bioRxiv preprint doi: https://doi.org/10.1101/777912; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Patterns of neural correlations in V1 vary with the number of objects

Na Young Jun (1-3)

Douglas A. Ruff (4-5)

Surya T Tokdar (6)

Marlene R Cohen (4-5)

Jennifer M Groh (1-3, 7-9)

1. Department of Neurobiology, Duke University
2. Center for Cognitive Neuroscience, Duke University
3. Duke Institute for Brain Sciences
4. Department of Neuroscience, University of Pittsburgh
5. Center for the Neural Basis of Cognition, University of Pittsburgh
6. Department of Statistical Science, Duke University
7. Department of Psychology and Neuroscience, Duke University
8. Department of Biomedical Engineering, Duke University
9. Department of Computer Science, Duke University

Acknowledgments:

Thanks to: Valeria C. Caruso, Yunran Chen, Jeffrey T. Mohl, Meredith N. Schmehl, Shawn M. Willett, for helpful comments on the analysis and manuscript. This work was supported by: National Institutes of Health grant nos. R00EY020844 (MRC), R01EY022930 (MRC); Core Grant P30 EY008098s (MRC); R01DC013906 (JMG, STT); and R01DC016363 (JMG, STT); and by support to MRC from the McKnight, Whitehall, Sloan and Simons Foundations.

Abstract

How the brain preserves information about more than one stimulus at a time remains poorly understood. We recently showed that when more than one stimulus is present, single neurons may fluctuate between coding one vs. the other(s) across some time period. A critical unanswered question is whether and how any such fluctuations are coordinated across the neural population. Here, we analyzed the spike count ("noise") correlations observed between pairs of primary visual cortex (V1) neurons under a variety of conditions. We report that when two separate grating stimuli are presented simultaneously, distinct distributions of positive and negative correlations emerge, depending on whether the two neurons in the pair both respond more strongly to the same vs. different individual stimuli. Neural pairs that shared the same stimulus preference were more likely to show positively correlated spike count variability whereas those with different preferences were more likely to show negative correlations, suggesting that the population response to one particular stimulus may be enhanced over the other on any given trial. This pattern of results was not present when the two gratings were superimposed and formed a single plaid, supporting the interpretation that the pattern of correlated fluctuations is related to the segregation of individual objects in the visual scene.

bioRxiv preprint doi: https://doi.org/10.1101/777912; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Key Words

Spike count correlations; variability; V1; multiplexing; attention; population coding

Introduction

Natural sensory scenes contain many objects. In general, the populations of neurons capable of responding to each object overlap with one another, raising the question of how information about each item is preserved. Logic suggests that information about each distinct stimulus must be segregated in the neural code in some fashion, either into exclusive neural subpopulations or into different epochs of time or some combination of both. We have recently presented the hypothesis that the nervous system might employ a form of neural turn-taking, in which individual neurons fluctuate between responding to each of the items in or near their receptive fields across various epochs of time (Caruso, Mohl et al. 2018, Mohl, Caruso et al. 2019). Such a coding scheme could preserve information about each stimulus across time and/or across neural subpopulations.

One open question raised by this theory is whether and how activity patterns fluctuate together in a neural ensemble. For example, pairs of neurons that exhibit fluctuating activity in response to multiple stimuli might show either positive or negative correlations with each other. The pattern of such correlations across the population, and in particular how the observed correlations depend on the sensitivity of the two neurons in the pair to the individual component stimuli in the stimulus "bundle", can reveal insights into the degree to which the overall neural representation on any given trial may be biased towards one stimulus vs. another.

A second open question concerns how the pattern of fluctuations and their correlation across the population might depend on whether multiple "stimuli" fuse as a single object or are perceived as distinct from each other (Milner 1974, Gray and Singer 1989, Von Der Malsburg 1994, Singer and Gray 1995, Gray 1999) (but see Palanca and DeAngelis 2005). Stimuli that segregate into distinct objects may be more likely to be associated with fluctuations in neural activity and their attendant correlations across neurons, whereas stimuli that fuse into a single distinct object may cause activity patterns that are akin to those observed when only one stimulus is present. Such a pattern would specifically implicate activity fluctuations in playing a role in the perceptual process of object segregation.

To address these questions, we turned to the primary visual cortex (V1). V1 allows for a strong test of these hypotheses since V1 neurons have comparatively small receptive fields and are therefore much less subject to the multiple-stimulus-overlap problem than more broadly tuned areas such as the inferior colliculus or IT cortex that were assessed in our previous report (Caruso, Mohl et al. 2018). Furthermore, the responses of V1 neurons are typically much less modulated by cognitive or other modulatory processes than neurons in higher visual areas (e.g. (Luck, Chelazzi et al. 1997, Cook and Maunsell 2002, Buffalo, Fries et al. 2010))

bioRxiv preprint doi: https://doi.org/10.1101/777912; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

We evaluated activity in V1 while monkeys viewed either single stimuli (gratings) or two different types of combined stimuli (superimposed vs. adjacent gratings). When the two gratings were superimposed, they presumably appeared as a single object, or plaid (Adelson and Movshon 1982, Rodman and Albright 1989, Heeger, Simoncelli et al. 1996, Busse, Wade et al. 2009, Lima, Singer et al. 2010), whereas when they were adjacent they appeared as two distinct objects. We then evaluated the degree and sign of the trial-wise spike count correlations (commonly referred to as "noise" correlations (Cohen and Kohn 2011)) observed between pairs of simultaneously recorded units in response to particular stimulus conditions. We report here that the pattern of correlations varied dramatically depending on whether the stimuli were superimposed or side-by-side (one object vs. two). In the two-object case, the distribution of spike count correlations was markedly different from previous reports involving single stimuli (c.f. Table 1, Cohen and Kohn 2011) , and encompassed a range spanning many negative correlations in addition to positive ones. Whether the correlations tended to be positive vs. negative depended on whether the two neurons in the pair preferred the same stimulus (median correlation +0.25) or preferred different stimuli (median correlation -0.05). The distribution of spike count correlation values was even more widely spread when only pairs of neurons that showed demonstrably fluctuating activity across trials (using the statistical analyses we previously developed) were analyzed (same preference: +0.49 and different preference: -0.14). In contrast, in the single object case, positive correlations predominated (median value ~0.22).

Overall, this pattern of results is consistent with the possibility that when two visual objects are presented in close proximity, a subpopulation of V1 neurons fluctuates in a coordinated fashion and biases the population code in favor of one object over the other on individual trials.

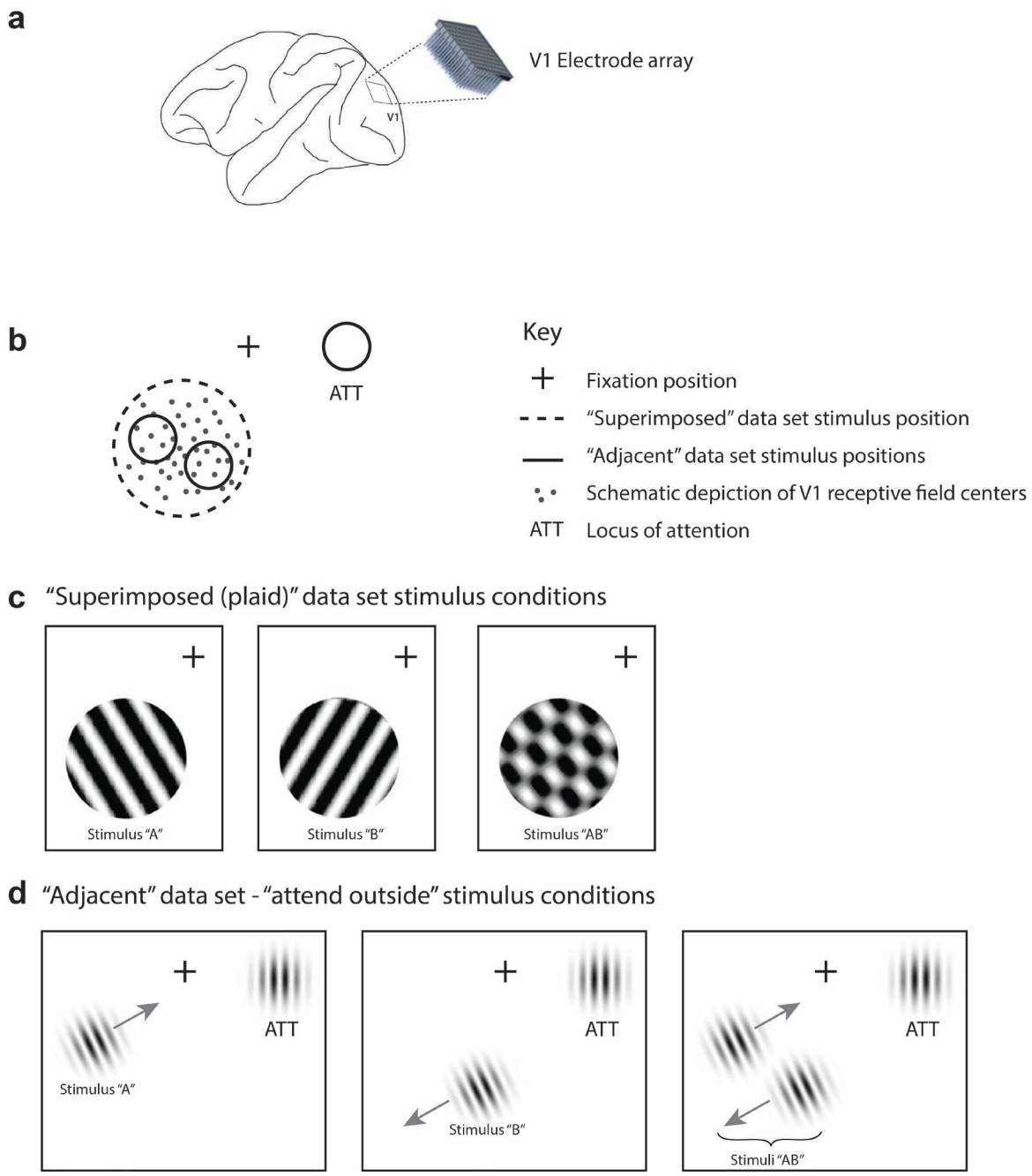
Results

General experimental design and theoretical rationale

The activity of neurons in V1 was recorded in two experimental designs in four monkeys (N=2 per experiment) using chronically implanted multielectrode arrays (Fig 1a). The original experiments are described in (Ruff, Alberts et al. 2016) and (Ruff and Cohen 2016). In both experimental datasets, drifting gratings were presented at locations that overlapped with the receptive fields of the recorded V1 neurons (Fig 1b). In one dataset ("superimposed"), the gratings were large, spanning the receptive fields of the recorded V1 neurons, and were presented either individually or in combinations of two orthogonal gratings at a consistent location on every trial (Fig 1c). When the two gratings were presented, they superimposed and formed a "plaid" stimulus. The monkeys passively fixated during these stimulus presentations.

