 and stronger, respectively, than electrode-recorded effects. Consequently, the differences in
319 population responses between monocular and binocular stimulations revealed by this study
320 might be less pronounced.
321
322 **Materials and Methods**
323 *Monkey preparation*
324 Monkey preparations were identical to those reported in a previous study (Guan, Ju,
325 Tao, Tang, & Yu, 2021; Ju, Guan, Tao, Tang, & Yu, 2021). Three rhesus monkeys (*Macaca*
326 *mulatta*) aged 4 to 6 years, respectively, were each prepared with two sequential surgeries
327 under general anesthesia and strictly sterile condition. In the first surgery, a 20-mm diameter
328 craniotomy was performed on the skull over V1. The dura was opened and multiple tracks of
329 100-150 nL AAV1.hSynap.GCaMP5G.WPRE.SV40 (AV-1-PV2478, titer 2.37e13 (GC/ml), Penn
330 Vector Core) were pressure-injected at a depth of ~350 μm. Then the dura was sutured, the
331 skull cap was re-attached with three titanium lugs and six screws, and the scalp was sewn up.
332 After the surgery, the animal was returned to the cage, treated with injectable antibiotics
333 (Ceftriaxone sodium, Youcare Pharmaceutical Group, China) for one week. Postop analgesia
334 was also administered. The second surgery was performed 45 days later. A T-shaped steel
335 frame was installed for head stabilization, and an optical window was inserted onto the
336 cortical surface. Data collection could start as early as one week later. More details of the
337 preparation and surgical procedures can be found in Li et al. (2017). The procedures were
338 approved by the Institutional Animal Care and Use Committee, Peking University.
339
340 *Behavioral task*
341 After a ten-day recovery from the second surgery, monkeys were seated in primate

342 chairs with head restraint. They were trained to hold fixation on a small white spot (0.1°)
343 with eye positions monitored by an ISCAN ETL-200 infrared eye-tracking system (ISCAN Inc.)
344 at a 120-Hz sampling rate (Monkeys A & B) or an Eyelink-1000 (SR Research) at a 1000-Hz
345 sampling rate (Monkey C). During the experiment, trials with the eye position deviated 1.5°
346 or more from the fixation before stimulus offset were discarded as ones with saccades and
347 repeated. For the remaining trials, the eye positions were mostly concentrated around the
348 fixation point, with eye positions in over 95% of trials within 0.5° from the fixation point. The
349 viewing was binocular.
350
351 *Visual stimuli*
352 For Monkeys A and B, visual stimuli were generated by the ViSaGe system (Cambridge
353 Research Systems) and presented on a 21'' Sony G520 CRT monitor (refresh rate = 80 Hz,
354 resolution = 1280 pixel × 960 pixel, pixel size = 0.31 mm × 0.31 mm). Because of the space
355 limit, the viewing distance and the monitor position varied depending on the stimulus spatial
356 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
357 Monkey C, visual stimuli were generated by Psychotoolbox 3 (Pelli & Zhang, 1991) and
358 presented on a 27'' Acer XB271HU LCD monitor (refresh rate = 80 Hz native, resolution =
359 2560 pixel × 1440 pixel native, pixel size = 0.23 mm × 0.23 mm). The viewing distance was 50
360 cm for lower frequencies (0.25 - 1 cpd) and 100 cm for higher frequencies (2 - 8 cpd). For
361 both monitors, the screen luminance was linearized by an 8-bit look-up table, and the mean
362 luminance was ~47 cd/m^2.
363 A drifting square-wave grating (SF = 4 cpd, contrast = full, speed = 3 cycles/s, starting
364 phase = 0°, and size = 0.4° in diameter) was first used to determine the location, eccentricity
365 (3.4° for Monkey A, 1.7° for Monkey B, and 1.1° for Monkey C) and size (0.8 - 1°) of the

366 population receptive field associated with a recording field of view (FOV), as well as ocular
367 dominance columns when monocularly presented to confirm the V1 location. This fast
368 process used a 4 × objective lens mounted on the two-photon microscope and revealed no
369 cell-specific information.

