Feature representation under crowding in macaque V1 and V4 neuronal populations

Christopher A Henry[1], **Adam Kohn**[1,2,3]

[1]Dominick P. Purpura Department of Neuroscience, Albert Einstein College of Medicine, Bronx, NY 10461, USA

[2]Department of Ophthalmology and Visual Sciences, Albert Einstein College of Medicine, Bronx, NY 10461, USA

[3]Department of Systems and Computational Biology, Albert Einstein College of Medicine, Bronx, NY 10461, USA

SUMMARY

Visual perception depends strongly on spatial context. A profound example is visual crowding, whereby the presence of nearby stimuli impairs discriminability of object features. Despite extensive work on perceptual crowding and the spatial integrative properties of visual cortical neurons, the link between these two aspects of visual processing remains unclear. To understand better the neural basis of crowding, we recorded simultaneously from neuronal populations in V1 and V4 of fixating macaque monkeys. We assessed the information available from the measured responses about the orientation of a visual target, both for targets presented in isolation and amid distractors. Both single neuron and population responses had less information about target orientation when distractors were present. Information loss was moderate in V1 and more substantial in V4. Information loss could be traced to systematic divisive and additive changes in neuronal tuning. Additive and multiplicative changes in tuning were more severe in V4; in addition, tuning exhibited other, non-affine transformations that were greater in V4, further restricting the ability of a fixed sensory readout strategy to extract accurate feature information across displays. Our results provide a direct test of crowding effects at different stages of the visual hierarchy. They reveal how crowded visual environments alter the spiking activity of cortical populations by which sensory stimuli are encoded and connect these changes to established mechanisms of neuronal spatial integration.

eTOC Blurb

Corresponding authors: Christopher Henry, christopher.aaron.henry@gmail.com; Adam Kohn (lead contact), adam.kohn@einsteinmed.edu, 1410 Pelham Parkway S., Rm 822, Bronx NY 10461, Tel: 718-430-2417, Fax: 718-430-8821.

Henry and Kohn investigate how visual crowding impairs the encoding of target orientation in macaque V1 and V4 neuronal populations. Distractor stimuli induce systematic additive and multiplicative changes in tuning, leading to moderate information loss for target discriminability in V1 and more substantial loss in V4.

INTRODUCTION

Crowding is a perceptual phenomenon in which a clearly visible stimulus (a target) becomes difficult to discern when other stimuli (distractors) are presented in spatial proximity[1-4]. Crowding has been studied extensively because it often limits visual function and because understanding its basis may provide insight into core aspects of visual processing.

Impairment of discriminability under crowding involves features from nearby stimuli being 'combined' with those of a target. The challenge for mechanistic descriptions of crowding is to identify the form of this 'combination' at distinct stages of visual processing. There are numerous proposals of how crowding might arise, each invoking different computations and proposing different loci in the visual system. These proposals, some of which have conceptual overlap, include: lateral interactions between neurons within a cortical area that represent targets and distractors[5-6]; spatial pooling by neurons whose receptive fields encompass both the target and distractors[1]; an encoding strategy in which visual stimuli are represented using a set of summary statistics[7-8] or simply averaged together[9-10]; or an erroneous readout of sensory information, involving either compulsory averaging or the substitution of signals encoding distractors for those encoding a target (e.g. [11]), perhaps arising from fluctuations in attention[12]. These various bases of crowding have been proposed to occur in primary visual cortex (V1)[5-6], V2[8, 13], V4[1], or a combination of these[14].

To date, how crowding affects neural representations has been studied almost exclusively using coarse or indirect measures of brain activity, such as EEG or functional MRI[15-20]. These approaches provide important information, but they lack the resolution to elucidate how crowding affects the representation of visual information by neuronal spiking activity, the activity that encodes and transmits visual information in the brain. As a result, we have a limited understanding of how crowding arises.

We recently assessed the degree to which crowded visual displays limit neuronal population information about the orientation of a target stimulus, in the primary visual cortex (V1) of anesthetized macaque monkeys[21]. We found that distractors reduced information through a combination of divisive, suppressive modulation of responsivity in some neurons and additive facilitation in others. The information loss in V1 could account for much, but not all, of the perceptual crowding evident in humans and monkeys performing a discrimination task using the same stimuli. Thus, we proposed that spatial contextual modulation in V1 is a major contributor to crowding.

How crowded displays influence stimulus encoding in midlevel visual cortex is unknown. The more extensive spatial pooling performed by neurons in higher visual cortex[22-23] suggests additional mechanisms by which sensory information may be lost, though precisely how those neurons pool visual inputs across space is not well understood[24-25]. Here we

assess how information about target orientation is affected by the presence of distractors in neuronal populations in area V4 of fixating macaque monkeys. We find that V4 populations show greater information loss than in simultaneously recorded V1 populations, under identical stimulus conditions. Further, this loss is evident for a greater range of spatial offsets between targets and distractors. Information loss in V4 involves some of the same mechanisms seen in V1, but V4 neurons also show a complex dependence on the precise set of stimuli presented in or near their spatial receptive fields.

RESULTS

We carried out simultaneous extracellular recordings from V1 and V4 neuronal populations via chronic microelectrode arrays implanted in two awake macaque monkeys (Figure 1A). Animals viewed stimuli passively, while maintaining fixation within a 1.3 degree diameter window about a small central point.

We first obtained maps of the spatial receptive fields of the sampled neuronal populations in dedicated recording sessions. As expected, individual V1 receptive fields were small, and the receptive fields of V4 neurons were many-fold larger (Figure 1B, top). Aggregate receptive field maps showed substantial overlap in the visual field representation of recorded V1 and V4 populations in each animal (Figure 1B, bottom).

To assess feature encoding, we used a small target grating (1 degree diameter), positioned to drive both cortical populations. The target grating was located at 2.8 degrees eccentricity in the lower visual field of one animal, and at 3.6 degrees in the other (indicated by the black circle, Figure 1B). On each trial, stimuli were shown in two epochs separated by 0.35 s (Figure 1C). Target stimuli were chosen from an 18 degree range straddling a reference orientation (changed in each session), in increments of 4-6 degrees (Figure 1D). On each presentation, the target stimulus could be shown either in isolation or surrounded by 4 distractor stimuli, whose orientation was ±10 degrees of the reference orientation. The size and relative orientations of target and distractor stimuli were chosen to produce robust perceptual crowding effects, as shown previously[21]. The spacing of distractors around the target was varied because target-distractor distance is known to influence perceptual crowding[26]. Identical stimuli, save for a rotation in reference orientation by 90 degrees, were shown on half the presentations, to minimize the effects of adaptation[27]. Thus, each recording session probed cortical responses to two narrow ranges of target orientation, yielding distinct responses from the recorded neuronal populations that were analyzed independently.

We recorded in 11 sessions, yielding 22 neuronal populations in each area. The size of V1 populations was on average 19 ± 1.9 neurons, and the size of V4 populations was 31 ± 3.4 neurons.

Distractors limit target discriminability

We first assessed how the presentation of distractors affected the discriminability of target orientation afforded by responses of individual V1 (n = 396) and V4 neurons (n = 691). To quantify discriminability, we used receiver operating characteristic (ROC) analysis to

distinguish between the spike counts evoked by two target gratings differing 18 degrees in orientation (Figure 2A, abscissa; see Figure S1 for sample response distributions). Chance performance is indicated by a value of 0.5 in the area under the ROC curve. We assumed the orientation closer to 0 would evoke stronger responses, so better performance is indicated by the deviation from 0.5, with values of 0 or 1 both indicating perfect discriminability between stimuli (for neurons with opposing tuning slopes over the sampled target range). Note that performance for many neurons was near chance, because the narrow range of target orientations did not include those that modulated the neuron's responses significantly (i.e. near the flank of the tuning function).

We compared single neuron discriminability for targets alone and targets with distractors. For V1 neurons, discriminability was better when targets were presented alone (Figure 2A), evident as a regression slope that was significantly less than 1 (slope = 0.60, 95% confidence interval: [0.50 0.69]). When distractors were moved further from the target, effects on discriminability were slightly more moderate (Figure 2C, farthest spacing, slope = 0.74, [0.66 0.82]). In V4, the loss of discriminability with nearby distractors was stronger (Figure 2B, slope = 0.20, [0.14 0.26]) and remained pronounced even when distractors were placed farther from the target (Figure 2C, farthest spacing, slope = 0.44, [0.37 0.50]). Thus, distractors led to a greater impairment in single neuron discriminability of targets in V4 than V1, over all spatial scales.

We next quantified how distractors affected neuronal population information about target orientation. We used logistic regression analysis to determine how well two target stimuli could be distinguished using the measured neuronal population responses. We optimized decoders separately for each pairing of target stimuli, when targets were presented alone or with distractors. By allowing for separate decoding of these sets of responses, we provide an upper-bound estimate on the target information accessible to a linear readout strategy of the measured responses.

To provide good estimates of how distractors affect representational accuracy, we aggregated decoding performance across target grating pairings and sessions from each animal, for each cortical area and target-distractor spacing. An example is shown in Figure 2D,E for V1 and V4 of one animal, comparing performance for responses to targets alone and to targets with the nearest distractors. In V4, when decoder performance was above chance (Figure 2E, abscissa values greater than 0.5), performance for responses to targets with distractors tended to be lower (ordinate; p<0.001 for difference; addressed further below). In V1, performance loss was less evident (p>0.05 for difference; addressed further below), due in part to the low discriminability for both targets alone and targets with distractors.

To estimate how the change in decoder performance might translate to the strength of perceptual crowding, we fit a single parameter model to the decoder performance values aggregated across sessions. The model consisted of a psychometric function for target discrimination whose slope could change with the appearance of distractors. The single parameter provided an estimate of the discrimination threshold elevation that would be consistent with the change in population performance (see Figure S2). Sample predicted

performance trends for different changes in threshold are shown in Figures 2D,E (gray curves).

In V1, distractors led to a modest loss of population information about target orientation whose magnitude was consistent with a moderate elevation in discrimination threshold (Figure 2F, blue). Distractors at the closest spacing yielded the largest threshold elevations, 1.31 ± 0.10; distractors at the farthest spacing yielded the smallest elevation, 1.15 ± 0.08. The estimated threshold elevation was statistically significant for all but the farthest target-distractor separation (permutation test, $p<0.05$). In V4 populations, the estimated threshold elevations were greater than those in V1 (Figure 2F, red): 2.52 ± 0.23 for the closest distractors and 1.89 ± 0.19 for the farthest distractors (significant for all distractor separations, permutation test, $p<0.001$).

We were concerned that our estimates of threshold elevation might be poorly constrained by the limited decoding performance values which reflected the small size of the recorded populations, particularly in V1. We thus repeated our analysis using responses of larger 'pseudo-populations', constructed by combining neurons across sessions. Pseudo-populations provided a broader range of performance values and yielded similar estimates of how distractors affected thresholds (Figure S3).

In summary, we found that distractors reduced population information about target orientation in both V1 and V4. In V1, there was a moderate threshold elevation for the closest distractors and little effect for distractors placed farther away. In V4, there was a greater threshold elevation, evident across a wider range of distractor spacings. We next turn to describing which changes in neuronal activity contribute to the loss of encoded feature information.

Distractors modulate neuronal responsivity and variability

The stimulus information encoded by neuronal populations (as extracted by linear decoders) is a function of the tuning and response variability of the individual neurons and of how their variability is shared[28-29]. All of these response statistics can be altered by changes in spatial context, such as the presentation of distractors [21,30-32]. To understand why V1 and V4 neuronal populations provide less information about target orientation in crowded displays, we assessed how distractors affected pairwise correlations (shared variability), individual neuronal variability, and neuronal responsivity, in turn.