bioRxiv preprint doi: https://doi.org/10.1101/777912; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

In the other dataset ("adjacent"), the stimuli were smaller Gabor patches and were presented either individually or adjacent to one another. Together they spanned the receptive fields of the V1 neurons being recorded in a fashion similar to the "superimposed" experiment (Fig 1d). In this experiment, the monkeys fixated while performing an orientation change discrimination task involving either one of these stimuli or a third stimulus presented in the ipsilateral hemifield. We focused our analyses on trials in which the monkeys attended to the third stimulus and judged its orientation, i.e. attention was consistently directed away from either of the two adjacent Gabor patches that elicited responses in the neurons under study. Trials in which the monkey was required to attend to one or the other of the adjacent Gabor patches were excluded from the analyses.



bioRxiv preprint doi: https://doi.org/10.1101/777912; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 1. Experimental design. A. Multiunit activity was recorded in V1 using chronically implanted 10x10 electrode arrays in four monkeys, two for each dataset. B. In both datasets, the stimuli were positioned to overlap with ("adjacent" dataset) or completely span ("superimposed" dataset) the centers of the receptive fields of the V1 neurons that were recorded. C. In the "superimposed" dataset, gratings were presented either individually or in combination at a consistent location and were large enough to cover the V1 receptive fields (stimulus size range: 2.5-7o). The combined gratings appeared as a plaid (rightmost panel). Monkeys maintained fixation throughout stimulus presentation and performed no other task. D. In the "adjacent" dataset, Gabor patches were smaller (typically <10) and were presented individually or side-by-side roughly covering the region of the V1 receptive fields. Monkeys maintained fixation while performing an orientation change detection task. The data analyzed in this study involved trials in which the monkeys were attending a third Gabor patch located in the ipsilateral hemifield to perform the orientation change detection.

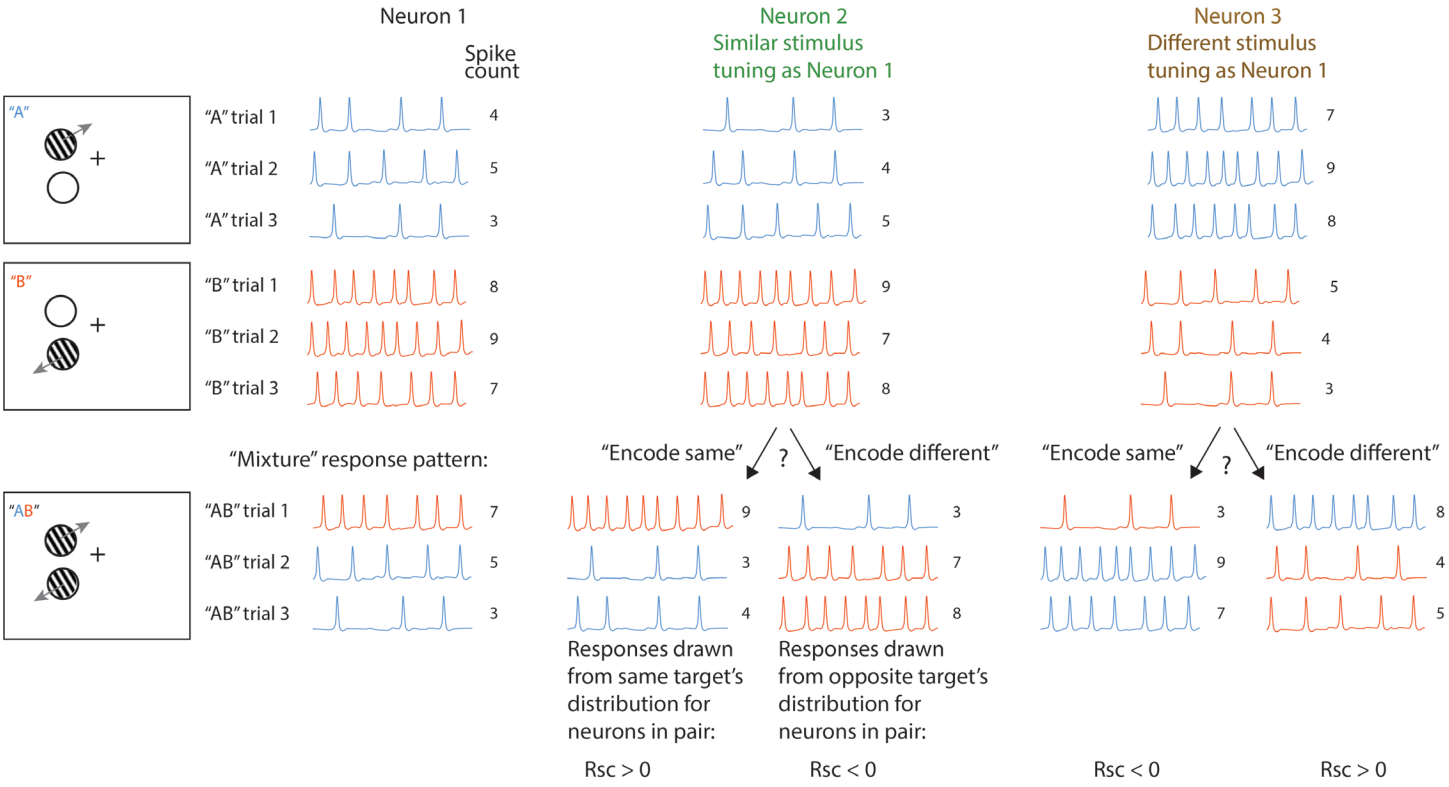


Figure 2. Schematic depiction of possible response patterns and interpretation. Three hypothetical neurons and their possible spiking activity on individual trials are shown, for three trials each of stimulus "A" alone, stimulus "B" alone, and stimuli "A" and "B" together. If neurons "pick" from their "A"-like and "B"-like response distributions when "A" and "B" are presented together, then several possible patterns of correlation might emerge between neurons, depending on whether the individual neurons in the pair "picks" the same stimulus ("Encode same") or different stimuli ("Encode different") and whether they share similar (Neurons 1 and 2) vs. different stimulus preferences (Neurons 1 and 3).

The possible results and their interpretation are illustrated schematically in Figures 2 and 3. For simplicity, we show only the "adjacent" stimulus case (left column), with the two unattended adjacent Gabor patches "A" and "B", and we focus on hypothetical neurons that appear to fluctuate between responding to "A" and responding to

bioRxiv preprint doi: https://doi.org/10.1101/777912; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

"B" on different trials. A statistical method for formally identifying such patterns is presented in Figure 5 and associated text (Caruso, Mohl et al. 2018, Mohl, Caruso et al. 2019).

In Figure 2, the column labeled "Neuron 1" depicts how one such neuron might respond to the "A" and "B" Gabor patches presented individually or in combination. For example, this neuron might respond rather modestly to "A", with spike counts ranging from 3-5 spikes (blue) across a series of 3 trials. When "B" is presented, perhaps it responds more vigorously, with 7-9 spikes (red). When "A" and "B" are presented in combination, it is theoretically possible that Neuron 1's spike counts are drawn from either its "A-like" or "B-like" spike count distributions and which distribution is somehow selected varies across trials. These spikes are illustrated in red and blue to indicate the distribution they are presumably drawn from. This is referred to as a "mixture" pattern. It is important to note that the mixture pattern depicted in Figure 2 is an exaggeration of what is likely to be found in the neural data, where actual response patterns can rarely be this easily identified as likely drawn from one distribution vs. another. It is also important to note that neurons whose responses are demonstrably consistent with fluctuating between A-like and B-like response patterns may represent a small fraction of the overall population of neurons (and will tend to be strongly normalized as previously described in this data set (Ruff, Alberts et al. 2016)). These points will be revisited in greater detail below.

With these caveats in mind, let us now consider a second neuron, Neuron 2. Suppose this neuron, like Neuron 1, responds more vigorously to "B" (red spikes) than to "A" (blue spikes), and also shows activity that appears drawn from either its A-like or B-like response distributions on combined AB trials (bottom groupings of red and blue spike trains). We can then ask whether, on trials when Neuron 1 "picks" A, does Neuron 2 also tend to "pick" A? And when Neuron 1 "picks" B, does Neuron 2 also tend to "pick" B?[1] If so, then there would tend to be a positive spike count correlation between Neuron 1 and Neuron 2 on combined stimulus trials ("encode same" branch). Note that this correlation could be larger than customary values of spike count correlations typically observed on single stimulus trials, depending on how reliably the two neurons "choose" the same stimulus and how different their firing rates are on "A" vs. "B" trials.

In contrast, if Neuron 2 and Neuron 1 tend to "pick" differently, with Neuron 2 exhibiting spike counts that appear to be drawn from its "B"-like distribution on trials that Neuron 1 exhibits spike counts that appear drawn from its "A"-like distribution and vice versa ("encode different" branch), then the spike count correlations between this pair on combined stimulus trials will be negative, and could be strongly so depending again on the reliability of

[1] We set aside as beyond our present scope what it means to "pick" one or the other distribution and what the neural mechanism that might accomplish such an operation might look like; see Caruso et al. (2018) for discussion of this point.

bioRxiv preprint doi: https://doi.org/10.1101/777912; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

the underlying anticorrelation together with the difference in the response distributions on the “A” and “B” trials alone.

