

Sensory and Motor Systems

The Effect of Inclusion Criteria on the Functional Properties Reported in Mouse Visual Cortex

Natalia Mesa,^{1,2}  Jack Waters,^{1,2} and Saskia E. J. de Vries^{1,2}<https://doi.org/10.1523/ENEURO.0188-20.2021>¹Allen Institute, Seattle, WA 98109 and ²University of Washington, Seattle, WA 98195

Abstract

Neurophysiology studies require the use of inclusion criteria to identify neurons responsive to the experimental stimuli. Five recent studies used calcium imaging to measure the preferred tuning properties of layer 2/3 pyramidal neurons in mouse visual areas. These five studies employed different inclusion criteria and reported different, sometimes conflicting results. Here, we examine how different inclusion criteria can impact reported tuning properties, modifying inclusion criteria to select different subpopulations from the same dataset of almost 17,000 layer 2/3 neurons from the Allen Brain Observatory. The choice of inclusion criteria greatly affected the mean tuning properties of the resulting subpopulations; indeed, the differences in mean tuning because of inclusion criteria were often of comparable magnitude to the differences between studies. In particular, the mean preferred temporal frequencies (TFs) of visual areas changed markedly with inclusion criteria, such that the rank ordering of visual areas based on their TF preferences changed with the percentage of neurons included. It has been suggested that differences in TF tuning support a hierarchy of mouse visual areas. These results demonstrate that our understanding of the functional organization of the mouse visual cortex obtained from previous experiments critically depends on the inclusion criteria used.

Key words: calcium imaging; data analysis; inclusion criteria; neurophysiology; visual cortex

Significance Statement

Inclusion criteria are widely used in physiological studies to limit analysis to active or responsive neurons, yet the impact of the criteria employed on the ensuing analyses are rarely considered. We have compared the effect of several inclusion criteria used in published studies comparing visual responses across cortical visual areas in the mouse cortex by applying these to a single dataset. The choice of inclusion criteria greatly affected the mean tuning properties of the resulting subpopulations; indeed, the differences in mean tuning because of inclusion criteria were often of comparable magnitude to the differences between studies.

Introduction

Five recent studies have employed two-photon calcium imaging to compare spatial frequency (SF) tuning, temporal frequency (TF) tuning, orientation selectivity, and directional selectivity of neurons across mouse visual cortical areas (Table 1; Fig. 1; Andermann et al., 2011; Marshel et al., 2011; Roth et al., 2012; Tohmi et al., 2014; Sun et al., 2016). Some results were consistent across studies, e.g., the mean preferred TF of neurons in area AL was greater than those in V1 (Fig. 1A), but there were also differences

between studies, e.g., some studies found that the mean preferred TF of neurons in PM was greater than those in V1 while others found the opposite. Further, the magnitudes of average TF tuning, orientation selectivity index (OSI), and direction selectivity index (DSI) in individual visual areas as well as the rank order of these properties between visual areas differed across studies (Fig. 1). All five studies imaged layer 2/3 of mouse visual cortex and activity was evoked with a drifting grating stimulus, but the studies differed in anesthesia state, calcium indicator, stimulus parameters, and in the inclusion criteria used in

Received May 8, 2020; accepted January 13, 2021; First published January 28, 2021.

The authors declare no competing financial interests.

Author contributions: N.M., J.W., and S.E.J.d.V. designed research; N.M. and S.E.J.d.V. analyzed data; N.M., J.W., and S.E.J.d.V. wrote the paper.