370 Neuronal responses were then measured with a high-contrast (0.9) Gabor grating
371 (Gaussian-windowed sinusoidal grating) drifting at 2 cycles/s in opposite directions
372 perpendicular to the Gabor orientation. The starting phase of the drafting Gabors was always
373 0°. The Gabor grating varied at 12 orientations from 0° to 165° in 15° steps, and 6 spatial
374 frequencies from 0.25 to 8 cpd in 1-octave steps. In addition, three stimulus sizes (with
375 constant stimulus centers) were used at each spatial frequency, for two purposes. First, our
376 pilot measurements suggested very strong surround suppression with larger stimuli.
377 Therefore, comparing the responses to different stimulus sizes helped approximate the RF
378 size of each neuron that produced maximal response and least surround suppression.
379 Second, for neurons whose RF centers and the stimulus center were misaligned, the RFs of
380 some misaligned neurons might be better covered by larger stimuli. Still the RFs of additional
381 neurons might have less overlap with the stimuli. These neurons would have weaker and less
382 orientation-tuned responses because of the Gaussian-blurred stimulus edge, and would most
383 likely be filtered out during our multiple steps of selection of orientation tuned neurons (see
384 below). Specifically, the σ of the Gaussian envelope of the Gabor were 0.64λ and 0.85λ at all
385 spatial frequencies, and was additionally smaller at 0.42λ when spatial frequencies were
386 0.25 - 1 cpd, and larger at 1.06λ when spatial frequencies were 2 - 8 cpd (λ: wavelength;
387 Gabors with the same σ in wavelength unit had the same number of cycles). Here at the
388 smallest σ (0.42λ), the Gabors still had sufficient number of cycles (frequency bandwidths =
389 1 octave) (Graham, 1989), so that the actual stimulus spatial frequencies were precise at

390 nominal values. In terms of visual angle, σ = 1.68°, 2.56°, and 3.36° at 0.25 cpd; 0.84°, 1.28°,
391 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°,
392 0.21°, and 0.26° at 4 cpd, and 0.08°, 0.11°, and 0.13° at 8 cpd, respectively.

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

414

415 *Two-photon calcium imaging*

416 Two-photon calcium imaging was performed with a Prairie Ultima IV (In Vivo) two-
417 photon microscope (Prairie Technologies) on Monkeys A and B, or a FENTOSmart two-photon
418 microscope (Femtonics) on Monkey C, and a Ti:sapphire laser (Mai Tai eHP, Spectra Physics).
419 GCaMP5 was chosen as the indicator of calcium signals because the fluorescence activities it
420 expresses are linearly proportional to neuronal spike activities within a wide range of firing
421 rates from 10-150 Hz (Li et al., 2017). One FOV of 850 x 850 μm^2 was selected in each animal
422 and imaged using a 1000-nm femtosecond laser under a 16 × objective lens (0.8 N.A., Nikon)
423 at a resolution of 1.6 μm/pixel. Fast resonant scanning mode (32 frames per second) was
424 chosen to obtain continuous images of neuronal activity (8 frames per second after averaging
425 every 4 frames). Recording was first performed at a shallower depth, and some neurons with
426 high brightness or unique dendrite patterns were selected as landmarks. In the next daily
427 session, the same FOV at the same depth was located before recording with the help of
428 landmarks, and the depth plane was lowered if recording was performed at a deeper depth
429 (Monkeys A & B). Because of the time limit, recordings at a specific FOV/depth with
430 monocular and binocular stimulations were completed in 2-3 consecutive daily sessions, but
431 the same neurons could be precisely tracked over multiple recording sessions with the use of
432 multiple landmark cues.

433

434 *Imaging data analysis: Initial screening of ROIs*

435 Data were analyzed with customized MATLAB codes. A normalized cross-correlation
436 based translation algorithm was used to reduce motion artifacts (Li et al., 2017). Then
437 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 (*ΔF = 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.