Now imagine a third neuron, Neuron 3, which shows an opposite stimulus selectivity as Neurons 1 and 2: this neuron responds more vigorously to “A” (blue spikes) than to “B” (red spikes). If this neuron also fluctuates between “A”-like activity and “B”-like activity on the combined “AB” trials, it too can be either positively or negatively correlated with Neuron 1. If it “encodes same” (left branch), its spike count correlations will be *negative* in sign (when paired with Neuron 1). If it “encodes different”, it will now be *positively* correlated in spike count with Neuron 1: The larger number of spikes associated with stimulus A in Neuron 3 will tend to occur on trials in featuring a large number of spikes associated with stimulus B in Neuron 1, and with the smaller numbers of B-like spikes for Neuron 3 co-occurring with smaller numbers of A-like spikes occurring for Neuron 1.

In short, positive vs. negative spike count correlations in response to combined stimuli will have different interpretations depending on whether the two neurons in the pair both respond more vigorously to the same component stimulus or to different component stimuli.

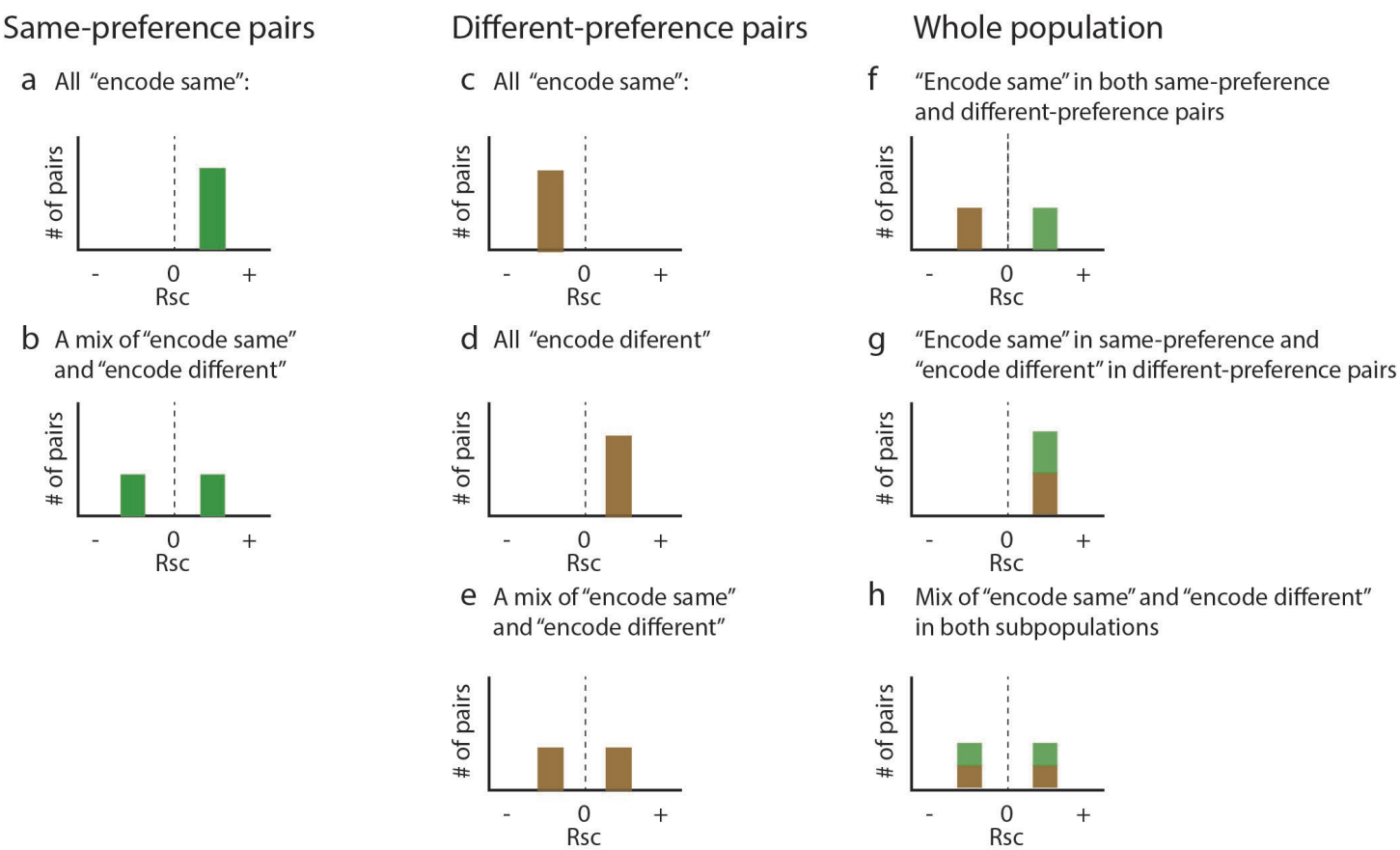


Figure 3. Population-level predictions under different scenarios. Schematic histograms of the possible patterns of spike count correlations across all the pairs of recorded neurons as a function of whether they share the same

bioRxiv preprint doi: https://doi.org/10.1101/777912; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

stimulus preferences or have different preferences, and whether they tend to "encode" the same stimulus or different stimuli on each trial.

The possible patterns of spike count correlations across a population of pairs of neurons are illustrated in Figure 3. If the population of pairs of neurons that prefer the same individual stimulus reliably follow the "encode-same" pattern, then the distribution of spike count correlations for this subpopulation should be positively skewed (Fig 3A). If the pairs of same-preference neurons include a mixture of "encode-same" and "encode-different" response patterns, then the overall distribution will range from negative to positive (Fig 3B). This is depicted here as distinct positive and negative bars, but it could easily simply appear as a widely spread distribution.

Finally, we note that it is not possible for all same-preference neuron pairs to involve "encode-different" response patterns for the two-stimulus case. Imagine another neuron, Neuron 4, which is like Neurons 1 and 2 in its response pattern to the individual stimuli "A" and "B". If Neurons 1 and 2 "encode different" and Neurons 2 and 4 also "encode-different" in relation to each other, then Neurons 1 and 4 will "encode-same" with respect to each other. Neurons 1 and 2 will be negatively correlated and neurons 2 and 4 will be negatively correlated on "AB" trials, but Neurons 1 and 4 will tend to be positively correlated.

In contrast, if "different preference" pairs tend to "encode-same", then the overall distribution of spike count correlations in this subgroup will be skewed towards negative values (Figure 3C), whereas if they tend to "encode-different", the distribution will skew positive (Figure 3D). A mixture of "encode-same" and "encode-different" will be a broader distribution ranging from negative to positive, depicted here schematically as two distinct categories for simplicity (Figure 3E).

Next we consider how the "same preference" and "different preference" subpopulations might look when combined. If both subgroups preferentially "encode same", then there will be two distinct distributions, negatively-skewed in the "different preference" subgroup and positively skewed in the "same preference" subgroup (Figure 3F). "Encode same" among "same preference" and "encode different" among "different preference" would yield an overall positively skewed distribution (Figure 3G), and a mixture of different strategies in both subpopulations would yield a mix of positive and negative values for both subgroups (Figure 3H).

It should be noted that it is likely that all spike count correlations between pairs of neurons ride on an overall wave of at least slight positivity, due to shared sensitivity to non-stimulus-related factors like overall arousal level or satiety-related signals that might accompany task performance. Thus, the negative- and positive-modes of a broad distribution may not be symmetric around 0 but slightly shifted towards the positive side.

With two objects, distinct distributions of positive and negative spike count correlations occur

bioRxiv preprint doi: https://doi.org/10.1101/777912; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

We now turn to the actual results, with these predictions in mind. We examined the Pearson's correlation between the trial-wise spike count responses of pairs of single or multiunits to the superimposed and adjacent stimuli trials in each data set. This spike count correlation (r_{sc}), also commonly called a noise correlation, was computed for a given pair of units and for a given single stimulus or combination of simultaneously presented stimuli. Each pair of units contributed multiple r_{sc} values to the population distribution, one value for each relevant stimulus condition. See methods for additional details.

Figure 4 A-F illustrates results involving three example units from the adjacent-stimuli dataset, selected for showing spike count distributions on dual trials (black lines) that reflect a mixture of the distributions evident on the corresponding single stimulus trials (red and blue lines; the formal test for "mixtures" is described in next section). The dual-stimulus distributions of spike counts are over-dispersed compared to what would be expected if the spikes on dual stimulus trials were generated from a similar Poisson process as the single stimulus trials, and a tendency for bimodality with modes near the modes for each of the individual stimulus trials is evident.

The units in panels a and b share a similar stimulus preference to each other: stimulus "A", elicits higher spike counts (red line) than stimulus "B" (blue line) for both. The unit in panel c has the opposite preference, with higher spike counts for "B" than for "A". Panels D-F show the activity of each of these units on individual trials plotted against the others. The pattern of spike count correlation on individual trials varies depending on the stimulus tuning preferences, with the one pairing between units that shared the same preference being positive (0.56, panel d) and the two pairings involving different tuning preferences yielding negative spike count correlations (-0.45, -.34, panels e and f). This pattern is borne out when the full set of pairings involving the unit in (a) and other units recorded at the same time that also showed "mixture" response patterns is considered (panel g): all the pairings that involve "same" tuning preferences yield positive correlations, and all of these correlations are individually significant (green bars, $p<0.05$). In contrast, all the pairings that involve "different" tuning preferences yield negative correlations (brown bars); as expected, these are slightly more weakly negative than the "same" pairings are positive, but 1 of 5 reaches individual significance (darker brown, $p<0.05$).