Table 1: Summary of the experimental conditions and inclusion criteria used in published studies

Paper	Anesthesia	Indicator	Stimulus	Stimulus repetitions	Responsiveness criteria	Percentage of responsive cells	# Cells Allen Brain Observatory
Study 1 Sun et al. (2016)	None	GCaMP6s	12-s full-field square grating TF: 0.5, 1 Hz SF 0.05cpd 8–16 directions	10	Mean $\Delta F/F > 10\%$	49% ($n = 1279/2609$)	17.5% ($n = 2962/16923$)
Study 2 Roth et al. (2012)	Urethane	OGB-1	5-s full-field sine wave grating TF: 0.5, 1, 2, 4 Hz SF: 0.01, 0.02, 0.04, 0.08, 0.16cpd 8 directions	4	In 50% of trials, mean $\Delta F/F > \text{baseline} + 3\sigma$ Mean response $> 5\%$	44% ($n = 399/973$)	6.3% ($n = 1068/16923$)
Study 3 Andermann et al. (2011)	None	GCaMP3	40° sine wave grating patches TF: 0.5, 1, 2, 4, 8, 15, 24 Hz SF: 0.02, 0.04, 0.08, 0.16, 0.32cpd Direction: upward	9–28	t test comparing grating response with blank sweep with Bonferroni correction ($p < 0.05/n$)	8% ($n = 28/340$)	29.6% ($n = 5015/16923$)
Study 4 Marshel et al. (2011)	Isoflurane	OGB-1	4-s full-field sine wave gratings SF: 0.04 cpd TF: 0.5, 1, 2, 4, 8 Hz 8 directions	5	Mean $\Delta F/F > 6\%$ Reliability > 1 (see Materials and Methods)	42% ($n = 586/1395$)	7.8% ($n = 1068/16923$)
Study 5 Tohmi et al. (2014)	Urethane	Fura-2	5-s ramping square and sine wave gratings SF 0.05 and 0.1cpd 8 directions	28–124	Max $\Delta F/F > 5\%$	41.2% ($n = 142/347$)	94.7% ($n = 16,033/16,923$)

Last column shows the number of neurons selected from the Allen Brain Observatory.

analysis (Table 1). It is likely that all these differences contribute to the contrasting results. Here, we leverage a single large and open dataset, the Allen Brain Observatory, to quantify the impact of the choice of inclusion criteria on the measurement of tuning properties of neurons in mouse visual areas.

Calcium imaging studies usually require the use of inclusion criteria to select neurons that are deemed to be “active” or “responsive” such that the derived analysis of their activity is relevant to the aims of the experiment and not a quantification of noise. As the measured fluorescence shows continuous fluctuations, these criteria serve to identify which fluctuations reflect signal rather than noise. Criteria are often based on the amplitude of the fluorescence change, e.g., a threshold on the mean or median change in fluorescence over multiple trials, or its reproducibility, e.g., a statistically significant stimulus-evoked change in fluorescence on a subset of trials. Naturally, some neurons exhibit large-amplitude changes in fluorescence on every trial in response to a preferred stimulus and fulfil both amplitude and reproducibility criteria (Fig. 2A–C). Many neurons display reproducible, small-amplitude changes (Fig. 2D–F) or large-amplitude

changes in fluorescence on only some trials (Fig. 2G–I). Although not often used as the basis for inclusion criteria, other features of the fluorescence traces, such as periodicity in the fluorescence in response to a periodic stimulus such as a drifting grating (Fig. 2I) and tuning to stimulus characteristics such as orientation and TF (Fig. 2C,H,I), may also be suggestive of stimulus-evoked activity (Niell and Stryker, 2008).

Each of the five studies used different inclusion criteria and it is unclear whether these different criteria select the same or different neurons and how they impact the distribution of measured responses to visual stimuli across the population. Here, we explore the effects of inclusion criteria on results from a single large dataset, eliminating the effects of different experimental conditions. We used recordings from the Allen Brain Observatory, a database of physiological activity in visual cortex measured with two-photon calcium imaging from adult GCaMP6f transgenic mice (de Vries et al., 2020). We found that tuning properties varied with inclusion criteria, in some cases changing the rank order of tuning properties across mouse cortical visual areas.