We first determined whether the presence of distractors altered the strength of shared variability among neurons. For each pair of neurons, we calculated the strength of 'noise correlations' as the average of the correlations calculated for each stimulus separately. Distractors significantly reduced correlations compared to those evident in responses to targets alone. In V1 (Figure 3A; see Figure S1 for sample responses), pairwise correlations decreased from 0.07 ± 0.003 for targets alone to 0.05 ± 0.002 with the nearest distractors (paired t-test, $p<0.001$). In V4 (Figure 3B), the corresponding decrease in correlations was from 0.11 ± 0.001 to 0.06 ± 0.001 (paired t-test, $p<0.001$). To determine whether altered correlations contributed to information loss, we fit decoders to population responses after trial shuffling responses to reduce correlations. In these shuffled data, distractors led

to threshold elevations similar to those seen in the original data (Figure 3C), indicating changes in correlations, though substantial, played a minimal role in the information loss we observed.

To quantify how response variability was affected by distractors, we measured the response variance normalized by its mean (Fano factor). The Fano factor for each neuron was calculated as the geometric mean of the Fano factors measured for each stimulus condition (combination of target and distractor orientations) separately. In V1, distractors at the closest spacing led to a slight reduction in variability (Figure 3D, from 1.05 ± 0.02 for targets alone to 0.98 ± 0.01, paired t-test, p<0.001). In V4, distractors at all spacings led to a significant reduction in variability (Figure 3E, 1.00 ± 0.01 for targets alone to 0.93 ± 0.01 with closest distractors, paired t-test, p<0.001 for each spacing). All else being equal, the reduction in variability with distractors should result in better discriminability not worse. Thus, changes in variability must not be responsible for information loss we observe.

We next asked how distractors influenced neuronal responsivity to targets. We calculated a modulation index defined as the ratio of the response to targets with distractors relative to that for targets alone. For each neuron, we summarized this index across targets as the geometric mean of the indices computed for each target orientation separately. In V1, neurons exhibited diversity of modulation, with both suppression and facilitation evident. For distractors at the closest spacing, suppression and facilitation occurred in roughly equal proportion (Figure 4A, dark blue). To assess whether the modulation indices were meaningful, we determined the probability observing each measured modulation index by chance. Suppressive effects were statistically significant in 23% of neurons (permutation test, p<0.05); facilitation was significant in 27% cases. Whether distractors suppressed or facilitated responses was related to the alignment of the receptive field with the target stimulus (Figure S4). For distractors placed further from the target (Figure 4A, lighter blues), modulation was weaker; only 2.5% cases of significant facilitation were observed (the expected false positive rate) and 13% cases of suppression.

The responses of V4 neurons to targets could also be facilitated or suppressed by the presentation of distractors. For the closest distractors, most V4 neurons showed facilitation (Figure 4B, dark red points, 1.11 ± 0.38, t-test, p<0.001; 51% of cases of facilitation were statistically significant and 14% of suppressive cases) and the average modulation index was larger than in V1 (one sided t-test, p<0.01). As distractor spacing increased, a roughly equal proportion of V4 neurons exhibited facilitation and suppression (Figure 4B, light red points, 0.97 ± 0.48; 24% and 32% cases involved significant facilitation and suppression, respectively).

To quantify the strength of modulation by distractors, regardless of whether it facilitated or suppressed responsivity, we averaged the absolute value of each neuron's modulation index (log-transformed). The modulation in both areas was statistically significant, for all target-distractor spacings (permutation test, p<0.01). The average modulation was stronger in V4 than in V1 at all separations (Figure 4C, one sided t-test, p=0.01 for closest distractors and p<0.001 for all other spacings). Further, modulation strength in V4 was largely maintained across target-distractor spacings (p=0.44, one sided t-test comparing nearest to farthest

distractors) whereas the farthest separation yielded weaker modulation in V1 than the nearest separation (one sided t-test, $p < 0.001$).

To understand how response modulation by distractors influences the discriminability of target orientation, we must consider the functional form of changes in orientation tuning induced by distractors. We therefore fit a model in which each neuron's target tuning amid distractors was described by either a multiplicative scaling and additive translation of the target alone tuning, or both (Figure 5A, shaded regions). The additive term was normalized, so that it reflected the proportion of the mean firing rate for targets alone that was needed to produce the appropriate translation in tuning. In this analysis, we only considered neurons that had reliable tuning over the narrow range of sampled orientations (see Methods).

We first considered a version of the model in which either multiplicative scaling or additive translation could occur. In prior work[21], we had found that V1 neurons whose responses were suppressed were better described by the multiplicative model whereas those that were facilitated by distractors were better described by an additive model. In the V1 and V4 responses reported here, we observed a similar outcome (Figure S5). We next determined how strongly these forms of modulation alter the encoding of target orientation by single neurons. Both forms of modulation reduce information[21] divisive suppression makes responses to different stimuli more similar; additive facilitation results in greater response variability, since variability in cortical responses scales with the mean. Though additive and multiplicative modulation did reduce information about target orientation, these effects—considered in isolation—could not fully capture the loss of discriminability we observed (Figure 2).

We therefore considered a model in which both multiplicative and additive effects could occur. Figure 5 shows the fitted parameters for this model, for V1 (Figure 5B; see Figure S1 for sample tuning functions) and V4 (Figure 5C) neurons, with each point representing a neuron and distractor spacing condition, and darker points representing closer distractors. For V1 neurons (Figure 5B), the amplitude of the multiplicative term was significantly less than 1 on average (abscissa, 0.59 ± 0.03), indicating a compression of the tuning curve, and the additive term was positive (0.39 ± 0.16), indicating an upward translation of the tuning. For V4 neurons (Figure 5C), the amplitude of the multiplicative term was also less than 1 (0.36 ± 0.01) and the additive term was positive (0.90 ± 0.02). The multiplicative term for V4 neurons was significantly smaller than for V1 neurons (t-test, $p < 0.001$) and the additive term was significantly larger for V4 than V1 (t-test, $p > 0.001$).

In both V1 and V4, there was a strong covariation between the additive and multiplicative terms, consistent with a flattening of the tuning function with maintained responsivity. For instance, for tuning that becomes flat with distractors, the multiplicative term would be 0 and the additive term would be positive to reproduce the mean firing rate. For tuning that was only partially flattened, the multiplicative term would be more positive (but less than 1), so a smaller additive term would be required to sum to reproduce the tuning. In summary, the presence of distractors compressed the dynamic range of neuronal tuning in both areas, via a multiplicative scaling, and maintained or enhanced responsivity, via an additive offset.

To test whether these two effects together capture the information loss we observe in V1 and V4, we simulated crowded responses by scaling and translating the target alone tuning with the fitted multiplicative and additive terms (see Methods). We then calculated the discriminability for targets alone and for targets with distractors, using the area under the ROC curve (as in Figure 2). This simulation yielded some loss of discriminability in V1 (Figure 5D) and more substantial effects in V4 (Figure 5E); the effects in simulations were similar but slightly stronger than the losses observed in the measured responses (Figure 2). In both areas, the loss of discriminability depended weakly on target-distractor spacing, as in the data. Together, these results demonstrate that divisive scaling and additive translation of target tuning in the presence of distractors limit the encoding of target orientation, and that additional impairment in V4 results from greater modulation by these factors compared to V1.

We note that the model fitting revealed cases in which the multiplicative term was negative; as a result, a large additive term was required to offset the negative responses and reproduce the firing rate. These cases occurred rarely in V1 but more often in V4 and correspond to changes in the slope of the tuning function. These striking changes in tuning are considered next.

Distractors alter neuronal tuning for target orientation

Both additive facilitation and divisive suppression of neuronal responsivity by distractors are consistent with well-established mechanisms of spatial integration by cortical neurons (see Discussion)[33-35]. Yet, previous work has also suggested more complex forms of spatial integration, particularly in V4. For instance, the tuning of V4 neurons for a small stimulus in the receptive field can change dramatically, and often seemingly idiosyncratically, when paired with other stimuli (e.g. [24-25]). Consistent with this description, we often observed that V4 tuning for target orientation could be altered substantially when distractors were present (Figure 6B). V1 tuning for targets, in contrast, was usually (Figure 6A, left) but not always (right) scaled or translated by distractors.

To quantify the similarity of tuning for targets alone and targets with distractors, we calculated their Pearson correlation. If target tuning is simply scaled or translated by distractors, correlation values should be high. If distractors distort tuning, correlations should be lower. To provide a fair comparison of V1 and V4 tuning, we selected subsets of neurons in each area that exceeded a criterion that captured the reliability of tuning for targets alone, defined by calculating the correlation between the tuning from even versus odd trials. In V1, tuning for targets alone was well correlated with the tuning for targets with nearby distractors (Figure 6C, blue points). For instance, for cells exceeding a criterion of 0.5, the correlation between tuning with and without distractors was 0.38 ± 0.03. Tuning correlation was significantly lower in V4 neurons (Figure 6C, red points), with the corresponding value being 0.22 ± 0.03 (p<0.001 for comparison with V1, one-sided t-test).

We next assessed whether the change in tuning depended on the particular set of distractor gratings. We computed the correlation of target tuning for displays with different target-distractor spacings (near and far). In this comparison too, correlations between target tuning were lower in V4 than V1 (Figure 6D, for cutoff of 0.5, correlations were 0.17 ± 0.02 and

0.36 ± 0.02, respectively, p<0.001, one-sided t-test). Even subtle changes in the composition of distractors could strongly alter tuning for the target. Specifically, we compared target tuning in the presence of two distractor configurations, for which the position of each distractor was fixed but its orientation differed by 20 degrees between configurations (with the same average orientation within each configuration). V4 tuning differed substantially between these two conditions, whereas V1 tuning was nearly unaffected (Figure 6E, for cutoff of 0.5, V4 correlation values were 0.24 ± 0.06 and V1 correlation values were 0.65 ± 0.06, p<0.001 for difference between areas, one-sided t-test). Together, these results suggest that distractors introduce configuration-dependent changes in neuronal tuning for target orientation, particularly in V4.

The change in V4 tuning with distractors need not result in a loss of population information about target orientation, because changes in tuning could be compensated for by adjusting the readout strategy. The population decoding we have considered thus far had this flexibility, as it involved fitting separate decoders to the activity evoked by each display. However, circuits reading out V1 and V4 responses are unlikely to flexibly alter their strategy, to optimally decode responses for each stimulus condition. For a fixed decoding strategy, changes in tuning with distractors, as described above, would be expected to limit discriminability further. To assess this possibility, we projected population responses onto two fixed decoding axes: one that was optimized for responses to targets alone and one optimized for targets with nearby distractors.

Figure 7A shows example population responses from a single V1 session projected onto these two decoding axes, with ellipses indicating response distributions for each of the 4 target orientations. In V1, projection of responses from either condition yielded similar separation of points on either decoding axis, indicating the information about target orientation can be decoded similarly well by a fixed decoder optimized either to responses to targets alone or to targets with distractors. In V4 (Figure 7B), projections were less aligned between the two decoding strategies, and the separation of points was reduced along the decoding axis of the opposing condition. Thus, in V4, projections depend heavily upon the decoding axis used, and decoding performance would be expected to generalize poorly between configurations.