While it is visually evident that the spiking responses of these three example units on combined AB trials appear drawn from a mixture of the A-like and B-like response distributions, evaluating this systematically across the population requires a formal statistical assessment. We developed such an assessment in our previous work concerning fluctuating activity in the context of encoding of multiple simultaneous stimuli (Caruso, Mohl et al. 2018, Glynn, Tokdar et al. 2019, Mohl, Caruso et al. 2019).

bioRxiv preprint doi: https://doi.org/10.1101/777912; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

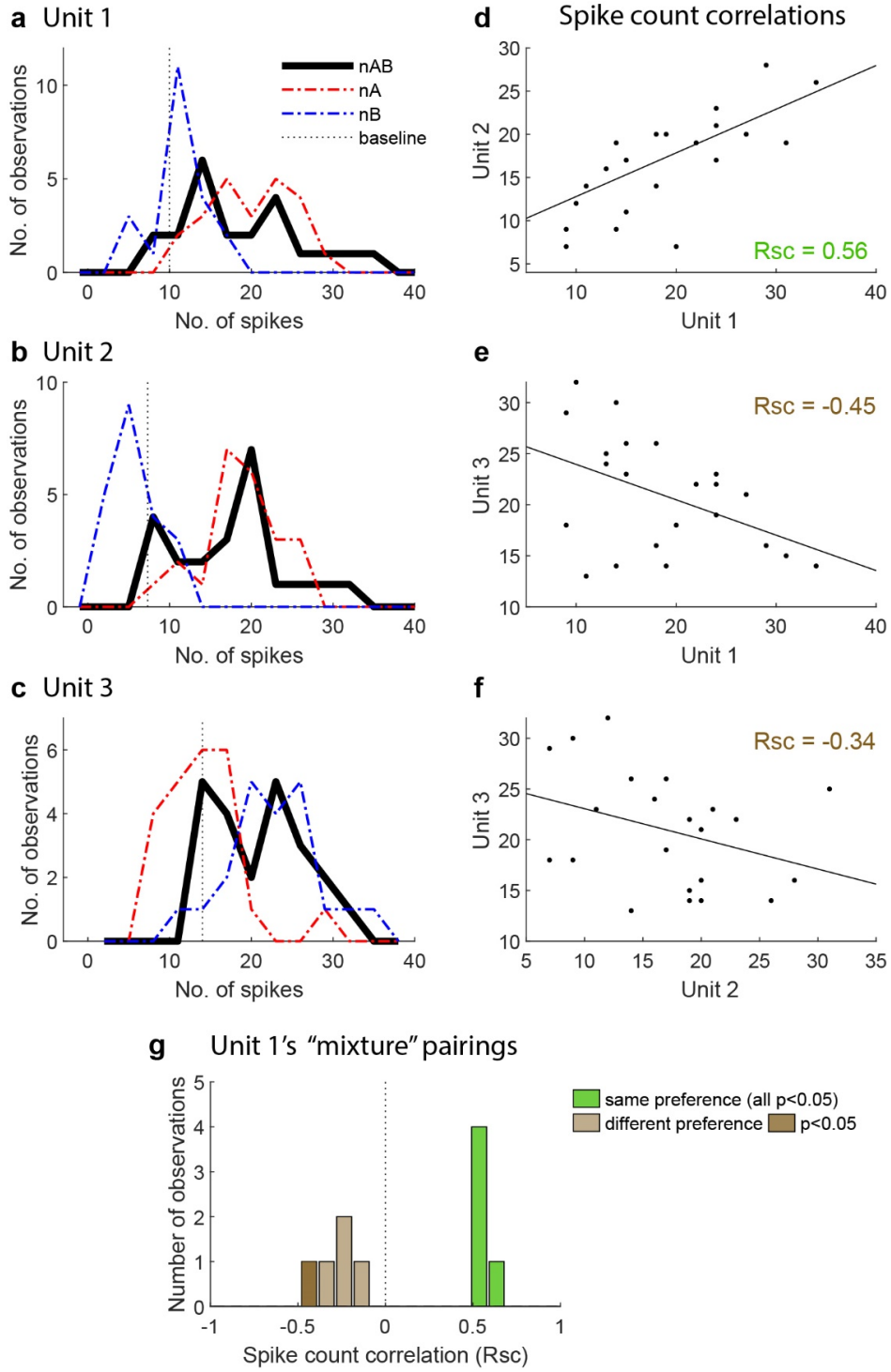


Figure 4. Example units and correlation patterns. A-C. Distribution of spike counts on single stimuli (red, blue) and dual adjacent stimulus trials (black) for three units recorded simultaneously, selected for showing a bimodal/overdispersed spike count pattern on the dual stimulus trials ("mixtures", see main text). The units in panels A and B shared a similar stimulus preference (the red curve is shifted to the right of the blue curve for both), whereas the unit shown in C had a different stimulus preference compared to A and B (red curve is to the left of the blue curve). Dashed lines indicate the baseline activity for these units. D-F. Spike counts on individual trials for the two units that shared a similar tuning preference exhibited positively correlated spike count variation on individual trials for the dual stimulus condition (green arrow and Rsc value), whereas both exhibited a negative correlation with the differently-tuned unit in c (brown arrows and Rsc values). D. Distribution of Rsc values for

bioRxiv preprint doi: https://doi.org/10.1101/777912; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

a subset of the units recorded simultaneously with the unit in (A), selected for also showing a spiking pattern that appears to be a mixture of the two individual stimulus response patterns, and color coded according to whether the stimulus preference was the same (green) or different (brown). All of the "same" pairs exhibited positive correlations, and 5 of 5 were individually significant (p<0.05). All of the "different" pairs exhibited negative correlations, and 1 of 5 was individually significant (p<0.05).

In particular, we can model the firing rate behavior of neurons when two simultaneous grating stimuli A and B are presented in relation to the firing rates that occur when stimuli A and B are presented individually. We assume that each single-stimulus condition has the firing rate distribution following a single Poisson distribution and we exclude cases where this assumption is violated. We use a Bayesian model comparison framework to consider four hypotheses concerning the combined AB trials (Figure 5A): 1) The responses to A and B together appear drawn from the same distribution as either A or B and consistently so on every trial, as if the unit responded to only one of the two stimuli ('*single'*). 2) Responses to A and B together appear drawn from a distribution '*outside'* the range spanned by the A and B response distributions; this is the predicted pattern if neurons were to linearly sum the A and B responses or if one stimulus strongly suppressed the response to the other. 3) The responses to A and B together are drawn from a single distribution with a mean at an '*intermediate'* value between the A-like and B-like response rates. This is the response pattern that would be expected under theories such as divisive normalization in which the responses of an individual neuron to a more favored stimulus are reduced when other stimuli are also present. 4) The responses to A and B together appear to be drawn from a '*mixture'* of the A-like and B-like response distributions. Mixtures are the category of interest for this analysis, as they indicate the presence of activity fluctuations on the trial-by-trial timescale.

We first consider the overall pattern of results for the adjacent vs. superimposed stimulus datasets (Figure 5 B-C). The most obvious difference between these two datasets is the dominance of "singles" in the adjacent stimulus dataset vs. "outsides" in the superimposed dataset. This suggests that the most common pattern of response in the adjacent gratings case was to respond to only one of the two gratings, whereas in the superimposed gratings dataset responses were generally greater in response to the plaid than to either grating presented alone. This difference almost certainly stems from differences in the size of the stimuli being presented in these two datasets -- typically only one of the two adjacent gratings was located within the receptive field of a given V1 unit whereas this was not the case for the superimposed dataset. This "single" vs. "outside" difference, then, is a side note to our main focus on any possibly fluctuating activity patterns.

We found that "mixtures" were not evident in V1 when only one "object" was present, i.e. the superimposed gratings dataset (Figure 5B, rightward bars). Among the 1252 triplets analyzed, the majority of the neurons were best explained by the 'outside' or 'single' hypotheses, while there were only 26 triplets that are in support of the 'mixture' hypothesis (2%). These results indicate that fluctuating activities are not evident in V1 when two stimuli are superimposed one another and likely perceived as one object.

bioRxiv preprint doi: https://doi.org/10.1101/777912; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