Materials and Methods

Stimulus and dataset

We used calcium imaging recordings from the Allen Brain Observatory, a publicly available dataset that surveys physiological activity in the mouse visual cortex (de Vries et al., 2020). We specifically used the responses to the drifting grating stimulus in this dataset. This stimulus consisted of a 2 s grating followed by a 1 s mean luminance gray period. Six TFs (1, 2, 4, 8, 15 Hz), eight different directions, and one SF (0.04 cpd) were used. Each grating condition was presented 15 times.

This work was supported by the Allen Institute.

Acknowledgements: We thank Dan Millman and Jun Zhuang for their feedback on this manuscript. We also thank Allen Institute founder, Paul G. Allen, for his vision, encouragement, and support.

Correspondence should be addressed to Saskia E. J. de Vries at saskiad@alleninstitute.org.

<https://doi.org/10.1523/ENEURO.0188-20.2021>

Copyright © 2021 Mesa et al.

This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution and reproduction in any medium provided that the original work is properly attributed.

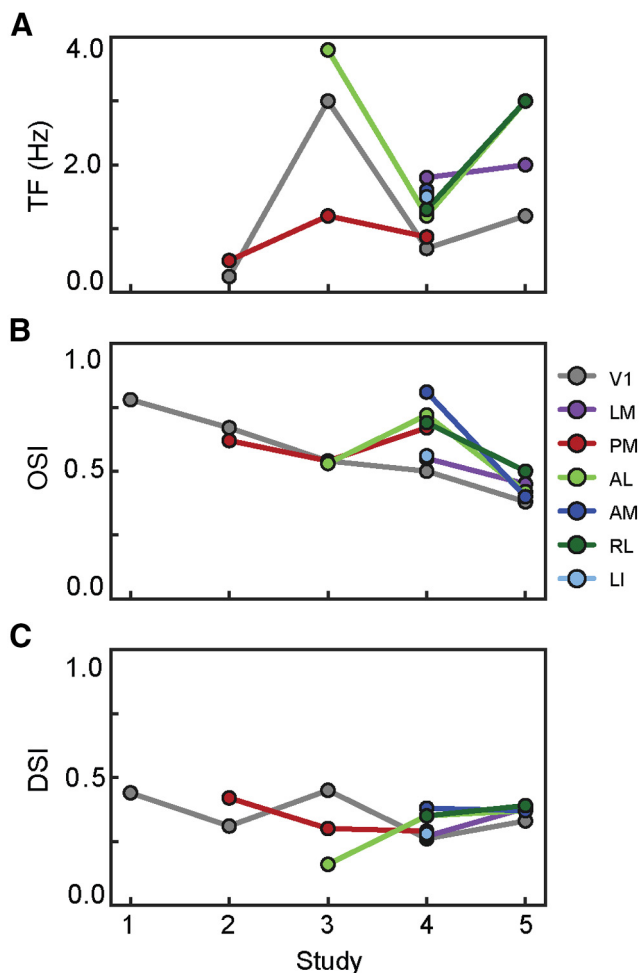


Figure 1. Tuning characteristics in published studies. **A**, Mean preferred TF tuning of seven visual areas reported in five published studies. **B**, **C**, Same as in **A** but reporting the OSI and the DSI.

Data analysis was performed in Python using the AllenSDK. The evoked response was defined as the mean dF/F during the 2-s grating presentation. Responses to all 15 stimulus presentations were averaged together to calculate the mean evoked response.

We restricted our analysis to cells in layer 2/3 (175–250 μm below pia) of transgenics lines Cux2-CreERT2; Camk2a-tTa;Ai93 and Slc17a7-IRES2-Cre;Camk2a-tTa; Ai93, which express GCaMP6f in neural populations in layer 2/3 and throughout neocortex, respectively. A total of $N = 16,923$ neurons from 66 mice (42 male, 24 female) were used for this analysis.