To quantify how well a common decoding strategy worked across configurations, we calculated the dprime value between response distributions for each pair of target stimuli, when projected onto a decoding axis defined by a matching display condition (same projection) or the opposite configuration (cross projection). Discriminability between stimuli was high in V1 and V4 when using the matched decoder ('same', Figures 7C,D). When data were projected onto the decoding axis of the opposing condition ('cross'), discriminability was largely preserved in V1 (dprime of 0.95 ± 0.13 for same projection and 0.88 ± 0.12 for cross projection, p=0.20, paired t-test). In contrast, discriminability dropped significantly for V4 (1.02 ± 0.13 vs. 0.67 ± 0.05, p=0.002, paired t-test). In additional analyses, we confirmed that the distinct decoding axes in V4 for different displays do not arise from relying on distinct subsets of neurons. (Figure S6). Thus, changes in V4 representation across displays do not result from dramatic changes in which neurons were relevant for encoding target

orientation, but rather from changes in the tuning of a common underlying pool of neurons between conditions.

In summary, V4 populations exhibit a more configuration-dependent representation of target orientation than V1 populations. Thus, not only does V4 show a greater loss of stimulus information for optimal decoders (Figure 2), but it also shows greater loss for a fixed linear readout of V4 responses for different displays (Figure 7).

DISCUSSION

We evaluated how neuronal population information about target orientation in areas V1 and V4 is affected by the presence of distractors. V1[5,6,17-19] and V4[1,25] have each been suggested as a possible locus for the bottleneck limiting perception in crowded scenes. We found that distractors cause a moderate loss of information about target orientation in V1 and a more pronounced loss in V4. V4 effects extended over a larger range of target-distractor separations than V1 effects. Distractors affected the responsivity, tuning, and variability of neurons in both areas. Responses were diversely modulated, but modulation often took the form of a divisive suppression and additive facilitation of responsivity, whose combined influence could explain the reduction in discriminability with distractors. Our results provide a direct test of crowding effects at different stages of the visual hierarchy using neuronal population spiking activity and connect changes in representational quality to established mechanisms of neuronal spatial integration.

The modulation of target responses by distractors is consistent with prior descriptions of how V1 and V4 neurons integrate visual signals spatially. Response suppression in V1 is expected when distractors fall in the receptive field surround, and this suppression acts divisively[35,36]. The neural mechanisms underlying surround suppression likely involve feedforward, lateral, and feedback components[37]. When targets and distractors fall within the receptive field, response summation—facilitation—would be expected[33,36]. Summation within the receptive field is additive, but modulated by normalization mechanisms (which act divisively)[34,35]. The combination of additive facilitation and divisive suppression would be expected to yield a 'flattened' tuning curve, as we often observed.

In V4, neuronal receptive fields are a few-fold larger than in V1, at matched eccentricity[23], and thus more likely to span multiple elements of crowded displays. Within the classical receptive field, V4 neurons have been shown to sum multiple visual inputs[38], as well as to show various forms of normalization[39,40]. In addition, V4 neurons possess a suppressive surround[41,42] that would be expected to be engaged by the distractors in our displays, particularly when they are far from the target.

We have previously shown that discriminability in neurons with Poisson-like variability can be reduced if distractors recruit divisive suppression or additive facilitation of tuning[21,43]. We also reported that, in V1, suppressed neurons were better explained by a divisive modulation, and facilitated neurons better explained by an additive modulation. We replicated those findings for the V1 and V4 responses in the present study and confirmed that these forms of modulation impaired discriminability (Figure S5). However,

the information loss expected from each of these mechanisms in isolation was insufficient to account for the effects observed in the measured responses. We therefore considered a model in which additive facilitation and divisive suppression could co-occur, which revealed most neurons in V1 and V4 are better described as having both multiplicative and additive changes in target tuning amid distractors. These combined additive and multiplicative changes are required to describe the strong information loss we observe, particularly in V4.

The combined influence of these two mechanisms was less evident in our prior V1 study because most of those cases involved distractors which fell entirely in the surround, recruiting only response suppression. In the recordings reported here, response facilitation was more evident. In V4, this occurred because of the larger receptive fields; in V1, facilitation was more common because of the placement of the target and distractor stimuli relative to the receptive field (Figure S4). Importantly, both the mechanisms reported previously and those brought to light here, yield similar conclusions concerning a moderate information loss in V1 for target orientation.

Nearby distractors affect both perceptual detectability and discriminability (i.e., crowding) of a target stimulus[1,4]. Both impairments depend on the configuration of stimuli but differ in the form of that dependency, suggesting the two recruit distinct neural mechanisms[44]: Impaired detectability has been associated with neural surround suppression[45]; crowding has often been linked to summation within the receptive field[1-4]. We have previously shown that neuronal surround suppression can also reduce information about the features of suprathreshold stimuli (i.e., discriminability)[21], as also observed here (see also [46]). In addition, we show here that how signals are summed within the receptive field, particularly in V4, does indeed contribute strongly to information loss. Importantly, we provide a quantitative description of the form of that summation and why it reduces discriminability.

Beyond these general forms of response modulation under crowding, we showed that V4 tuning for targets is more sensitive to the stimulus configuration than that in V1. V4 neurons have been shown to be selective either for elongated contours[47] or for the curvature of an object within their receptive fields[48-51]. Additionally, V4 responses depend on the joint stimulus statistics shown within their receptive fields, as in studies of selectivity for stimulus textures[52,53]. Together, these studies suggest V4 neurons encode the relationship between spatially adjacent image features, to extract mid-level representations of object shape and surfaces[54-58]. While each of these studies probes distinct forms of selectivity, together they reinforce the notion that V4 tuning might differ strongly for a small target stimulus within the receptive field when it appears in isolation or amid other nearby stimuli. Indeed, a recent study[25] showed that V4 neuronal responses to letter stimuli were strongly altered when the arrangement of the constituent line segments was changed. To the extent that downstream areas implement a fixed readout of V4 activity across displays, this context-dependent tuning will limit further the resolution of target features available for perception.

Our experiments involved targets and distractors of a fixed size. Would our results have been different if the displays were scaled down so that, for instance, multiple image features fell within individual V1 receptive fields? Crowding effects at a given eccentricity scale

with the spacing between target and distractors[26]; when stimuli are small relative to V1 receptive field sizes, discrimination performance can be strongly impaired[9]. It remains unclear whether information loss in such displays would be stronger in early cortical areas, or instead remain more prominent in areas like V4. In contrast, if displays were scaled up such that target stimuli were similar in size to the receptive fields of V4 neurons, we would expect the response changes we observed in V1 and V4 to be greatly diminished, consistent with weaker perceptual crowding effects from distractors at increased spatial separation from the target[26].

Previous work has suggested that crowding occurs in V1, V2, or V4, or that it arises from interactions across the visual hierarchy[14]. What do our results imply for where the bottleneck responsible for crowding resides? Using an optimal linear readout of neuronal population activity, we found moderate information loss in V1 and more pronounced loss in V4. We have previously shown[21] that perceptual crowding in our displays, in humans and monkeys, produces a threshold elevation of roughly 50%, slightly greater than that evident in our V1 data and weaker than that evident in V4. Additionally, perceptual crowding occurs when the spacing between targets and distractors is within roughly half the target eccentricity[26]; we found substantial information loss in V4 for distractors placed farther away than this distance (2.6 deg for target eccentricity of roughly 3 degrees).

One interpretation of our results might be that crowding arises at some stage between V1 and V4, such as V2. However, discriminability is a function of both population information and the readout strategy. Observers have been shown to employ suboptimal readout strategies[59,60]. Deploying slightly suboptimal strategies for reading out responses could easily lead to stronger threshold elevation in V1 or V4 (or weaker if the readout of responses to targets alone was more suboptimal). Indeed, many aspects of crowding may be attributable to mechanisms of sensory readout. Attention[12], substitution of distractor features for that of the target[11], and perceptual averaging of nearby image features[9] may all reflect sensory readout strategies that are either fixed and suboptimal or fluctuate trial-to-trial.

Together, this suggests a conceptual framework whereby crowding does not reflect a unitary bottleneck on spatial feature perception at a defined locus within the visual system, but instead the result of a series of diverse transformations in visual processing. Visual signals are encoded in cortical populations, then repeatedly re-coded in novel representational formats downstream. Those representations are affected by mechanisms of spatial integration both within each neuron's receptive field and from the surround. Finally, these representations are interpreted through the lens of a readout strategy to determine perception in cluttered, natural environments. The recoding of visual information, the mechanisms of spatial integration, and the readout strategy can all contribute to the loss of sensory information under crowding.

STAR METHODS

RESOURCE AVAILABILITY

Lead contact—Further information and request for resources should be directed to and will be fulfilled by the lead contact, Adam Kohn (adam.kohn@einsteinmed.edu).

Materials availability—This study did not generate unique reagents.

Data and code availability—Recordings reported in this study have been deposited in the CRCNS data sharing web site [identifier to be added in proofs].

Any additional information required to reanalyze the data in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

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

METHOD DETAILS

Task and visual display—Animals were first implanted with a titanium headpost to allow for head restraint. Surgery was performed under isoflurane anesthesia, following strict sterile procedures. Antibiotics (ceftiofur or enrofloxacin) and an analgesic (buprenorphine) were provided postoperatively. Animals recovered for a minimum of 6 weeks before behavioral training began.

Animals were trained to enter a primate chair (Crist Instruments) and the head was stabilized via the implanted headpost. Inside a recording booth, animals viewed a CRT monitor (Iiyama; 100 Hz refresh rate; 1024 x 768 screen resolution; 40 cd/m^2 mean luminance) with linearized output luminance, from a distance of 64 cm. Visual stimuli were generated using custom software based on OpenGL libraries (Expo). Eye position was monitored using a video-based eye-tracking system (SR Research) running at 1 kHz sampling rate.

To initiate a trial, animals had to direct their gaze to a small central fixation point (0.2-degree diameter) and maintain fixation within a 1.3-degree diameter window around that point for the duration of the trial (eye positions were typically < 0.5 degrees from the fixation point). 0.2 s after fixation onset, a visual stimulus was presented in the lower right visual field for 0.25 s. Stimulus offset was followed by 0.35 s of a blank screen (mean grey), and then a second stimulus shown for 0.25 s. If visual fixation was maintained for an additional 0.15 s after the last stimulus offset, animals were given a small liquid reward. Eye movements outside the fixation window led to an aborted trial and no reward; these trials were not analyzed. A typical session consisted of 1000-1300 successful trials.

Neurophysiological recording—Once fixation training yielded stable behavioral performance, we performed a second sterile surgery, making a craniotomy and durotomy in the left hemisphere of each animal and implanting microelectrode arrays (Blackrock Microsystems) into V1 and V4. Array placement was targeted to yield overlapping retinotopic locations in the two areas, based on anatomical landmarks and prior physiological studies[22-23]. Each electrode array had either a 10 x 10 or a 6 x 8 configuration. In one animal, we implanted 1 array in V4 (48 channel) and 2 in V1 (48 and 96 channel). In the other, we implanted 2 arrays in V4 (both 48 channel) and 2 in V1 (both 96 channel).

After insertion, we sutured the dura over the arrays and covered it with a matrix to promote dural repair (DuraGen; Integra LifeSciences). The craniotomy was covered with titanium mesh.

Extracellular voltage signals from the arrays were amplified and recorded using either Cerebus (Blackrock Microsystems) or Grapevine Scout (Ripple Neuro) data acquisition systems. Signals were bandpass filtered between 250 Hz and 7.5 kHz, and those exceeding a user-defined threshold were digitized at 30 kHz and sorted offline (Plexon Offline Sorter). The local field potential on each electrode was also recorded but not analyzed in this study. Criteria for sorting of individual units included a stable spike waveform and no spike within the neuronal refractory period. We confirmed that the changes in neuronal tuning shown in Figure 6 were evident when our analyses were limited to the most well-isolated neurons, as defined by those units having a waveform SNR in the top 33th percentile[61].