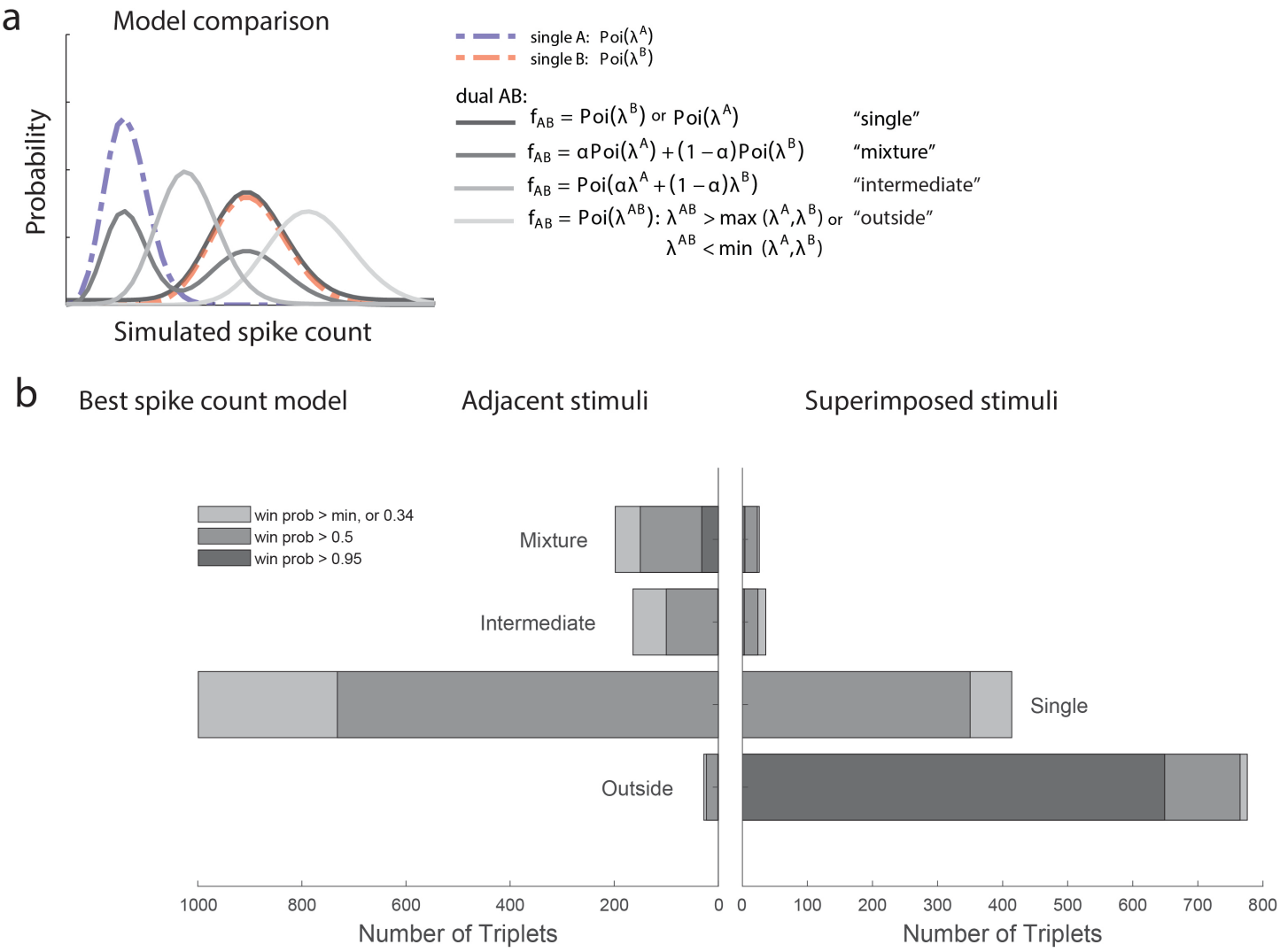


Figure 5. Bayesian model comparison regarding spike count distributions and results. A. We evaluated the distribution of spike counts on combined stimulus trials in relation to the distributions observed on when individual stimuli were presented alone. Four possible models were considered as described in the equations and text. Only one case of "single" (B-like) and "outside" ($\lambda^{AB} > \max(\lambda^A, \lambda^B)$ are shown. B. Best spike count models for the adjacent and superimposed stimulus datasets. Shading indicates the confidence level for the wining model, with darker shading indicating a higher probability that the chosen model is correct. While "singles" dominated in the adjacent stimulus dataset and "outsides" dominated in the superimposed stimulus datasets, we focus on the presence of a "mixtures" as an important minority subpopulation present in the "adjacent" stimulus dataset.

However, a substantially greater (though still overall modest) number of mixtures did appear when two "objects" were present: the two adjacent stimulus dataset yielded 198 out of 1389 "mixtures", or 14% (Figure 5B left panel). Although the proportion of conditions showing mixtures of activity patterns remained small (as expected given that V1 has comparatively small receptive fields which tends to minimize the evoked-activity overlap problem between multiple objects), it was a sufficiently large subgroup to permit analyses of spike count correlations between pairs of neurons that were both classified as mixtures.

bioRxiv preprint doi: https://doi.org/10.1101/777912; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Focusing now on the formally identified "mixtures" subgroup in the adjacent stimulus dataset (Figure 6A), we can see that the pattern observed for the example cells in Figure 4 holds at the population level: neural pairs in which both units responded better to the same individual stimuli tended to have positive correlations with each other (green bars, median 0.486), whereas those that had different stimulus preferences tended to exhibit negative correlations (brown bars, median -0,14). The spread of values is broad, with many of the pairs of "different-preference" neurons in particular exhibiting positive values, (a point we will return to in Figure 8).

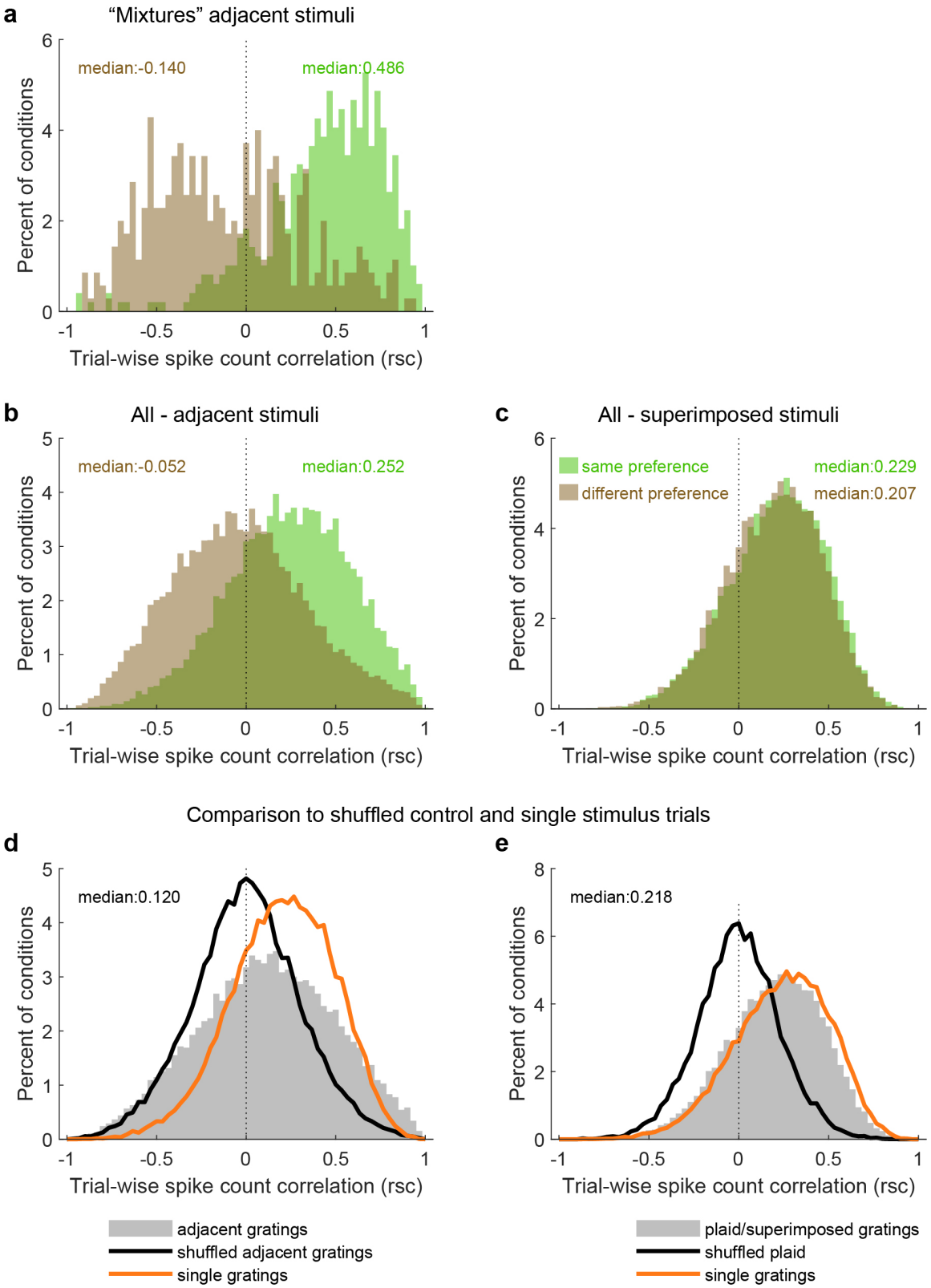
Because of the lack of an adequate population of "mixture"-classified pairs in the superimposed gratings dataset, we next compared the populations as a whole (Figure 6B-C). The patterns are quite different between these two datasets. In the adjacent-stimulus dataset, the overall broad distribution and distinction between same-preference and different-preference subgroups holds even when not selecting for "mixture" fluctuating patterns (green bars vs. brown bars, median rsc 0.252, -0.052). However, this is much less true of the superimposed-gratings dataset: here, there is very little difference between the "same-preference" and "different preference" pairs of neurons (median "same-preference" rsc = 0.229, median "different-preference" rsc = 0.207), nor is there much difference between the spike count correlations observed on dual-gratings trials vs. individual gratings trials for this dataset (Figure 6E). In contrast, there is a distinct difference between spike count correlations observed on the dual-stimuli vs. individual-stimulus trials in the adjacent-stimulus dataset (Figure 6D).

We next considered whether this overall pattern was robust to the classification categories emerging from the Bayesian model comparison, especially given that "singles" reflect the largest category in the adjacent stimulus dataset. Figure 7 illustrates the median spike count correlation by model classification category, for "same-preference" and "different-preference" pairs of neurons. We excluded the "outside" category from this analysis as there were too few units that were classified as such. We found that all 6 remaining combinations of classifications yielded positive median spike count correlations among "same preference" pairs and negative correlations among "different preference pairs". Thus, the overall pattern of results described above does not rest critically on the particular details of the model comparison we implemented here, and is present even among units that could not be formally shown to be fluctuating fully between "A"-like and "B"-like response distributions.

Returning to the predictions laid out in Figure 3, the implication of same-preference units being on average positively correlated and different-preference units being on average negatively correlated from a coding perspective is that V1's representation (among "mixture" units) may be slightly biased towards one or the other stimulus on each individual trial, most closely resembling the schematic depiction in Figure 2F. However, it should be noted that the actual data involves a broad distribution with positive spike count correlations among the different-preference pairs and negative spike count correlations among the same-preference pairs, which is therefore also consistent with the schematic depiction in Figure 2H, and suggests that while the overall pattern of activity among "mixture" units is biased towards one stimulus over the other on individual trials, there are

bioRxiv preprint doi: https://doi.org/10.1101/777912; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

ample cases of units that do not follow this pattern, and these exceptions may be sufficient to preserve information about the other stimulus.



bioRxiv preprint doi: https://doi.org/10.1101/777912; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Figure 6. Patterns of spike count correlations among pairs of V1 neurons in different subgroups and conditions. A. Neural pairs in which both units met the "mixture" classification showed distinct positive and negative patterns of correlation in response to adjacent stimuli. Positive correlations were more likely to occur among pairs of neurons that responded more strongly to the same individual stimuli (green bars, median rsc = 0.486), and negative correlations were more likely to occur among pairs of neurons that responded more strongly to different individual stimuli (brown bars, median rsc = -0.14). B. This pattern of results held even when all the unit pairs were considered in aggregate ("same preference" pairs, median rsc = 0.252; "different preference" pairs, median rsc = -0.052). C. However, among pairs recorded during presentation of superimposed gratings, this pattern was not apparent: unit pairs tended to show positive correlations in both cases ("same preference" median rsc =0.229, "different preference" median rsc = 0.207). D, E. Comparison of the rsc values for pairs of neurons recorded during presentation of stimulus combinations vs. individual stimuli (orange traces) and vs. a shuffled control. The correlations observed in the adjacent stimulus dataset for adjacent Gabor stimulus presentations differed from those observed for individual stimulus presentations (panel D, gray bars vs. orange line), whereas this was much less the case for the superimposed gratings dataset (panel E, gray bars vs. orange line).