Metrics

The preferred direction and TF condition was defined as the grating condition that evoked the largest mean response. In order to compute the average TF tuning of a population of neurons, these TF values were first converted an octave scale (base 2), averaged, then converted back to a linear scale and reported.

Direction selectivity was computed for each neuron as the following:

$$DSI = \frac{R_{pref} - R_{null}}{R_{pref} + R_{null}},$$

where R_{pref} is the mean response at to the preferred direction and R_{null} is the mean response to the opposite direction.

Orientation selectivity was computed for each neuron using the global OSI (OSI; Ringach et al., 1997), defined as the following:

$$OSI = \frac{\sum R_{\theta} e^{2i\theta}}{\sum R_{\theta}},$$

where R_{θ} is the mean response at each orientation θ .

The coefficient of variance (CV) was used as our metric to determine robustness. CV was calculated for each neuron as the ratio of SD of the 15 responses to the preferred condition (mean dF/F over the 2-s stimulus presentation) to the mean evoked response (see above). A low CV would indicate high robustness.

Metrics were either computed using all available trials, or with cross validation. When using cross validation, half of the trials (chosen at random, without replacement) were used to identify the preferred direction and TF, and the other half of the trials were used to compute the metrics using those preferred conditions. This was iterated 50 times, and the resulting metrics were averaged together.

When examining the effects of the number of trials, for each number of trials (n), n trials were chosen at random (without replacement), and the cross-validation was done as described above.

Inclusion criteria

Published studies used the following inclusion criteria, which we applied to cells in the Allen Brain Observatory dataset in the following manner:

Study 1: The mean evoked response (dF/F) to the preferred stimulus condition is $>10\%$ (Sun et al., 2016).

Study 2: In 50% of trials, the response is (1) larger than the $3\times$ the SD of the prestimulus baseline and (2) larger than 5% dF/F (Roth et al., 2012).

Study 3: Paired t test ($p < 0.05$) with Bonferroni correction comparing the mean evoked response during the blank sweeps with mean evoked responses to preferred stimulus condition (Andermann et al., 2011).

Study 4: (1) The mean response (dF/F) to any stimulus condition is $>6\%$. And (2) reliability >1 where:

$$reliability = \frac{R_{pref} - R_{blank}}{\sigma_{pref} + \sigma_{blank}} \text{ (Marshall et al., 2011).}$$

Study 5: The maximum fluorescence change (dF/F) during the 2-s stimulus presentation block to any stimulus condition was $>4\%$ (Tohmi et al., 2014).

Code availability

The code used in this paper is available at https://github.com/nataliamv2/inclusion_criteria.