461

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

463 The ratio of fluorescence change (*ΔF/F0*) was calculated as a neuron's response to a
464 specific stimulus condition. For a specific cell's response to a specific stimulus condition, the
465 $F0_n$ of the n-th trial was the average of 4 frames before stimulus onset, and F_n was the
466 average of 5th-8th frames after stimulus onset. $F0_n$ was then averaged across 12 trials to
467 obtain the baseline F0 for all 12 trials (for the purpose of reducing noises in the calculations
468 of responses), and $\Delta F_n/F0 = (F_n-F0)/F0$ was taken as the neuron's response to this stimulus at
469 this trial. The final response was averaged over 11 trials, minusing the 12th trial that showed
470 the weakest and often negative response. For a small portion of neurons (e.g., ~3% in
471 Monkeys A, ~8% in monkey B and ~2% in Monkey C) showing direction selectivity as their
472 responses to two opposite directions differed significantly ($p < 0.05$, Friedman test), the 6
473 trials at the preferred direction was considered for calculations of $\Delta F_n/F0$ as the cell's
474 responses to a particular stimulus. F0 was still averaged over 12 trials at two opposite
475 directions.

476 Several steps were then taken to decide whether a neuron was tuned to orientation
477 and/or spatial frequency, and if so, its ocular dominance index. For each monocular
478 condition, first the orientation, SF, and size (σ) producing the maximal response among all
479 conditions were selected. Then responses to other 11 orientations and 5 SFs were decided at
480 the selected SF and size. Second, to select orientation and/or SF tuned neurons, a non-
481 parametric Friedman test was performed to test whether a neuron's responses at 12
482 orientations or 6 SFs were significantly different from each other at least under one
483 monocular stimulation. To reduce Type-I errors, the significance level was set at $\alpha = 0.01$.
484 Third, for those showing significant orientation difference, the trial-based orientation

485 responses of each neuron were fitted with a Gaussian model with a MATLAB nonlinear least-
486 squares function: lsqnonlin:

487 $$\mathrm{R}(\theta) = a_1 2^{-(\frac{\theta-\theta_0}{\sigma})^2} + b$$

488 where R(θ) was the response at orientation θ, free parameters a_1, θ_0, σ, and b were the
489 amplitude, peak orientation, standard deviation of the Gaussian function, and minimal
490 response of the neuron, respectively. Only neurons with goodness of fit $R^2 > 0.5$ at least
491 under one stimulation condition were finally selected as orientation tuned neurons. Fourth,
492 for those showing significant SF difference, the trial-based SF responses of each neuron were
493 further fitted with a Difference-of-Gaussian model.

494 $$\mathrm{R}(sf) = a_1 e^{-(\frac{sf}{\sigma_1})^2} - a_2 e^{-\left(\frac{sf}{\sigma_2}\right)^2} + b$$

495 where R(sf) was a neuron's response at spatial frequency sf, free parameters a_1, σ_1,
496 a_2, and σ_2 were amplitudes and standard deviations of two Gaussians, respectively, and b
497 was the minimal response among 6 spatial frequencies. Only those with goodness of fit R^2 >
498 0.5 at least under one monocular stimulation were selected as SF tuned neurons.

499 The ocular dominance index (ODI) was calculated to characterize each orientation
500 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
501 neuron's respective peak responses at the best orientation and SF to ipsilateral and
502 contralateral stimulations on the basis of data fitting. Here ODI = -1 and 1 would indicate
503 complete contralateral and ipsilateral eye preferences, respectively, and ODI = 0 would
504 indicate equal preferences to both eyes.

505

506 *Model fitting*

507 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.

Acknowledgments

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*, 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 e384.
Graham, N. (1989). *Visual Pattern Analyzers (Oxford Psychology Series, No. 16)*. New York: Oxford University Press.
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*, 102120.
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*, e15075.
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 e1043.
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., Carlson, B. M., Westerberg, J. A., Cox, M. A., & Maier, A. (2023). A role for ocular dominance in binocular integration. *Curr Biol, 33*, 3884-3895 e3885.
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*, 104182.
Moradi, F., & Heeger, D. J. (2009). Inter-ocular contrast normalization in human visual cortex. *J Vis, 9*, 13 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.
Poggio, G. F., & Fischer, B. (1977). Binocular interaction and depth sensitivity in striate and prestriate cortex of behaving rhesus monkey. *J Neurophysiol, 40*, 1392-1405.
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.

Shatz, C. J., & Stryker, M. P. (1978). Ocular dominance in layer IV of the cat's visual cortex and the effects of monocular deprivation. *J Physiol, 281*, 267-283.

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