Visual stimuli—In initial sessions, we mapped the neuronal receptive fields. We presented small grating stimuli on a 9 x 9 degree grid of locations (sampled at 1 degree resolution; each stimulus shown at 4 orientations separated by 45 degrees, luminance contrast 50 or 100%, stimulus size 0.5-1.0 degrees). Using the resultant maps, we then placed the target stimuli to drive both the V1 and V4 populations. We also conducted a brief, coarse mapping of each neuron's receptive field at the beginning of each recording session.

Target stimuli were drifting sinusoidal gratings (spatial frequency of 1.5 cycles per degree, temporal frequency of 4 Hz, 1 degree diameter) of 50% luminance contrast. For each session, a reference orientation was chosen, and target orientations were sampled from an 18-degree range spanning that reference. On a subset of trials, 4 distractor stimuli (of the same size and spatiotemporal frequency as the target) were shown adjacent to the target stimulus. The orientation of the distractors was fixed at +/−10 degrees from the reference, and tilt was spatially alternated clockwise around the target. All stimuli had simultaneous onset and offset, and started their drift from a fixed phase. Across trials, the center-to-center spacing of distractors from the target was varied from 1.3 to 2.6 degrees. A second set of stimuli, identical to the first except for a 90-degree rotation of the ensemble reference orientation, was included in each session, to mitigate adaptation effects from repeated presentation of one range of orientations. All stimuli (targets alone, targets with distractors, and changes in distractor spacing and reference orientation) were randomly interleaved. In a separate set of experiments (N = 4 sessions, reported only in Figure 6E), we fixed the spacing of the distractor stimuli but instead varied their orientation, drawing from two distractor arrays with counterbalanced orientations.

QUANTIFICATION AND STATISTICAL ANALYSIS

Analysis was restricted to neurons that were driven by the target stimuli in isolation (responses   1 s.d. above spontaneous rate). All analyses were performed on the total spike count of each neuron during stimulus presentation on each trial (from 0-0.25 s after stimulus onset) using analysis code in Matlab 2014a (Mathworks).

Discriminability of target orientation in single neuron responses was quantified by applying receiver operating characteristic analysis[62] to the spike count distributions evoked from

presentations of the two most extreme target orientations shown (18-degree separation). We used the area under the ROC curve as the measure of discriminability. We then used the slope of the regression line relating discriminability for targets alone to discriminability in the presence of distractors to quantify the average change in discriminability across neurons. 95% confidence intervals for the slope were estimated using bootstrap with replacement across neurons.

Target orientation discriminability in neuronal populations was quantified using logistic regression, with Lasso regularization. The regularization parameter was set to the largest value that yielded maximal cross-validated performance on held out trials from the training set. For each pairwise combination of target orientations shown (and each distractor spacing), we fit a decoding model and evaluated performance on held out test data. The reported decoding performance is an average over cross-validation folds.

Crowding effect size for a given target-distractor separation was estimated by combining decoding performance values for target alone and crowded conditions across pairings of target stimuli and across recording sessions. We fit a single parameter model to the combined data to characterize the threshold elevation under crowding that best explained the change from target alone to crowded performance, using maximum likelihood estimation (validation on synthetic data is shown in Figure S2). Specifically, a target only discrimination model consisted of a cumulative normal distribution adjusted for chance performance in the 2AFC task. For each population and target only performance, we found the stimulus that yielded this performance by inverting the discrimination model. Then, for a given model threshold elevation under crowding, we could generate a predicted discrimination performance. This predicted model performance, assuming a binomial distribution, yielded a probability of the observed discrimination performance under crowding given this model of threshold elevation. For each model, we combined data points by taking the product of likelihoods, and threshold elevation was estimated to be the model with the maximum likelihood. We used a nonparametric bootstrap procedure (sampling from the data points with replacement) to generate confidence intervals for these estimates.

Spike count correlations (r_{SC}) were computed using trials in which the responses of both neurons were within 3 s.d. of their mean responses, to avoid contamination by outlier responses[63]. Correlation values were calculated separately for each target orientation and distractor arrangement, and then averaged.

The statistical significance of each modulation ratio was evaluated using a permutation test. Specifically, for each neuron and distractor offset, we combined responses to targets alone and targets with distractors. We then sampled counts with replacement from this joint distribution, equal in number to those obtained experimentally. We calculated the modulation index for the sampled 'target' and 'target with distractor' responses. We repeated this 1000 times; significance was determined by the rank order of the observed ratio in this null distribution. To assess the significance of the modulation strength, we used a similar procedure. We calculated the modulation strength for each neuron by sampling from the combined distribution, and then averaged the modulation strength value across neurons. We

repeated this calculation 1000 times; significance of the reported values was determined by their rank order in the respective null distribution.

We assessed how the target orientation tuning of each neuron changed with distractors by fitting a model of the following form: $r_{TD} = m * r_T + a * \overline{r_T}$, where r_{TD} are the predicted responses, r_T are the measured responses to each isolated target (i.e. the response to each target stimulus), and $\overline{r_T}$ is the mean response averaged across isolated targets (i.e. the mean of the tuning function). Models were rectified at 0 to prevent negative firing rates. The values of the free parameters m and a were chosen by maximum likelihood, where the log-likelihood of each model was calculated as:

$$logL = \sum log\left[\frac{r_{TD}^{r_T} e^{-r_T}}{r_T!}\right]$$

which is the log-likelihood under the assumption of Poisson variability. We performed this analysis on neurons whose tuning was consistent, defined by calculating the correlation between the target-only tuning computed on even versus odd trials. Neurons with a correlation above 0.5 were included. In these neurons, the fit quality was 0.63 and 0.73 in V1 and V4, respectively, where the fit quality was defined as a normalized likelihood, between a 'null' model which predicts the response to be the mean across all stimuli (a value of 0) and an 'oracle' model which predicts the measured response (a value of 1)[64].

To understand how distractor-induced modulation of tuning affected discriminability, we conducted a simple simulation. We measured the area under the ROC curve for responses generated from a Poisson distribution, whose rate was defined by either the measured mean response to targets alone or the modulation of this mean by the model defined in the preceding paragraph (with parameters chosen optimally for each cell separately).

To assess the similarity of the population representation for targets under different crowding conditions, we compared population decoding performance across conditions using a shared, fixed decoding strategy. We fit linear decoders with regularization, as previously described, to the population responses to the two most differing target orientations. One decoder was estimated using responses to targets in isolation, and a second decoder ('crowded') was estimated from responses to target stimuli with nearby distractors. Responses on held out trials were projected onto these two readout axes, as were an equal number of trials from the other target orientations and conditions of crowding. Population discriminability for two target stimuli along a specific readout axis was quantified by the d-prime between the response distributions along the decoding axis, where $d\,prime = (\mu_1 - \mu_2) / \sqrt{(0.5 * (\sigma_1^2 + \sigma_2^2))}$. We compared d-prime measures for responses projected onto the readout axis for the matching condition ('same decoder') to those projected onto an opposing crowding condition ('cross decoder').

We observed similar results in both animals so pooled the sampled neurons for most analyses. Effects of distractors on population information are shown separately for each

animal in Figure 2E; changes in single-neuron discriminability and tuning are shown separately for each animal in Figure S7.

All indications of variance in the text and figures are SEM unless otherwise noted.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements:

REFERENCES

1. Levi DM (2008) Crowding--an essential bottleneck for object recognition: a mini-review. Vision Res. 48: 635–654. [PubMed: 18226828]

2. Pelli DG, Tillman KA (2008) The uncrowded window of object recognition. Nat Neurosci. 11: 1129–1135. [PubMed: 18828191]

3. Whitney D, Levi DM (2011) Visual crowding: a fundamental limit on conscious perception and object recognition. Trends Cogn Sci. 15: 160–168. [PubMed: 21420894]

4. Manassi M, Whitney D (2018) Multi-level crowding and the paradox of object recognition in clutter. Curr Biol. 28: R127–R133. [PubMed: 29408262]

5. Pelli DG (2008) Crowding: a cortical constraint on object recognition. Curr Opin Neurobiol. 18: 445–451. [PubMed: 18835355]

6. Levi DM, Carney T (2009) Crowding in peripheral vision: why bigger is better. Curr Biol. 19: 1988–1993. [PubMed: 19853450]

7. Balas B, Nakano L, Rosenholtz R (2009) A summary-statistic representation in peripheral vision explains visual crowding. J Vis. 9(12): 13.1–18. [PubMed: 19761328]

8. Freeman J, Simoncelli EP (2011) Metamers of the ventral stream. Nat Neurosci. 14: 1195–1201. [PubMed: 21841776]

9. Parkes L, Lund J, Angelucci A, Solomon JA, Morgan M (2001) Compulsory averaging of crowded orientation signals in human vision. Nat Neurosci. 4: 739–744. [PubMed: 11426231]

10. Greenwood JA, Bex PJ, Dakin SC (2009) Positional averaging explains crowding with letter-like stimuli. Proc Natl Acad Sci USA 106: 13130–13135 [PubMed: 19617570]

11. Ester EF, Zilber E, Serences JT (2015) Substitution and pooling in visual crowding induced by similar and dissimilar distractors. J Vis. 15(1): 15.1.4.

12. He S, Cavanagh P, Intriligator J (1996) Attentional resolution and the locus of visual awareness. Nature 383: 334–337. [PubMed: 8848045]

13. He D, Wang Y, Fang F (2019) The critical role of V2 population receptive fields in visual orientation crowding. Curr Biol. 29: 2229–2236.e3. [PubMed: 31231052]

14. Freeman J, Donner TH, Heeger DJ (2011b) Inter-area correlations in the ventral visual pathway reflect feature integration. J. Vis 11.

15. Bi T, Cai P, Zhou T, Fang F (2009) The effect of crowding on orientation selective adaptation in human early visual cortex. J. Vis 9, 13.1–13.10. [PubMed: 19761328]

16. Anderson EJ, Dakin SC, Schwarzkopf DS, Rees G, Greenwood JA (2012) The neural correlates of crowding-induced changes in appearance. Curr. Biol 22: 1199–1206. [PubMed: 22658599]

17. Millin R, Arman AC, Chung ST, Tjan BS (2014) Visual crowding in V1. Cereb. Cortex 24: 3107–3115. [PubMed: 23833128]

18. Chen J, He Y, Zhu Z, Zhou T, Peng Y, Zhang X, Fang F (2014) Attention-dependent early cortical suppression contributes to crowding. J. Neurosci 34: 10465–10474. [PubMed: 25100582]

19. Kwon M, Bao P, Millin R, Tjan BS (2014) Radial–tangential anisotropy of crowding in the early visual areas. J. Neurophysiol 112: 2413–2422. [PubMed: 25122703]

20. Chicherov V, Plomp G, Herzog MH (2014) Neural correlates of visual crowding. Neuroimage 93: 23–31. [PubMed: 24582921]

21. Henry CA, Kohn A (2020) Spatial contextual effects in primary visual cortex limit feature representation under crowding. Nat Commun. 11(1): 1687. doi: 10.1038/s41467-020-15386-7. [PubMed: 32245941]

22. Van Essen DC, Newsome WT, Maunsell JH (1984) The visual field representation in striate cortex of the macaque monkey: asymmetries, anisotropies, and individual variability. Vision Res. 24: 429–448. [PubMed: 6740964]

23. Gattass R, Sousa AP, Gross CG (1988) Visuotopic organization and extent of V3 and V4 of the macaque. J Neurosci. 8: 1831–1845. [PubMed: 3385477]

24. Bushnell BN, Harding PJ, Kosai Y, Pasupathy A (2011) Partial occlusion modulates contour-based shape encoding in primate area V4. J Neurosci. 31: 4012–4024. [PubMed: 21411644]

25. Motter BC (2018) Stimulus conflation and tuning selectivity in V4 neurons: a model of visual crowding. J Vis. 18(1): 15. doi: 10.1167/18.1.15.