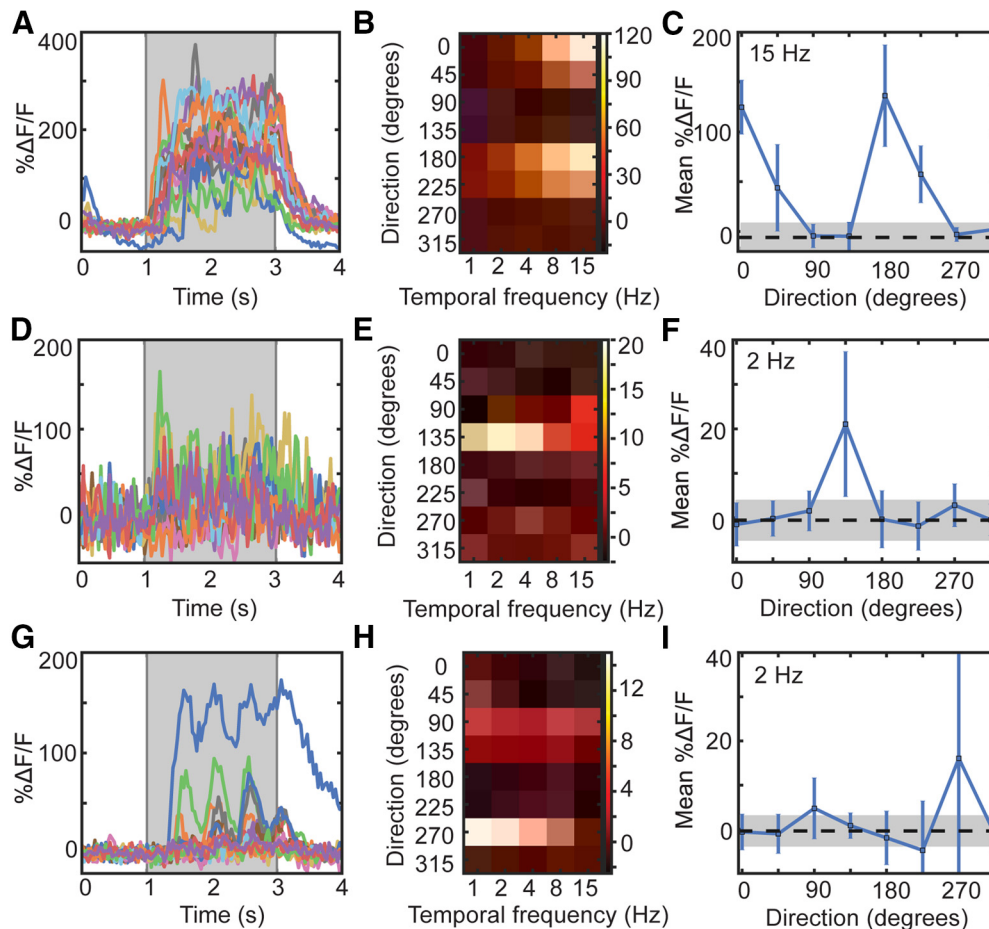


Figure 2. Example cells that pass inclusion criteria exclusively. **A**, All DF/F responses to the preferred stimulus condition (TF and direction) of a cell that passes all published inclusion criteria. **B**, Heatmap of mean %DF/F responses to each stimulus condition (TF \times direction). **C**, Mean %DF/F responses (\pm SEM) to stimuli of different grating directions in the same example cell. **D–F**, Same as in **A** but with a cell that passes most criteria, but not Study 2. **G–I**, Same as in **A** but with a cell that only passes Study 1 criteria.

Results

The five studies employed a range of inclusion criteria, selecting 8–49% of the neurons in their respective studies (Table 1). The inclusion criteria were based on one or both of the amplitude and the trial-to-trial variability of the evoked responses and we therefore calculated the mean and SD of the response of each neuron to its peak stimulus condition (the direction and TF that evoked the largest mean response). We applied the five different inclusion criteria to the Allen Brain Observatory, a large two-photon calcium imaging data set. We restricted our analysis to layer 2/3 excitatory neurons imaged 175–250 μ m below the pia in *Cux2-CreERT2;Camk2a;Ai93* and *Slc17a7-IRES2-Cre;Camk2a;Ai93* mice, yielding a dataset of fluorescence recordings from 16,923 neurons. Different inclusion criteria selected different, often overlapping populations of neurons (6–94% of 16,923 neurons; Table 1, column 7), readily visualized by plotting the mean against the SD of the response (Fig. 3A). The results derived using these different criteria covered similar ranges to those in the published studies, consistent with the idea that effects of inclusion criteria could contribute to the disparate results across published studies (Fig. 3B).