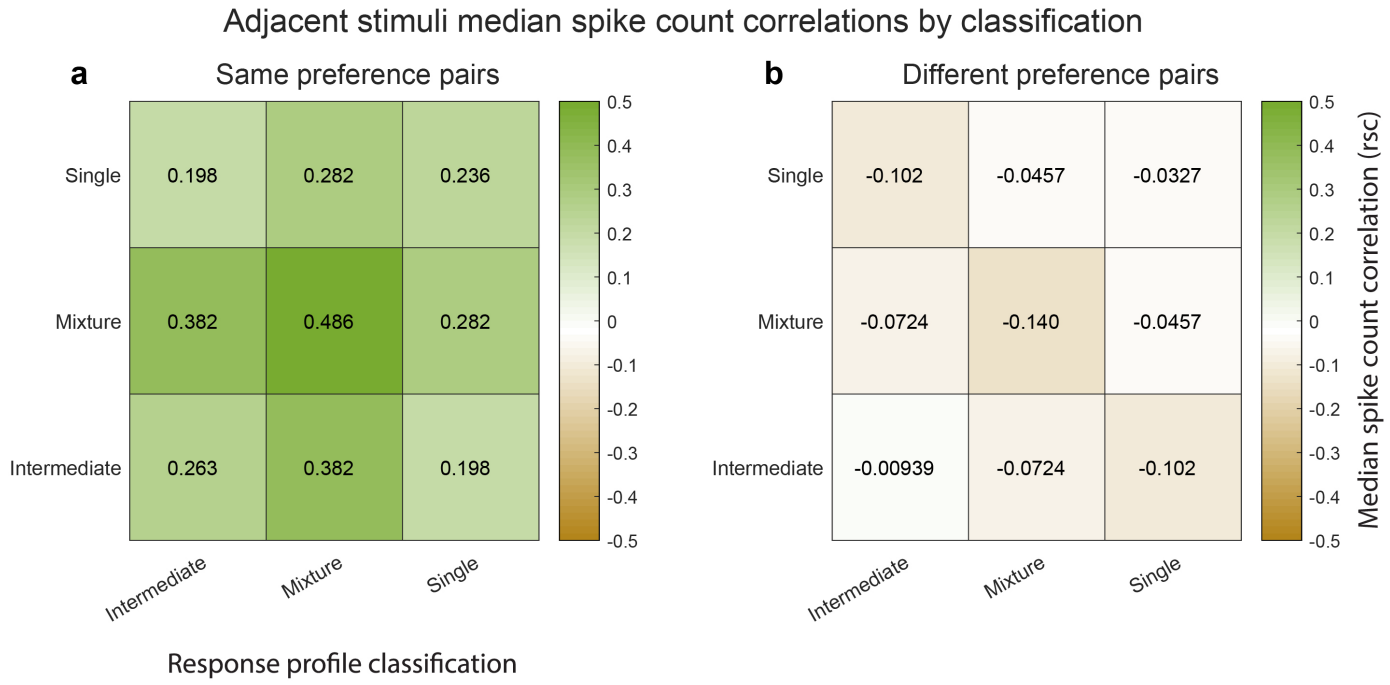


Figure 7. Median spike count correlations as a function of same-different preference (panel a vs. panel b) and as a function of the response profile classification resulting from the Bayesian model comparison for the adjacent stimulus dataset. The "mixture"-"mixture" combinations produced the strongest positive (same preference pairs) and strongest negative (different preference pairs) median spike count correlations, but all other combinations also involved positive median correlations for same preference pairs and negative median correlations for different preference pairs.

To visualize this in another way, we repeated the calculation of Pearson's correlations between pairs of unit-conditions classified as mixtures using not the spike counts on each trial but an assignment score concerning how "A"-like vs "B"-like the spike count was on an individual trial (ranging from 0 to 1; see methods). Plotted this way, a positive correlation indicates that the two units in the pair tended to exhibit response patterns consistent

bioRxiv preprint doi: https://doi.org/10.1101/777912; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

with the same object at the same time, whereas a negative correlation indicates that the two units tended to exhibit responses consistent with different objects at the same time. The overall pattern in the data is positively skewed (Figure 8), but with a long tail on the negative side, consistent with the population of units giving an edge to one stimulus over the other on each individual trial, but not to the exclusion of the other stimulus.

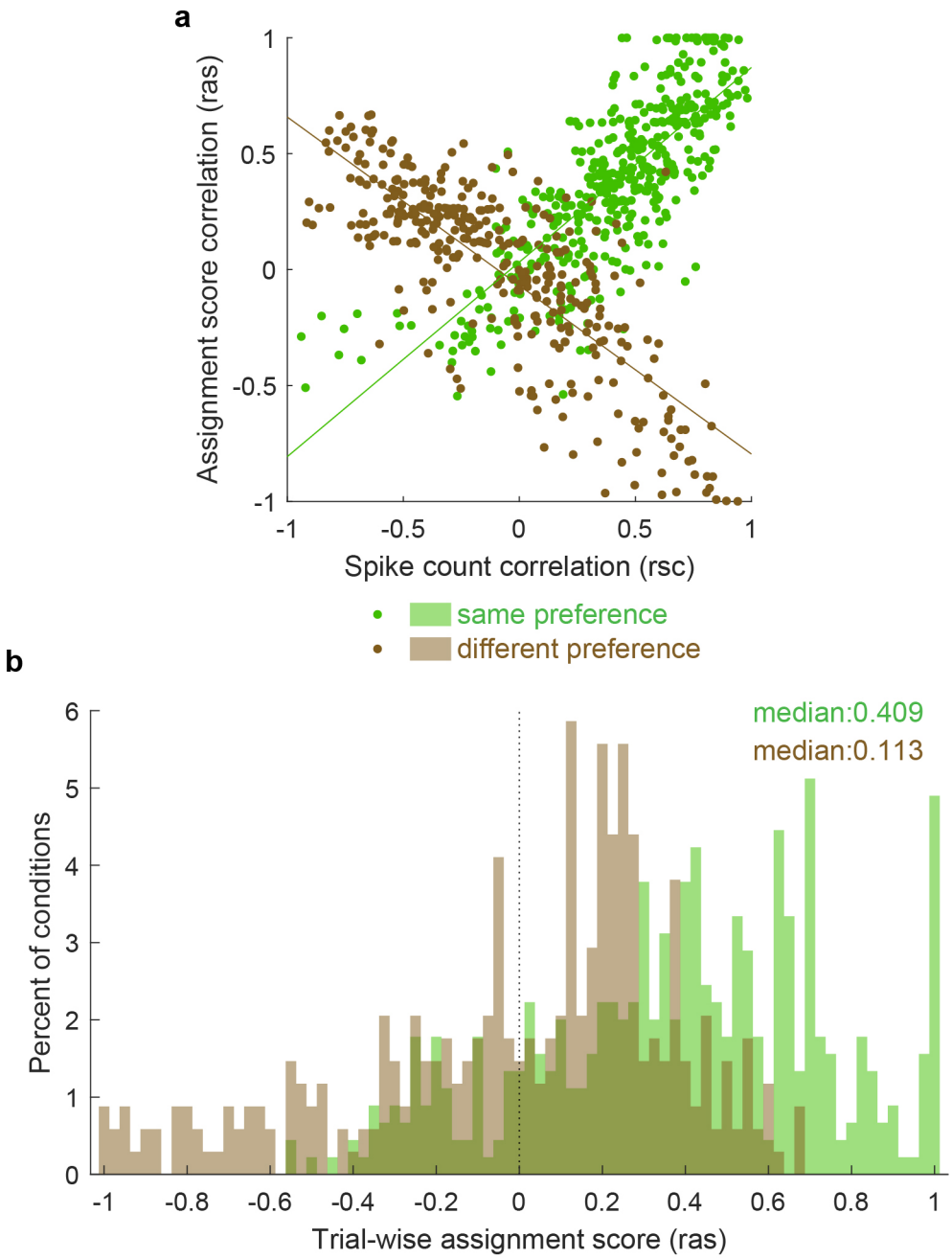


Figure 8. Trial-wise fluctuations in "mixture" pairs of unit-conditions are consistent with a bias towards both units in the pair tending to signal the same stimulus at the same time. This analysis involved the Pearson's correlation coefficients computed on trial-wise assignments scores (r_{as}), which take into account whether the response on combined "AB" trials is more "A-like" vs. "B-like". For two units that share a similar preference (e.g both respond better to A or both respond better to B), this correlation will have the same sign as the spike count correlation

bioRxiv preprint doi: https://doi.org/10.1101/777912; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

(panel a, green points, positively sloped best fit line). For two units that prefer different stimuli, this correlation will be opposite in sign to the spike count correlation (panel a, brown points, negatively sloped best fit line). The overall positive skew in the assignment score correlations for both the "same" and "different" preferring unit-condition pairs (panel b) therefore indicates a bias for the same stimulus at the same time. The long negative tail indicates this is only a bias and the other stimulus is also represented on a subpopulation of neurons.

Discussion

The central observation in this paper is that the pattern of spike count correlations observed in V1 differs based on the stimulus/stimuli that are presented and whether the two neurons in a pair share the same or different tuning preferences for those stimuli. Superimposed gratings yielded correlation patterns very similar to previous reports in the literature (Cohen and Kohn 2011, Ruff, Alberts et al. 2016, Ruff and Cohen 2016), as well as being very similar to trials involving single gratings, and the correlations did not greatly depend on whether the two units in the pair preferred the same individual stimulus or different ones. At the population level, there was little evidence that units exhibited "mixture"-type response patterns on superimposed gratings trials, i.e. fluctuating across trials between the responses observed in conjunction with each grating alone.