26. Bouma H (1970) Interaction effects in parafoveal letter recognition. Nature 226: 177–178. [PubMed: 5437004]

27. Kohn A (2007) Visual adaptation: physiology, mechanisms, and functional benefits. J Neurophysiol. 97: 3155–3164. [PubMed: 17344377]

28. Averbeck BB, Latham PE, Pouget A (2006) Neural correlations, population coding and computation. Nat Rev Neurosci. 7: 358–366. [PubMed: 16760916]

29. Kohn A, Coen-Cagli R, Kanitscheider I, Pouget A (2016) Correlations and neuronal population information. Annu Rev Neurosci. 39: 237–256. [PubMed: 27145916]

30. Henry CA, Joshi S, Xing D, Shapley RM, Hawken MJ (2013) Functional characterization of the extraclassical receptive field in macaque V1: contrast, orientation, and temporal dynamics. J Neurosci. 33: 6230–6242. [PubMed: 23554504]

31. Snyder AC, Morais MJ, Kohn A, Smith MA (2014) Correlations in V1 are reduced by stimulation outside the receptive field. J Neurosci. 34: 11222–11227. [PubMed: 25143603]

32. Trott AR, Born RT (2015) Input-gain control produces feature-specific surround suppression. J Neurosci. 35: 4973–4982. [PubMed: 25810527]

33. Movshon JA, Thompson ID, Tolhurst DJ (1978) Spatial summation in the receptive fields of simple cells in the cat's striate cortex. J Physiol. 283: 53–77. [PubMed: 722589]

34. Carandini M, Demb JB, Mante V, Tolhurst DJ, Dan Y, Olshausen BA, Gallant JL, Rust NC (2005) Do we know what the early visual system does? J Neurosci. 25: 10577–10597. [PubMed: 16291931]

35. Carandini M, Heeger DJ (2011) Normalization as a canonical neural computation. Nat Rev Neurosci. 13: 51–62. [PubMed: 22108672]

36. Cavanaugh JR, Bair W, Movshon JA (2002) Nature and interaction of signals from the receptive field center and surround in macaque V1 neurons. J Neurophysiol. 88: 2530–2546. [PubMed: 12424292]

37. Angelucci A, Bressloff PC (2006) Contribution of feedforward, lateral and feedback connections to the classical receptive field center and extra-classical receptive field surround of primate V1 neurons. Prog Brain Res. 154: 93–120. [PubMed: 17010705]

38. Ghose GM, Maunsell JH (2008) Spatial summation can explain the attentional modulation of neuronal responses to multiple stimuli in area V4. J Neurosci. 28: 5115–5126. [PubMed: 18463265]

39. Verhoef BE, Maunsell JHR (2017) Attention-related changes in correlated neuronal activity arise from normalization mechanisms. Nat Neurosci. 20: 969–977. [PubMed: 28553943]

40. Ruff DA, Alberts JJ, Cohen MR (2016) Relating normalization to neuronal populations across cortical areas. J. Neurophysiol 116: 1375–1386. [PubMed: 27358313]

41. Zanos TP, Mineault PJ, Monteon JA, Pack CC (2011) Functional connectivity during surround suppression in macaque area V4. Annu Int Conf IEEE Eng Med Biol Soc. 2011:3342–45. [PubMed: 22255055]

42. Ni AM, Ray S, Maunsell JH (2012) Tuned normalization explains the size of attention modulations. Neuron 73: 803–13. [PubMed: 22365552]

43. Arandia-Romero I, Tanabe S, Drugowitsch J, Kohn A, Moreno-Bote R (2016) Multiplicative and additive modulation of neuronal tuning with population activity affects encoded information. Neuron 89: 1305–1316. [PubMed: 26924437]

44. Pelli DG, Palomares M, Majaj NJ (2004) Crowding is unlike ordinary masking: distinguishing feature integration from detection. J. Vis. 4: 1136–1169. [PubMed: 15669917]

45. Petrov Y, Carandini M, McKee S (2005) Two distinct mechanisms of suppression in human vision. J Neurosci. 25: 8704–8707. [PubMed: 16177039]

46. Ziemba CM, Simoncelli EP (2021) Opposing effects of selectivity and invariance in peripheral vision. Nat Commun. 12: 4597. doi: 10.1038/s41467-021-24880-5. [PubMed: 34321483]

47. Chen M, Yan Y, Gong X, Gilbert CD, Liang H, Li W (2014b) Incremental integration of global contours through interplay between visual cortical areas. Neuron 82: 682–694. [PubMed: 24811385]

48. Pasupathy A, Connor CE (2002) Population coding of shape in area V4. Nat Neurosci. 5: 1332–1338. [PubMed: 12426571]

49. Nandy AS, Sharpee TO, Reynolds JH, Mitchell JF (2013) The fine structure of shape tuning in area V4. Neuron 78: 1102–1115. [PubMed: 23791199]

50. El-Shamayleh Y, Pasupathy A (2016) Contour curvature as an invariant code for objects in visual area V4. J Neurosci. 36:5532–5543. [PubMed: 27194333]

51. Hu JM, Song XM, Wang Q, Roe AW (2020) Curvature domains in V4 of macaque monkey. Elife 9: e57261. [PubMed: 33211004]

52. Okazawa G, Tajima S, Komatsu H (2017) Gradual development of visual texture-selective properties between macaque areas V2 and V4. Cereb Cortex. 27: 4867–4880. [PubMed: 27655929]

53. Kim T, Bair W, Pasupathy A (2022) Perceptual texture dimensions modulate neuronal response dynamics in visual cortical area V4. J Neurosci 42: 631–642. [PubMed: 34862189]

54. Desimone R, Schein SJ (1987) Visual properties of neurons in area V4 of the macaque: sensitivity to stimulus form. J Neurophysiol. 57: 835–68. [PubMed: 3559704]

55. Ghose GM, Ts'o DY (1997) Form processing modules in primate area V4. J Neurophysiol. 77: 2191–2196. [PubMed: 9114265]

56. Hegdé J, Van Essen DC (2006) Temporal dynamics of 2D and 3D shape representation in macaque visual area V4. Vis Neurosci. 23: 749–763. [PubMed: 17020631]

57. Roe AW, Chelazzi L, Connor CE, Conway BR, Fujita I, Gallant JL, Lu H, Vanduffel W (2012) Toward a unified theory of visual area V4. Neuron. 74: 12–29. [PubMed: 22500626]

58. Pasupathy A, Popovkina DV, Kim T (2020) Visual functions of primate area V4. Annu Rev Vis Sci. 6: 363–385. [PubMed: 32580663]

59. Beck JM, Ma WJ, Pitkow X, Latham PE, Pouget A (2012) Not noisy, just wrong: the role of suboptimal inference in behavioral variability. Neuron 74: 30–39. [PubMed: 22500627]

60. Jin M, Beck JM, Glickfeld LL (2019) Neuronal adaptation reveals a suboptimal decoding of orientation tuned populations in the mouse visual cortex. J Neurosci. 39: 3867–3881. [PubMed: 30833509]

61. Kelly RC, Smith MA, Samonds JM, Kohn A, Bonds AB, Movshon JA, Lee TS (2007) Comparison of recordings from microelectrode arrays and single electrodes in the visual cortex. J Neurosc. 27: 261–264.

62. Green DM, Swets JA (1966) Signal detection theory and psychophysics. John Wiley.

63. Kohn A, Smith MA (2005) Stimulus dependence of neuronal correlation in primary visual cortex of the macaque. J Neurosci. 25: 3661–3673. [PubMed: 15814797]

64. El-Shamayleh Y, Movshon JA (2011) Neuronal responses to texture-defined form in macaque visual area V2. J Neurosci. 31: 8543–8555. [PubMed: 21653858]

Highlights

- Crowding impairs encoding of target orientation in macaque V1 and V4 populations

- Target discriminability was moderately impaired in V1 and more substantially in V4

- The presence of distractors led to additive and multiplicative changes in tuning