Using CV ($CV = SD/mean$) as a measure of response robustness, we asked how increasing the number of neurons selected, from the most robust (lowest CV) to the least (highest CV), affects the computed tuning metrics. For some metrics, including more neurons affected tuning properties by almost as much as the differences between studies. For example, increasing included neurons changed the mean preferred TF for V1, PM, and AL as well as the rank order of these three areas, such that AL and PM display different mean TFs when only the top decile are included, but have the same mean TF when all neurons are included (Fig. 4A–D,M). Within V1, the change in mean TF reflects the fact that the highest decile (10% with highest CV) shows a broader distribution of preferred TF than the lowest decile (Fig. 4B,C). In contrast, the effect on OSI was smaller and more consistent across areas, having a smaller effect on the value or the rank order across areas (Fig. 4E–H,M). Finally, increasing the number of neurons included increased the mean DSI, and did so consistently and significantly across all visual areas (Fig. 4I–L,M). The increase in DSI reflects the fact that many of the neurons in the lowest decile have a DSI of 1, whereas the neurons in the highest decile have a uniform distribution of DSIs (Fig. 4J,K).

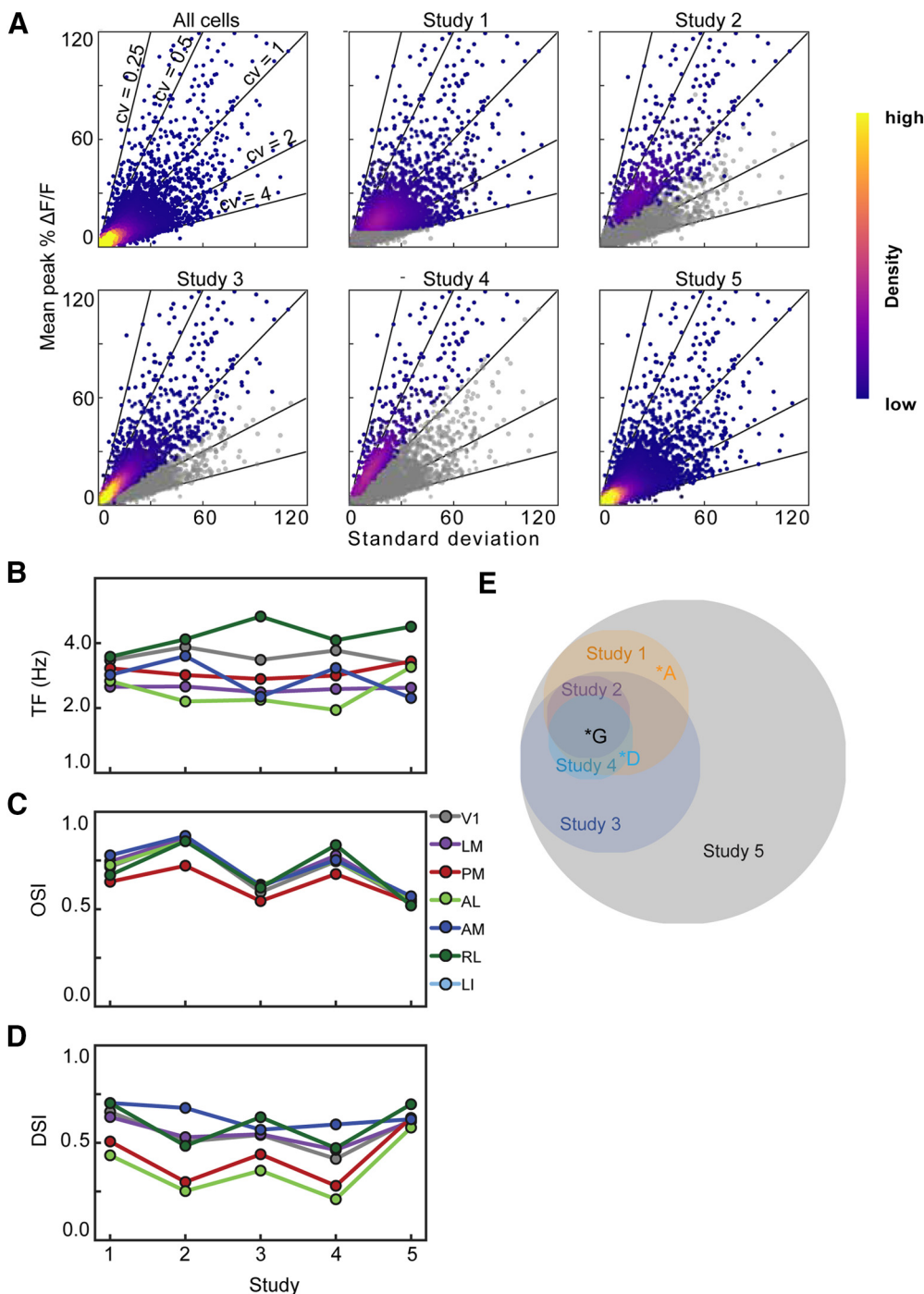


Figure 3. Most studies select for neurons along similar axes of the data. **A**, Six density plots of the mean response at the preferred stimulus condition (%DF/F) against the SD of the responses at the preferred stimulus condition where each point represents a single neuron. For each study, colored neurons are those selected for by inclusion criteria. Heatmap represents the density of neurons. **B–D**, Tuning characteristics after inclusion criteria are applied to Allen Brain Observatory. **B**, Mean TF tuning of six visual areas when different inclusion criteria are applied. **C**, **D**, show the mean OSI and DSI of six visual areas, respectively. **E**, Venn Diagram of neurons that were selected for by each inclusion criteria. Area of circles represents the number of neurons. Letters indicate example neurons from Figure 2.

None of the inclusion criteria used in the published studies apply a threshold on the CV specifically, but some incorporate measurements of reliability that might have a similar effect. If criteria are selecting neurons based primarily on reliability, one might expect that selecting a

population of neurons with matched mean CV would result in similar tuning properties and would replicate the differences observed between the studies. We selected populations of neurons that had the same mean CV as those chosen by each inclusion criteria, for each area