In contrast, when two stimuli were presented adjacent to one another other, the spike count correlations differed from the superimposed-stimulus dataset. In this dataset, two distinct distributions emerged based on whether the two units in the pair preferred the same individual stimulus (associated with generally positive spike count correlations) vs different individual stimuli (associated with generally negative spike count correlations). This pattern was observed in the population as a whole, but was especially pronounced in the subset of units that exhibited "mixture"-type response patterns indicating fluctuating across trials between the response distributions associated with each of the individual stimuli.

We interpret these observations under the conceptual framework of the challenge that V1 faces when representing a visual scene that contains either individual stimuli, combinations of stimuli that bind to form a single object, or combinations of stimuli that remain perceptually distinct from each other. The pattern of positive and negative correlations exhibited between pairs of such units is consistent with a population code biased towards one of the two stimuli on any given trial, but that preserves information about the other stimulus as well.

There has been a rich literature concerning correlated activity between visually-responsive neurons in recent decades. One school of thought considers correlations in the context of the variability of neural firing. Under this "noise correlation" view, positive correlations have historically been seen as detrimental for encoding information at the population level (Shadlen and Newsome 1994, Zohary, Shadlen et al. 1994). Such views have also seen notable refinement and qualification since these early studies (Romo, Hernandez et al. 2003, Averbeck and Lee 2004, Averbeck, Latham et al. 2006, Moreno-Bote, Beck et al. 2014, Kanitscheider, Coen-Cagli et al. 2015, Kohn, Coen-Cagli et al. 2016). Arguably closer to the current work is a different school of thought, the

bioRxiv preprint doi: https://doi.org/10.1101/777912; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

temporal correlation hypothesis (Milner 1974, Gray and Singer 1989, Von Der Malsburg 1994, Singer and Gray 1995, Gray 1999). This theory focused on the need to connect the brain's representation of different attributes of a given object together, and proposed that such binding might be mediated through precise synchrony of spikes among neurons responding to the same object. This view, then, sees correlated activity as both useful and specifically relevant to object vision. Studies exploring this hypothesis have, however, primarily focused on within-trial temporal synchrony of spikes on the order of milliseconds, whereas the noise correlation literature has focused on spike counts in the domain of hundreds of milliseconds and analysis at the level of the ensemble of trials. By evaluating trial-wise spike-count variation and comparing the results as a function of the number of stimuli/objects, the present work forges a bridge between these two areas of the literature.

It should be noted that the adjacent-stimulus and superimposed-stimulus datasets also differed along other dimensions besides whether the gratings were likely perceived as one object or two. In particular, the use of an attentional task differed across these two datasets. Indeed, our results are consistent with the hypothesis that on the timescale of trials (or likely fixations in natural viewing), populations of neurons shift focus from one stimulus to another when the two stimuli are perceived as separate objects (adjacent condition) but not when they are perceived as the same object (superimposed condition). One possibility is that despite the attentional target in the ipsilateral hemifield, the animals were also switching attention between the two objects in the V1 receptive field zone, which could contribute to the observed activity patterns (Ecker, Denfield et al. 2016, Engel, Steinmetz et al. 2016, Denfield, Ecker et al. 2018). Mildly arguing against this interpretation is that all of the analyzed trials were correctly performed, indicating that the monkey was maintaining attention in the ipsilateral hemifield on those trials. Since "mixtures" were identified on the basis of the spike counts fluctuating between the "A"-like and "B"-like trials at the whole-trial time scale, they might have led to errors in performing those trials correctly if such fluctuations corresponded fully to attentional focus.

Neural responses to attended and unattended stimuli can often be modeled as a weighted combination of the responses to those stimuli when presented alone (Boynton 2009, Lee and Maunsell 2009, Reynolds and Heeger 2009, Ni, Ray et al. 2012, Ni and Maunsell 2017, Ni and Maunsell 2019). Furthermore, the magnitude of the attention effects has been found to be correlated with the values of those weights, with the strongest effects of attention occurring in neurons that appeared to exhibit averaging responses (Lee and Maunsell 2009, Ni, Ray et al. 2012, Ni and Maunsell 2017, Ni and Maunsell 2019). Such averaging responses (see also Xiao, Niu et al. 2014, Xiao and Huang 2015) would be most consistent with the mixture and intermediate classification we used here, and may actually indicate fluctuations occurring on either the trial or sub-trial time scales. In future studies, it will be interesting to determine the relationship between the fluctuating responses we show here, response normalization, and attention.

bioRxiv preprint doi: https://doi.org/10.1101/777912; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Recent work suggests another more perceptual tie between studies of attention and the findings we report here. The likelihood of detecting a brief near-threshold visual stimulus varies with the phase of the brain wave oscillations at the time the stimulus is presented (Busch, Dubois et al. 2009, Busch and VanRullen 2010, Vanrullen, Busch et al. 2011, Fiebelkorn, Saalmann et al. 2013, Fiebelkorn, Pinsk et al. 2018, Helfrich, Fiebelkorn et al. 2018, Fiebelkorn and Kastner 2019) (see also Engel, Steinmetz et al. 2016). This might reflect a perceptual consequence of a brain mechanism in which neurons are slightly biased towards representing some stimuli in the visual scene over others in a naturally-occurring oscillatory fashion. In our previous study (Caruso, Mohl et al. 2018), we found that the LFP signal prior to stimulus onset was predictive of whether neurons would "pick" A vs. B on a given trial. Future work will be needed to ascertain whether a similar phenomenon occurs in V1.

Finally, it is worth noting here that considering how the brain preserves information about *two* visual stimuli presented is still a far cry from understanding how the myriad details present in a natural scene are encoded. When the number of objects gets too great, it is unlikely that neurons can fluctuate between all of them, and this is likely to have consequences for perception, perhaps accounting for well-known limits on the number of objects we can attend to and remember (e.g. Miller 1956). Future studies incorporating many stimuli and investigating how this changes the pattern of fluctuating activity and correlations between units are needed to shed light on how our brains operate outside the rarefied environment of the laboratory.

Methods

Electrophysiological recordings and visual stimuli

The full experimental procedures are described in Ruff et al.(2016) and Ruff & Cohen (2016) and summarized below. All animal procedures were approved by the Institutional Animal Care and Use Committees of the University of Pittsburgh and Carnegie Mellon University. Each of the two data sets consisted of multielectrode recordings from the V1 of two adult male rhesus monkeys (different animals for each data set). Recordings were made using a chronically implanted a 10 x 10 microelectrode array (Blackrock Microsystems) in V1. The electrode shafts were 1mm long, and the minimum distance between the nearest electrodes was 400 μm. In some sessions, recordings were also made using other electrodes in areas MT and V4, but these data are not included in the current analyses.

The visual stimuli and behavioral experiment for the superimposed stimulus dataset are fully described in Ruff et al.(2016). Monkeys were rewarded for passively viewing individual or superimposed orthogonal drifting gratings, positioned to span the receptive fields of the entire population of V1 neurons under study (size range: 2.5-7°). Trials with microsaccades were excluded from further analysis. In the full dataset, multiple contrast levels were presented, most of which were not included for analysis in the present study. Here, we included

bioRxiv preprint doi: https://doi.org/10.1101/777912; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

trials in which one grating had a contrast of 0 (i.e. was not visible) and the other had a contrast of 0.5 ("A"-alone and "B"-alone cases) or both gratings had a contrast of 0.5 ("AB"). In most sessions, each stimulus lasted for 200 ms; a few sessions with 1000 ms stimuli were also included. This data set consists of 3605 single and multi-units recorded over 65 experimental sessions.

The visual stimuli and behavioral experiment for the adjacent stimulus dataset are fully described in (Ruff and Cohen 2016). The animals performed an attentional task in which they were cued in blocks of trials to attend to small drifting Gabor patches (~1°) in various locations and respond when the orientation of the attended location changed. In the current study, we analyzed trials in which attention was directed to a Gabor patch located in the hemisphere ipsilateral to the recorded V1 neurons (i.e. well away from those neurons' receptive fields). On these trials, two unattended Gabor patches were presented in close proximity to each other within the area covered by the receptive fields of the recorded V1 neurons. These patches had the same orientation but drifted in opposite directions. The patches flashed on for 200 ms and off for 200-400 ms within the V1 receptive fields, and we analyzed responses to all stimuli before the orientation change, excluding the first stimulus in every trial. Again, only correctly performed trials with no microsaccades were included for analysis.

Analysis of spike count distributions

The full description of the statistical evaluation of spike count distributions on combined stimulus trials can be found in (Caruso, Mohl et al. 2018, Mohl, Caruso et al. 2019). Briefly, we deployed a Bayesian procedure for modeling the distribution of spike counts in response to combined stimuli. Assuming that the spike counts corresponding to condition A and condition B are both Poisson-distributed with the rate parameters λ^A and λ^B respectively (and excluding exceptions, see below), the four hypotheses for the spike count distributions for condition AB consist of:

1. "**single**": A single Poisson distribution $\mathrm{Poi}(\lambda^{AB})$, where λ^{AB} exactly equals either λ^A or λ^B.
2. "**outside**": A single Poisson distribution $\mathrm{Poi}(\lambda^{AB})$, where λ^{AB} lies outside the range between λ^A and λ^B.
3. "**intermediate**": A single Poisson distribution $\mathrm{Poi}(\lambda^{AB})$, where λ^{AB} lies inside the range between λ^A and λ^B.
4. "**mixture**": A mixture of the two single-stimulus distribution with an unknown mixing rate α: $\alpha\ \mathrm{Poi}(\lambda^A) + (1-\alpha)\ \mathrm{Poi}(\lambda^B)$

For each "triplet" or combination of A, B, and AB conditions, the hypothesis with the highest posterior probability was selected, based on the intrinsic Bayes factor[cite] and using the equal prior (¼) among the four hypotheses, Jeffrey's prior[cite] for the Poisson rate parameters, and a uniform prior in [0, 1] for the mixing weight α.

bioRxiv preprint doi: https://doi.org/10.1101/777912; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Only the triplets satisfying two additional criteria are used: (1) the single-stimulus distribution follows Poisson distribution, and (2) the single-stimulus rate parameters λ^A and λ^B are substantially separated. The first criterion was tested using Monte Carlo p-value calculation for a Chi-square goodness-of-fit test, and the second criterion was tested by whether the intrinsic Bayes factor of the model $\lambda^A \neq \lambda^B$ is more than three times higher than that of the model $\lambda^A = \lambda^B$. These exclusion criteria were applied to all the analyses in the paper, even those that did not build specifically on this model classification, to ensure that comparisons between subpopulations of the data were not affected by differences in data selection criteria.