- Fixed readout strategies were limited in performance by non-affine tuning changes
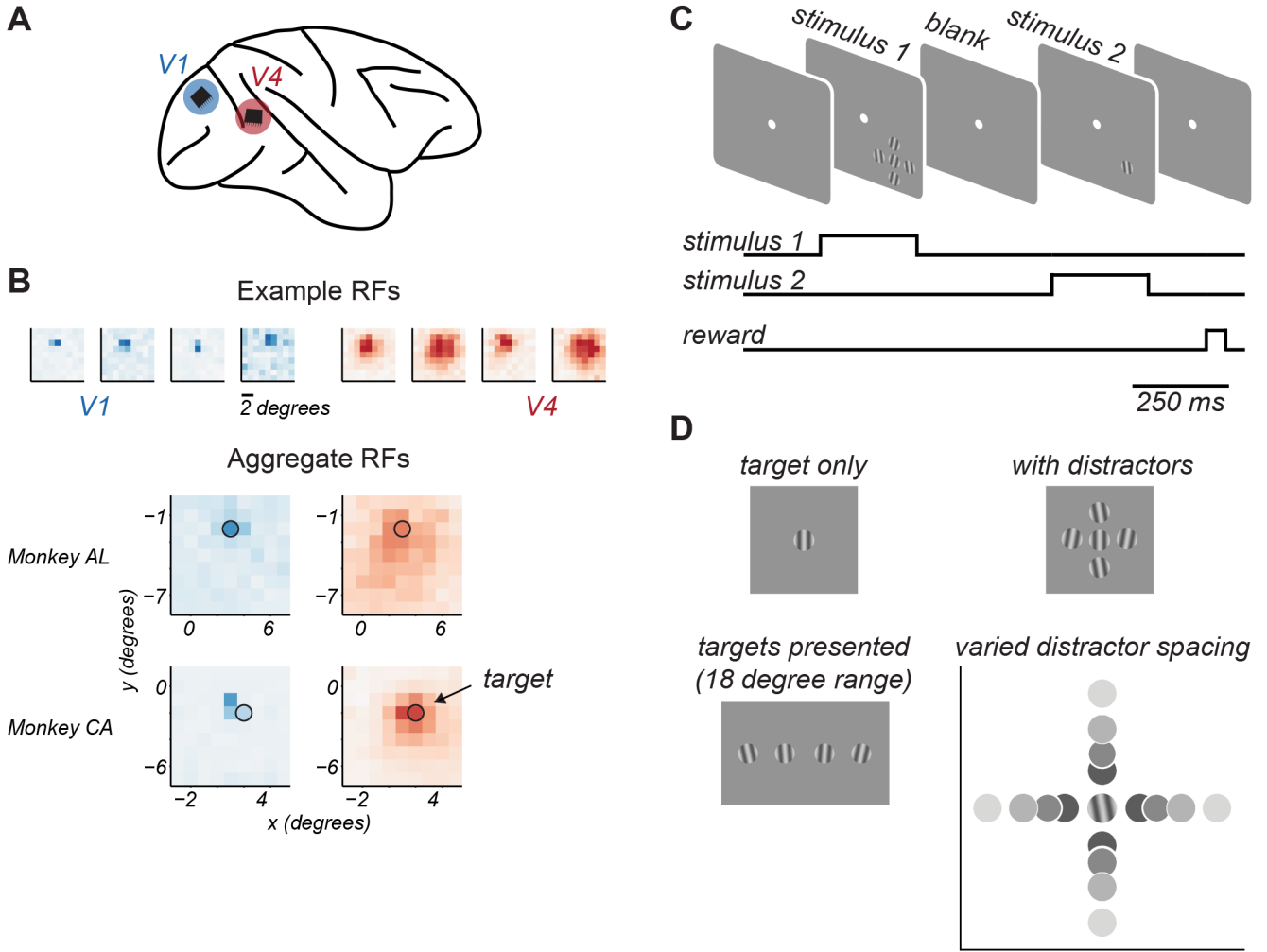
Figure 1. Experimental approach
A. Microelectrode arrays were chronically implanted into V1 and V4 of one hemisphere of each animal. B. Top. Spatial receptive field maps of 4 example neurons from V1 (blue) and V4 (red). Darker colors indicate higher firing rates. Bottom. Aggregate receptive field maps of recorded neuronal populations, by cortical area (columns) and animal (rows). Black circle indicates target stimulus location and approximate size. C. Task structure. Animals maintained fixation of a central spot while target stimuli were presented in two 0.25s epochs on each trial, separated by a 0.35s blank period. D. Target stimuli were drifting gratings chosen from an 18 degree range around a reference orientation. When shown, distractors were presented at ±10 deg relative to this reference orientation. Target-distractor spacing was varied.
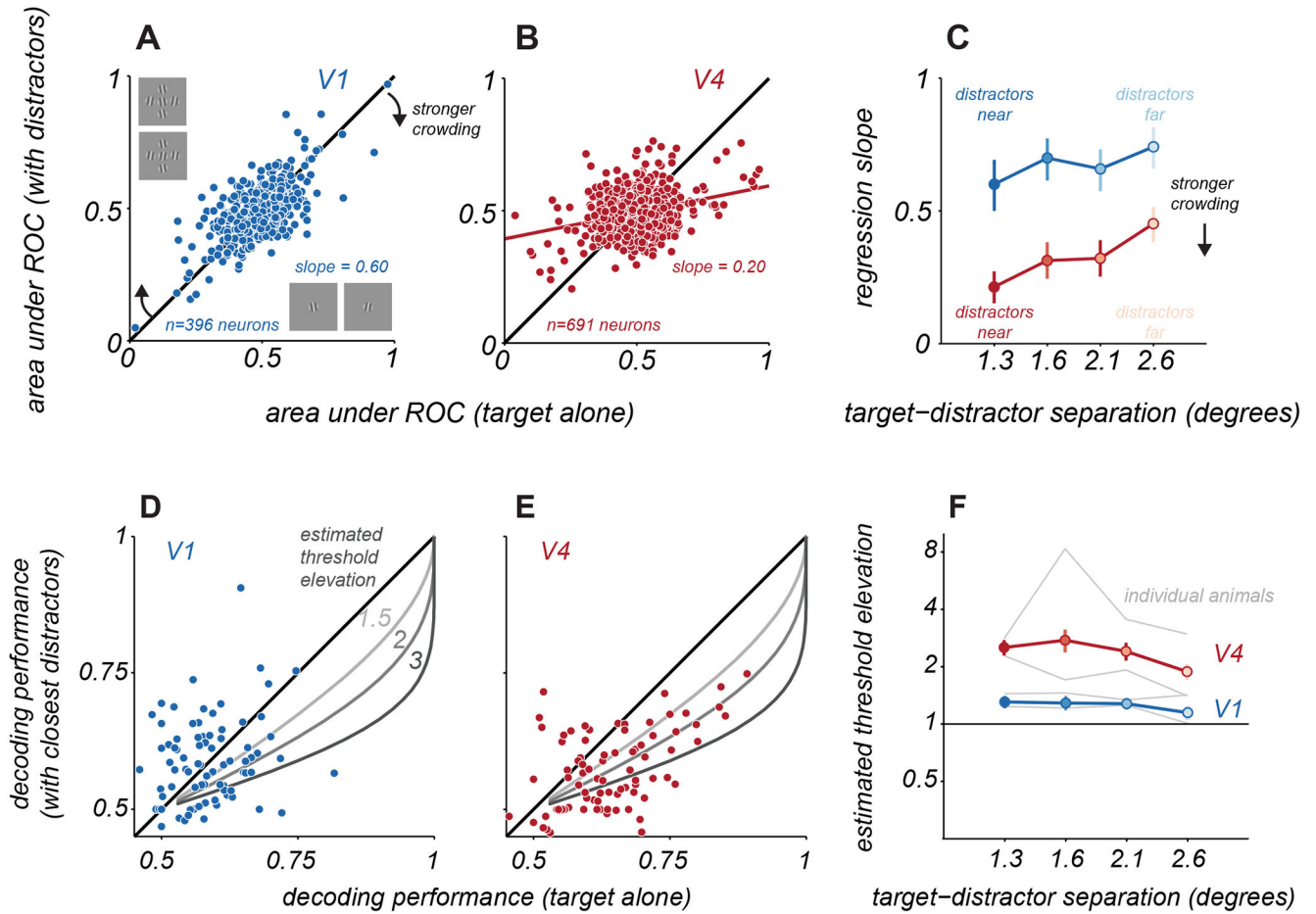
Figure 2. Single neuron and population discriminability of target orientation

A. Target discriminability of V1 neurons for targets alone (abscissa) and targets with near distractors (ordinate), using responses to the two most differing target orientations. Solid blue line indicates linear regression fit to the data. B. Same as A, but for V4 neurons. C. Change in discriminability as a function of target-distractor spacing for V1 (blue) and V4 (red), quantified using the slope of the regression line. Error bars indicate 95% confidence interval. D. Cross-validated V1 population discrimination performance for targets alone (abscissa) and targets with distractors (ordinate). Each point represents a single pairing of two target stimuli. Data are combined across pairings and sessions, for one animal and the nearest target-distractor spacing (n=84 cases). Gray lines indicate expected performance for the indicated threshold elevations. E. Same as D. but for V4. F. Estimated threshold elevation for populations as a function of target-distractor separation. Points represent the mean across both animals (V1, blue; V4, red); error bars indicate s.e.m., estimated via bootstrapped resampling of the data. Gray lines represent the mean values for each animal and cortical area. See also Figures S1, S2, S3, and S7.
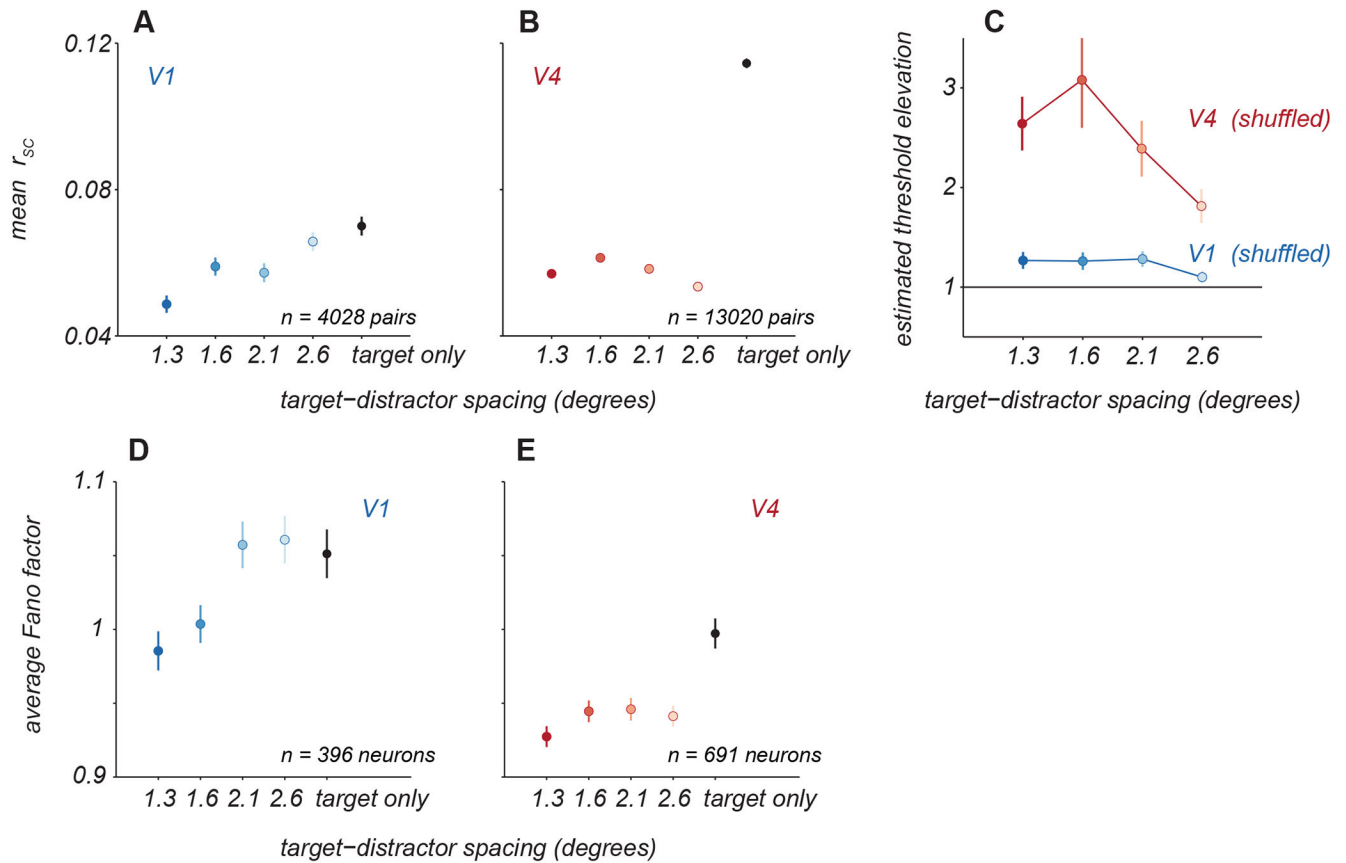
Figure 3. Effect of distractors on neuronal variability

A. Average pairwise 'noise' correlations (r_{SC}) for V1 neuronal pairs for targets alone (black) and with distractors of varied spacing (blue). Points and error bars represent mean ± s.e.m. B. Average correlations for V4 neuronal pairs, same conventions as in A. C. Crowding effect size for V1 and V4 populations when decoding was performed on trial-shuffled population activity. Points and error bars represent mean ± s.e.m. D. Individual neuron variability was quantified as the average Fano factor for V1 neurons for targets alone (black) and with distractors at various spacings (blue). Points and error bars represent the geometric mean ± s.e.m. over all neurons. E. The average Fano factor for V4 neurons, same conventions as in A.
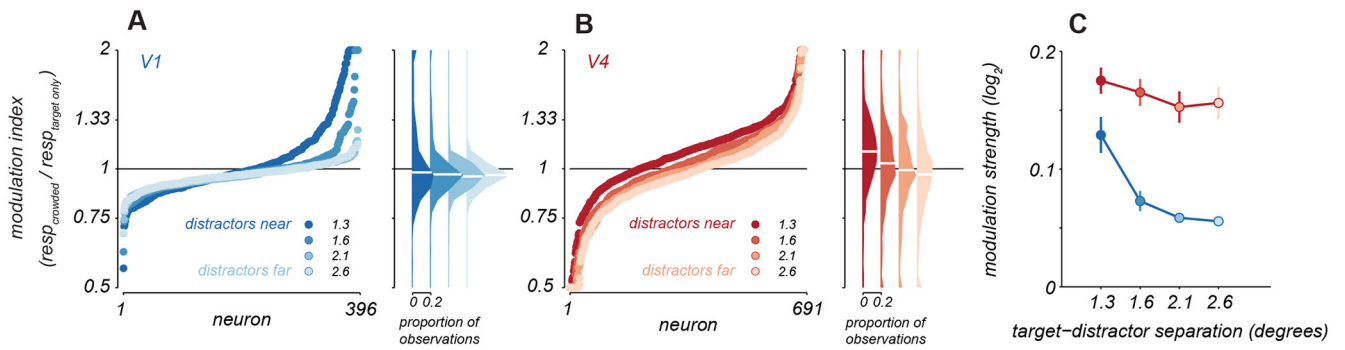
Figure 4. Response modulation with distractors