Correlation analysis

We calculated spike count correlations between pairs of units recorded at the same time in the same experiment. The Pearson correlation coefficient was calculated on the trial-wise spike counts for each relevant stimulus combination. Trials in which one or both units in the pair exhibited an “outlier” response, i.e. more than 3 standard deviations from the mean, were excluded from the analysis. The spike count correlations for particular unit pairs for different stimulus combinations were included in the population analyses as separate observations and were not averaged together. For example, in the adjacent stimulus dataset, pairs were typically tested with two separate adjacent stimulus combinations, differing in the direction of motion, potentially yielding two values of the spike count correlation (assuming both conditions passed the Poisson and response-separation exclusion criteria noted above).

Same or different preference

Preference of a unit for a particular stimulus was determined by higher spike counts. Unit pairs that both exhibited more spikes in response to stimulus A than to B, or both exhibited more spikes in response to stimulus B than to A, were defined as “same preference”. Unit pairs in which one responded with more spikes to A and the other with more spikes to B were defined as “different preference”.

bioRxiv preprint doi: https://doi.org/10.1101/777912; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

References

Adelson, E. H. and J. A. Movshon (1982). "Phenomenal coherence of moving gratings." Nature, Lond. **300**: 523-525.

Averbeck, B. B., P. E. Latham and A. Pouget (2006). "Neural correlations, population coding and computation." Nat Rev Neurosci **7**(5): 358-366.

Averbeck, B. B. and D. Lee (2004). "Coding and transmission of information by neural ensembles." Trends Neurosci **27**(4): 225-230.

Boynton, G. M. (2009). "A framework for describing the effects of attention on visual responses." Vision Res **49**(10): 1129-1143.

Buffalo, E. A., P. Fries, R. Landman, H. Liang and R. Desimone (2010). "A backward progression of attentional effects in the ventral stream." Proc Natl Acad Sci U S A **107**(1): 361-365.

Busch, N. A., J. Dubois and R. VanRullen (2009). "The phase of ongoing EEG oscillations predicts visual perception." J Neurosci **29**(24): 7869-7876.

Busch, N. A. and R. VanRullen (2010). "Spontaneous EEG oscillations reveal periodic sampling of visual attention." Proc Natl Acad Sci U S A **107**(37): 16048-16053.

Busse, L., A. R. Wade and M. Carandini (2009). "Representation of concurrent stimuli by population activity in visual cortex." Neuron **64**(6): 931-942.

Caruso, V. C., J. T. Mohl, C. Glynn, J. Lee, S. M. Willett, A. Zaman, A. F. Ebihara, R. Estrada, W. A. Freiwald, S. T. Tokdar and J. M. Groh (2018). "Single neurons may encode simultaneous stimuli by switching between activity patterns." Nat Commun **9**(1): 2715.

Cohen, M. R. and A. Kohn (2011). "Measuring and interpreting neuronal correlations." Nat Neurosci **14**(7): 811-819.

Cook, E. P. and J. H. Maunsell (2002). "Attentional modulation of behavioral performance and neuronal responses in middle temporal and ventral intraparietal areas of macaque monkey." J Neurosci **22**(5): 1994-2004.

Denfield, G. H., A. S. Ecker, T. J. Shinn, M. Bethge and A. S. Tolias (2018). "Attentional fluctuations induce shared variability in macaque primary visual cortex." Nat Commun **9**(1): 2654.

Ecker, A. S., G. H. Denfield, M. Bethge and A. S. Tolias (2016). "On the Structure of Neuronal Population Activity under Fluctuations in Attentional State." J Neurosci **36**(5): 1775-1789.

Engel, T. A., N. A. Steinmetz, M. A. Gieselmann, A. Thiele, T. Moore and K. Boahen (2016). "Selective modulation of cortical state during spatial attention." Science **354**(6316): 1140-1144.

Fiebelkorn, I. C. and S. Kastner (2019). "A Rhythmic Theory of Attention." Trends Cogn Sci **23**(2): 87-101.

Fiebelkorn, I. C., M. A. Pinsk and S. Kastner (2018). "A Dynamic Interplay within the Frontoparietal Network Underlies Rhythmic Spatial Attention." Neuron **99**(4): 842-853 e848.

Fiebelkorn, I. C., Y. B. Saalmann and S. Kastner (2013). "Rhythmic sampling within and between objects despite sustained attention at a cued location." Curr Biol **23**(24): 2553-2558.

Glynn, C., S. Tokdar, A. Zaman, V. C. Caruso, J. T. Mohl, S. M. Willett and J. M. Groh (2019). "Analyzing second order stochasticity of neural spiking under stimuli-bundle exposure." Submitted.

Gray, C. M. (1999). "The temporal correlation hypothesis of visual feature integration: still alive and well." Neuron **24**(1): 31-47, 111-125.

Gray, C. M. and W. Singer (1989). "Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex." Proc Natl Acad Sci U S A **86**(5): 1698-1702.

Heeger, D. J., E. P. Simoncelli and J. A. Movshon (1996). "Computational models of cortical visual processing." Proceedings of the National Academy of Sciences of the United States of America **93**(2): 623-627.

Helfrich, R. F., I. C. Fiebelkorn, S. M. Szczepanski, J. J. Lin, J. Parvizi, R. T. Knight and S. Kastner (2018). "Neural Mechanisms of Sustained Attention Are Rhythmic." Neuron **99**(4): 854-865 e855.

bioRxiv preprint doi: https://doi.org/10.1101/777912; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Kanitscheider, I., R. Coen-Cagli and A. Pouget (2015). "Origin of information-limiting noise correlations." Proc Natl Acad Sci U S A **112**(50): E6973-6982.

Kohn, A., R. Coen-Cagli, I. Kanitscheider and A. Pouget (2016). "Correlations and Neuronal Population Information." Annu Rev Neurosci **39**: 237-256.

Lee, J. and J. H. Maunsell (2009). "A normalization model of attentional modulation of single unit responses." PLoS One **4**(2): e4651.

Lima, B., W. Singer, N. H. Chen and S. Neuenschwander (2010). "Synchronization dynamics in response to plaid stimuli in monkey V1." Cereb Cortex **20**(7): 1556-1573.

Luck, S. J., L. Chelazzi, S. A. Hillyard and R. Desimone (1997). "Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex." J Neurophysiol **77**(1): 24-42.

Miller, G. A. (1956). "The magical number seven, plus or minus two: some limits on our capacity for processing information." Psychological Review **63**(2): 81-97.

Milner, P. M. (1974). "A model for visual shape recognition." Psychol Rev **81**(6): 521-535.

Mohl, J. T., V. C. Caruso, S. Tokdar and J. M. Groh (2019). "Sensitivity and specificity of a Bayesian single trial analysis for time varying neural signals." BioRxiv.

Moreno-Bote, R., J. Beck, I. Kanitscheider, X. Pitkow, P. Latham and A. Pouget (2014). "Information-limiting correlations." Nat Neurosci **17**(10): 1410-1417.

Ni, A. M. and J. H. R. Maunsell (2017). "Spatially tuned normalization explains attention modulation variance within neurons." J Neurophysiol **118**(3): 1903-1913.

Ni, A. M. and J. H. R. Maunsell (2019). "Neuronal Effects of Spatial and Feature Attention Differ Due to Normalization." J Neurosci **39**(28): 5493-5505.

Ni, A. M., S. Ray and J. H. Maunsell (2012). "Tuned normalization explains the size of attention modulations." Neuron **73**(4): 803-813.

Palanca, B. J. and G. C. DeAngelis (2005). "Does neuronal synchrony underlie visual feature grouping?" Neuron **46**(2): 333-346.

Reynolds, J. H. and D. J. Heeger (2009). "The normalization model of attention." Neuron **61**(2): 168-185.

Rodman, H. R. and T. D. Albright (1989). "Single-unit analysis of pattern-motion selective properties in the middle temporal visual area (MT)." Exp. Brain Res. **75**: 53-64.

Romo, R., A. Hernandez, A. Zainos and E. Salinas (2003). "Correlated neuronal discharges that increase coding efficiency during perceptual discrimination." Neuron **38**(4): 649-657.

Ruff, D. A., J. J. Alberts and M. R. Cohen (2016). "Relating normalization to neuronal populations across cortical areas." J Neurophysiol **116**(3): 1375-1386.

Ruff, D. A. and M. R. Cohen (2016). "Attention Increases Spike Count Correlations between Visual Cortical Areas." J Neurosci **36**(28): 7523-7534.

Shadlen, M. and W. Newsome (1994). "Noise, neural codes and cortical organization." **Current Opinion in Neurobiology** **4**: 569-579.

Singer, W. and C. M. Gray (1995). "Visual feature integration and the temporal correlation hypothesis." Annu Rev Neurosci **18**: 555-586.

Vanrullen, R., N. A. Busch, J. Drewes and J. Dubois (2011). "Ongoing EEG Phase as a Trial-by-Trial Predictor of Perceptual and Attentional Variability." Front Psychol **2**: 60.

Von Der Malsburg, C. (1994). The correlation theory of brain function. Models of neural networks, Springer**:** 95-119.

Xiao, J. and X. Huang (2015). "Distributed and Dynamic Neural Encoding of Multiple Motion Directions of Transparently Moving Stimuli in Cortical Area MT." J Neurosci **35**(49): 16180-16198.

Xiao, J., Y. Q. Niu, S. Wiesner and X. Huang (2014). "Normalization of neuronal responses in cortical area MT across signal strengths and motion directions." J Neurophysiol **112**(6): 1291-1306.

bioRxiv preprint doi: https://doi.org/10.1101/777912; this version posted September 23, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Zohary, E., M. N. Shadlen and W. T. Newsome (1994). "Correlated neuronal discharge rate and its implications for psychophysical performance." Nature **370**(6485): 140-143.