A. Modulation indices were calculated as the average response to targets with distractors over the response to targets alone. Points represent sorted modulation indices for all V1 neurons, with target-distractor spacing indicated by blue shading. Distributions of modulation indices are shown at the right, with median values for each distribution indicated by white lines. B. Modulation indices for V4 neurons in red, same conventions as in A. C. Average modulation strength as a function of target-distractor spacing for V1 (blue) and V4 (red). Points represent mean ± s.e.m (V1: n=396, V4: n=691 neurons). See also Figure S4.
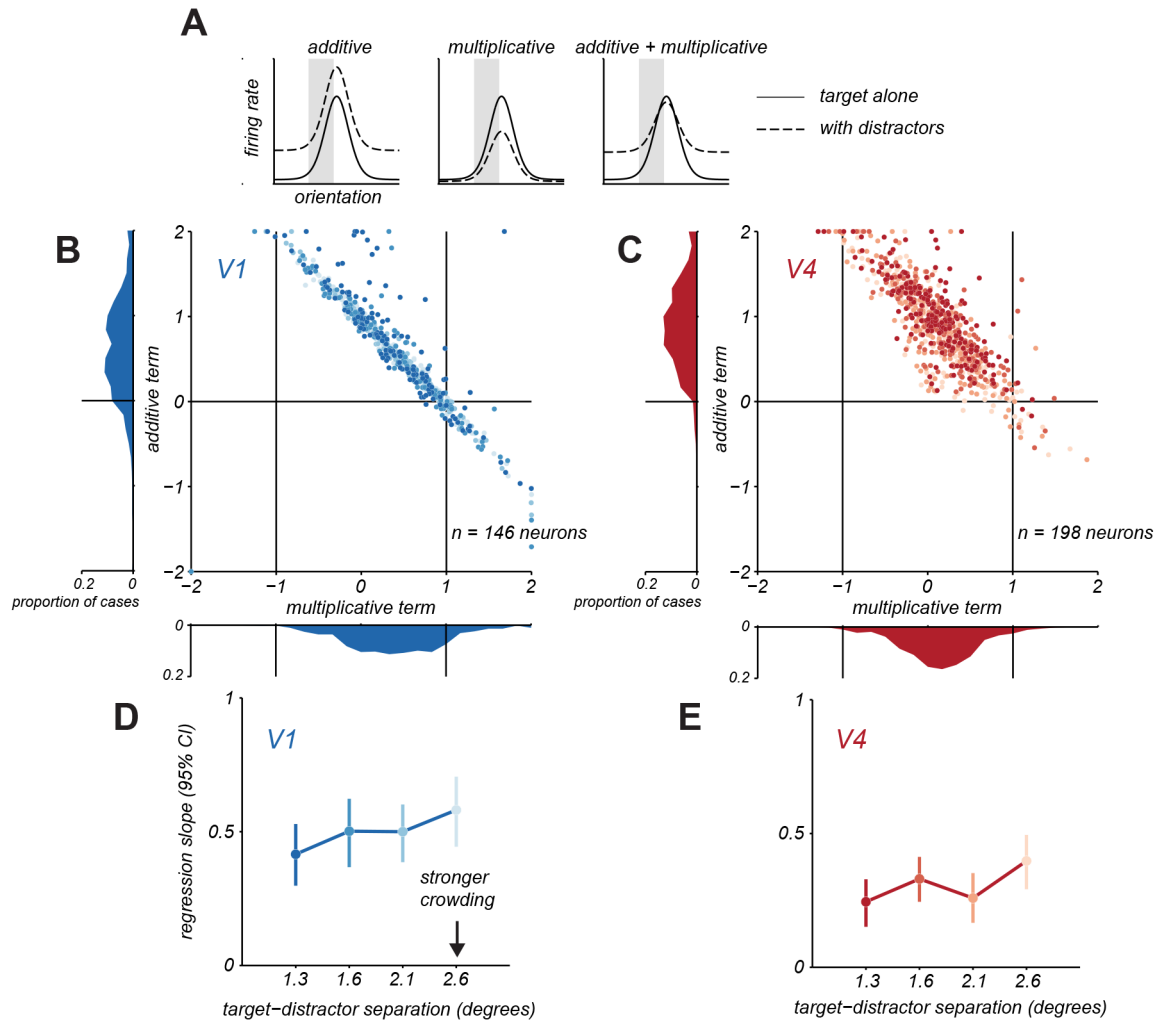
Figure 5. Additive and multiplicative changes in target tuning with distractors
A. Schematic illustration of how target tuning amid distractors (dashed line) might involve additive translation, multiplicative scaling, or combined changes in target only tuning (solid line). B. Multiplicative and additive terms fit to V1 neurons under crowding. Each point represents one neuron and distractor spacing condition, darker shades indicate closer distractor spacing. Marginal distributions of parameters are shown along the axes. C. Same as in B, but for V4 neurons. D. Change in V1 discriminability as a function of target-distractor spacing by applying the additive and multiplicative changes in target only tuning from B. Change in discriminability was quantified using the slope of the regression line relating discriminability for target only and target with distractors conditions. Error bars indicate 95% confidence interval. E. Same as in D, but for V4. See also Figures S5 and S7.
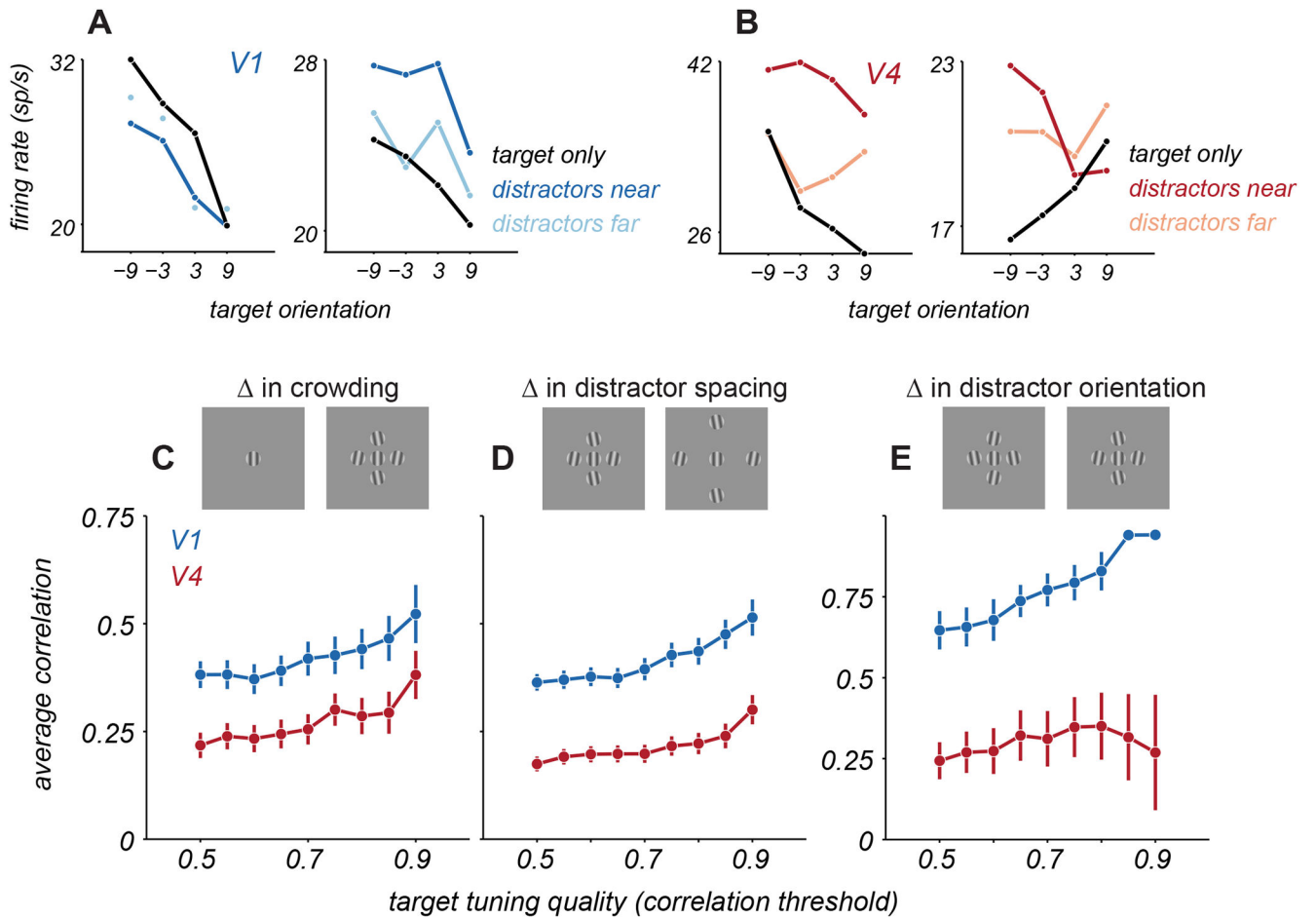
Figure 6. Change in target tuning with distractors
A. Target orientation tuning for 2 example V1 neurons for targets alone (black) and for targets with near (dark blue) or far (light blue) distractors. B. Target orientation tuning for 2 example V4 neurons for targets alone (black) and with distractors (red). C. Average correlation between tuning for targets alone and targets with near distractors for V1 (blue) and V4 (red), as a function of target tuning quality (since the ability to detect a change in tuning will depend on how consistent and well-defined the tuning is for targets alone). Target tuning quality is defined as the correlation between the tuning for targets along, calculated from even and odd trials (i.e., a measure of whether tuning was similar in each half of the data, a measure of tuning quality or consistency). Points and error bars represent mean ± s.e.m. D. Average correlation between target tuning with near and far distractors. E. Average correlation between target tuning for two near distractor configurations, counterbalanced for distractor orientation.
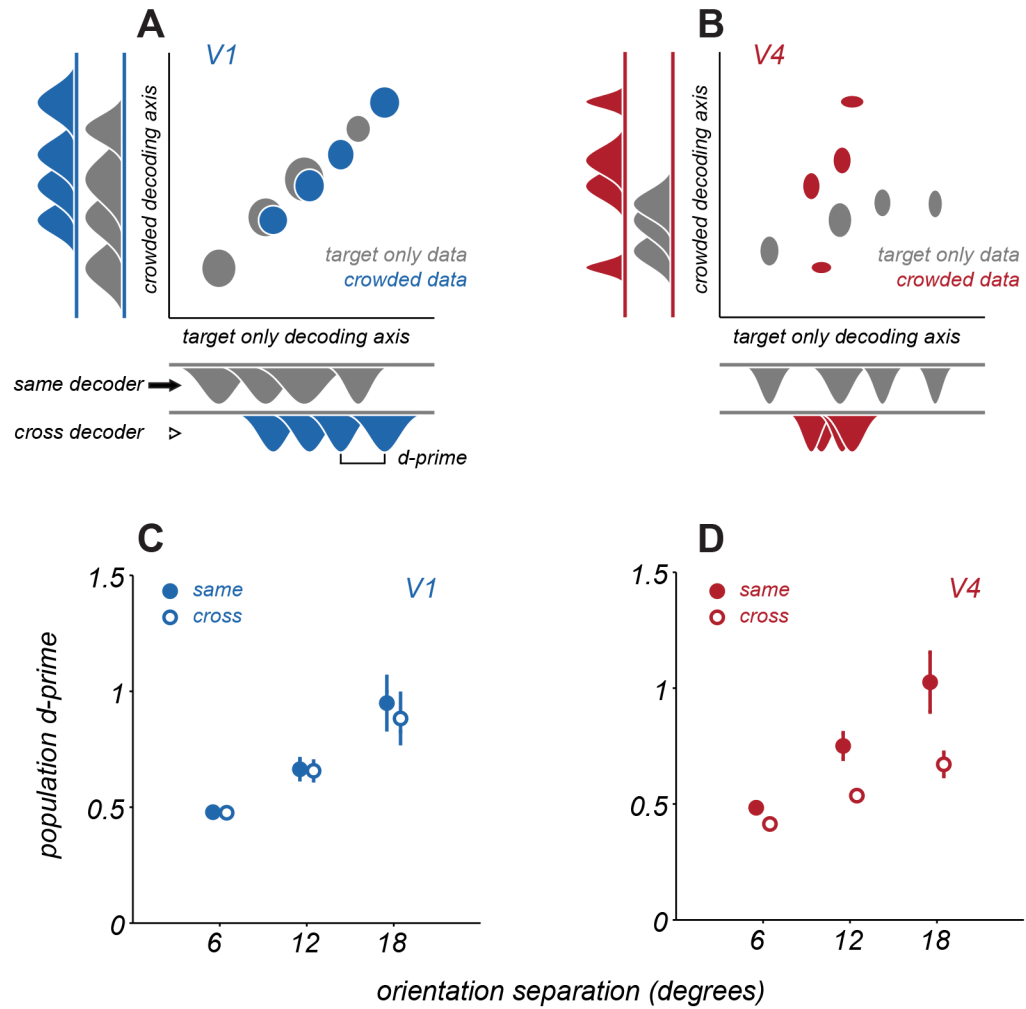
Figure 7. Generalization of discriminability across decoding axes

A. An example V1 population response projected onto two distinct decoding axes, one optimized for target only responses (abscissa) and one optimized for responses to targets with distractors (ordinate). Gray ellipses represent mean ± s.e.m. of the spike counts projected onto the relevant decoding axis, over repeated trials of a single target stimulus. The 4 ellipses represent the distinct target orientations shown. Blue ellipses represent the same, but for responses to targets with distractors. Marginal distributions are shown at the sides of the plot. B. An example V4 population response, same conventions as in A. C. Average V1 discriminability is measured as the d-prime between two distributions, when projected onto the same decoding axis as the data (same, filled circles) or the decoding axis from the opposing condition of crowding (cross, open circles). Points and error bars represent mean ± s.e.m. Points are offset slightly along abscissa for visualization. D. Average V4 discriminability, same conventions as in C. See also Figure S6.

KEY RESOURCES TABLE

LIFE SCIENCE TABLE WITH EXAMPLES FOR AUTHOR REFERENCE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit monoclonal anti-Snail	Cell Signaling Technology	Cat#3879S; RRID: AB_2255011
Mouse monoclonal anti-Tubulin (clone DM1A)	Sigma-Aldrich	Cat#T9026; RRID: AB_477593
Rabbit polyclonal anti-BMAL1	This paper	N/A
Bacterial and virus strains		
pAAV-hSyn-DIO-hM3D(Gq)-mCherry	Krashes et al.[1]	Addgene AAV5; 44361-AAV5
AAV5-EF1a-DIO-hChR2(H134R)-EYFP	Hope Center Viral Vectors Core	N/A
Cowpox virus Brighton Red	BEI Resources	NR-88
Zika-SMGC-1, GENBANK: KX266255	Isolated from patient (Wang et al.[2])	N/A
Staphylococcus aureus	ATCC	ATCC 29213
Streptococcus pyogenes: M1 serotype strain: strain SF370; M1 GAS	ATCC	ATCC 700294
Biological samples		
Healthy adult BA9 brain tissue	University of Maryland Brain & Tissue Bank; http://medschool.umaryland.edu/btbank/	Cat#UMB1455
Human hippocampal brain blocks	New York Brain Bank	http://nybb.hs.columbia.edu/
Patient-derived xenografts (PDX)	Children's Oncology Group Cell Culture and Xenograft Repository	http://cogcell.org/
Chemicals, peptides, and recombinant proteins		
MK-2206 AKT inhibitor	Selleck Chemicals	S1078; CAS: 1032350-13-2
SB-505124	Sigma-Aldrich	S4696; CAS: 694433-59-5 (free base)
Picrotoxin	Sigma-Aldrich	P1675; CAS: 124-87-8
Human TGF-β	R&D	240-B; GenPept: P01137
Activated S6K1	Millipore	Cat#14-486
GST-BMAL1	Novus	Cat#H00000406-P01
Critical commercial assays		
EasyTag EXPRESS 35S Protein Labeling Kit	PerkinElmer	NEG772014MC
CaspaseGlo 3/7	Promega	G8090
TruSeq ChIP Sample Prep Kit	Illumina	IP-202-1012
Deposited data		
Raw and analyzed data	This paper	GEO: GSE63473
B-RAF RBD (apo) structure	This paper	PDB: 5J17

Human reference genome NCBI build 37, GRCh37	Genome Reference Consortium	http://www.ncbi.nlm.nih.gov/projects/genome/assembly/grc/human/
Nanog STILT inference	This paper; Mendeley Data	http://dx.doi.org/10.17632/wx6s4mj7s8.2
Affinity-based mass spectrometry performed with 57 genes	This paper; Mendeley Data	Table S8; http://dx.doi.org/10.17632/5hvpvspw82.1
Experimental models: Cell lines		
Hamster: CHO cells	ATCC	CRL-11268
D. melanogaster: Cell line S2: S2-DRSC	Laboratory of Norbert Perrimon	FlyBase. FBtc0000181
Human: Passage 40 H9 ES cells	MSKCC stem cell core facility	N/A
Human: HUES 8 hESC line (NIH approval number NIHhESC-09-0021)	HSCI iPS Core	hES Cell Line: HUES-8
Experimental models. Organisms/strains		
C. elegans: Strain BC4011: srl-1(s2500) II; dpy-18(e364) III; unc-46(e177)rol-3(s1040) V.	Caenorhabditis Genetics Center	WB Strain: BC4011; WormBase: WBVar00241916
D. melanogaster: RNAi of Sxl: y[1] sc[*] v[1]; P{TRiP.HMS00609}attP2	Bloomington Drosophila Stock Center	BDSC:34393; FlyBase: FBtp0064874
S. cerevisiae: Strain background: W303	ATCC	ATTC: 208353
Mouse: R6/2: B6CBA-Tg(HDexon1)62Gpb/3J	The Jackson Laboratory	JAX: 006494
Mouse: OXTRfl/fl: B6.129(SJL)-Oxtr[tm1.1Wsy]/J	The Jackson Laboratory	RRID: IMSR_JAX:008471
Zebrafish: Tg(Shha:GFP)t10: t10Tg	Neumann and Nuesslein-Volhard[3]	ZFIN: ZDB-GENO-060207-1
Arabidopsis: 35S::PIF4-YFP, BZR1-CFP	Wang et al.[4]	N/A
Arabidopsis: JYB1021.2: pS24(AT5G58010)::cS24:GFP(-G):NOS #1	NASC	NASC ID: N70450
Oligonucleotides		
siRNA targeting sequence: PIP5K I alpha #1: ACACAGUACUCAGUUGAUA	This paper	N/A
Primers for XX, see Table SX	This paper	N/A
Primer: GFP/YFP/CFP Forward: GCACGACTTCTTCAAGTCCGCCATGCC	This paper	N/A
Morpholino: MO-pax2a GGTCTGCTTTGCAGTGAATATCCAT	Gene Tools	ZFIN: ZDB-MRPHLNO-061106-5
ACTB (hs01060665_g1)	Life Technologies	Cat#4331182
RNA sequence: hnRNPA1_ligand: UAGGGACUUAGGGUUCUCUCUAGGGACUUAGGGUUCUCUCUAGGGA	This paper	N/A
Recombinant DNA		
pLVX-Tight-Puro (TetOn)	Clonetech	Cat#632162
Plasmid: GFP-Nito	This paper	N/A
cDNA GH111110	Drosophila Genomics Resource Center	DGRC:5666; FlyBase:FBcl0130415
AAV2/1-hsyn-GCaMP6- WPRE	Chen et al.[5]	N/A
Mouse raptor: pLKO mouse shRNA 1 raptor	Thoreen et al.[6]	Addgene Plasmid #21339
Software and algorithms		

ImageJ	Schneider et al.[7]	https://imagej.nih.gov/ij/
Bowtie2	Langmead and Salzberg[8]	http://bowtie-bio.sourceforge.net/bowtie2/index.shtml
Samtools	Li et al.[9]	http://samtools.sourceforge.net/
Weighted Maximal Information Component Analysis v0.9	Rau et al.[10]	https://github.com/ChristophRau/wMICA
ICS algorithm	This paper; Mendeley Data	http://dx.doi.org/10.17632/5hvpvspw82.1
Other		
Sequence data, analyses, and resources related to the ultra-deep sequencing of the AML31 tumor, relapse, and matched normal	This paper	http://aml31.genome.wustl.edu
Resource website for the AML31 publication	This paper	https://github.com/chrisamiller/aml31SuppSite

PHYSICAL SCIENCE TABLE WITH EXAMPLES FOR AUTHOR REFERENCE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, peptides, and recombinant proteins		
QD605 streptavidin conjugated quantum dot	Thermo Fisher Scientific	Cat#Q10101MP
Platinum black	Sigma-Aldrich	Cat#205915
Sodium formate BioUltra, 99.0% (NT)	Sigma-Aldrich	Cat#71359
Chloramphenicol	Sigma-Aldrich	Cat#C0378
Carbon dioxide (^{13}C, 99%) (<2% ^{18}O)	Cambridge Isotope Laboratories	CLM-185-5
Poly(vinylidene fluoride-co-hexafluoropropylene)	Sigma-Aldrich	427179
PTFE Hydrophilic Membrane Filters, 0.22 μm, 90 mm	Scientificfilters.com/Tisch Scientific	SF13842
Critical commercial assays		
Folic Acid (FA) ELISA kit	Alpha Diagnostic International	Cat# 0365-0B9
TMT10plex Isobaric Label Reagent Set	Thermo Fisher	A37725
Surface Plasmon Resonance CM5 kit	GE Healthcare	Cat#29104988
NanoBRET Target Engagement K-5 kit	Promega	Cat#N2500
Deposited data		
B-RAF RBD (apo) structure	This paper	PDB: 5J17
Structure of compound 5	This paper; Cambridge Crystallographic Data Center	CCDC: 2016466
Code for constraints-based modeling and analysis of autotrophic *E. coli*	This paper	https://gitlab.com/elad.noor/sloppy/tree/master/rubisco
Software and algorithms		
Gaussian09	Frish et al.[1]	https://gaussian.com
Python version 2.7	Python Software Foundation	https://www.python.org
ChemDraw Professional 18.0	PerkinElmer	https://www.perkinelmer.com/category/chemdraw

Weighted Maximal Information Component Analysis v0.9	Rau et al.[2]	https://github.com/ChristophRau/wMICA
Other		
DASGIP MX4/4 Gas Mixing Module for 4 Vessels with a Mass Flow Controller	Eppendorf	Cat#76DGMX44
Agilent 1200 series HPLC	Agilent Technologies	https://www.agilent.com/en/products/liquid-chromatography
PHI Quantera II XPS	ULVAC-PHI, Inc.	https://www.ulvacphi.com/en/products/xps/phi-quantera-ii/