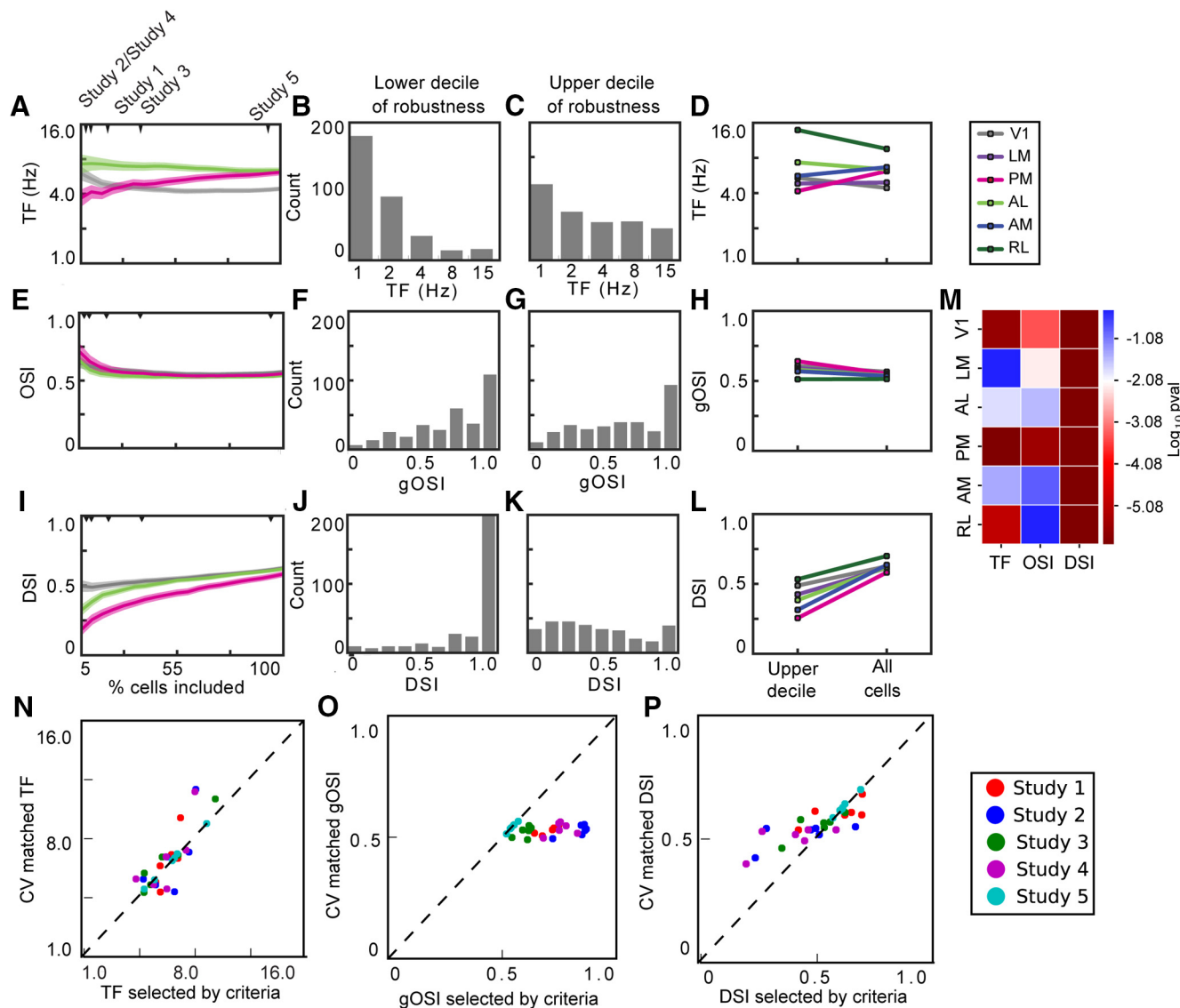


Figure 4. Tuning characteristics of neurons based on robustness. **A, E, I,** Mean TF, OSI, and DSI tuning of neurons in V1, AL, and PM based on what percentage of most robust cells (cells with low CV) are included in the analysis. Shaded regions indicate SEM. The minimum percentage most robust cells displayed is 5%. **B, C,** Distribution of TF tuning of 10% least robust cells. **F, G, J, K,** Same as in **B, C** but with OSI and DSI. **D, H, L,** Mean TF, OSI, and DSI tuning of neurons in all visual areas comparing the 10% most robust neurons to the entire population of neurons. **M,** Heat map displaying p values for Mann–Whitney U test comparing the 10% most robust neurons and the entire population of neurons. The color scale is centered at $p = 0.05/6$ to account for Bonferroni correction. **N,** Mean DSI calculated for neurons selected to match the mean CV for each the neurons selected by each criterion, for each area, compared with the mean DSI for the neurons selected by that criteria and area. **O, P,** Same as in **M** but for OSI and TF.

separately, and compared the tuning properties for that population to the tuning properties for the neurons chosen by the criteria. For some metrics, there was a high correlation between these values, namely mean preferred TF and mean DSI ($r = 0.82$ Pearson’s correlation for both; Fig. 4N,P). For preferred TF the values were close to unity, indicating that selecting neurons by their CV closely matched the differences between studies. For DSI, however, the range of DSI values was more limited. Thus, while there was a high correlation between the values for neurons selected by CV to those for neurons selected by the criteria, the shallow slope of this relationship made it

less predictive. Further, for the mean OSI, there was no correlation between these values ($r = 0.09$; Fig. 4O). Thus, some of the differences between the published studies could result from the inclusion criteria effectively selecting neurons based on their reliability at different threshold. However, it is clear that the criteria did not select neurons exclusively based on the reliability, as captured by the CV, as CV alone cannot account for all of the differences between the studies.

Selection by CV displayed a greater effect on preferred TF and DSI than on OSI, likely because the measurements of preferred TF and DSI are more susceptible to noise.

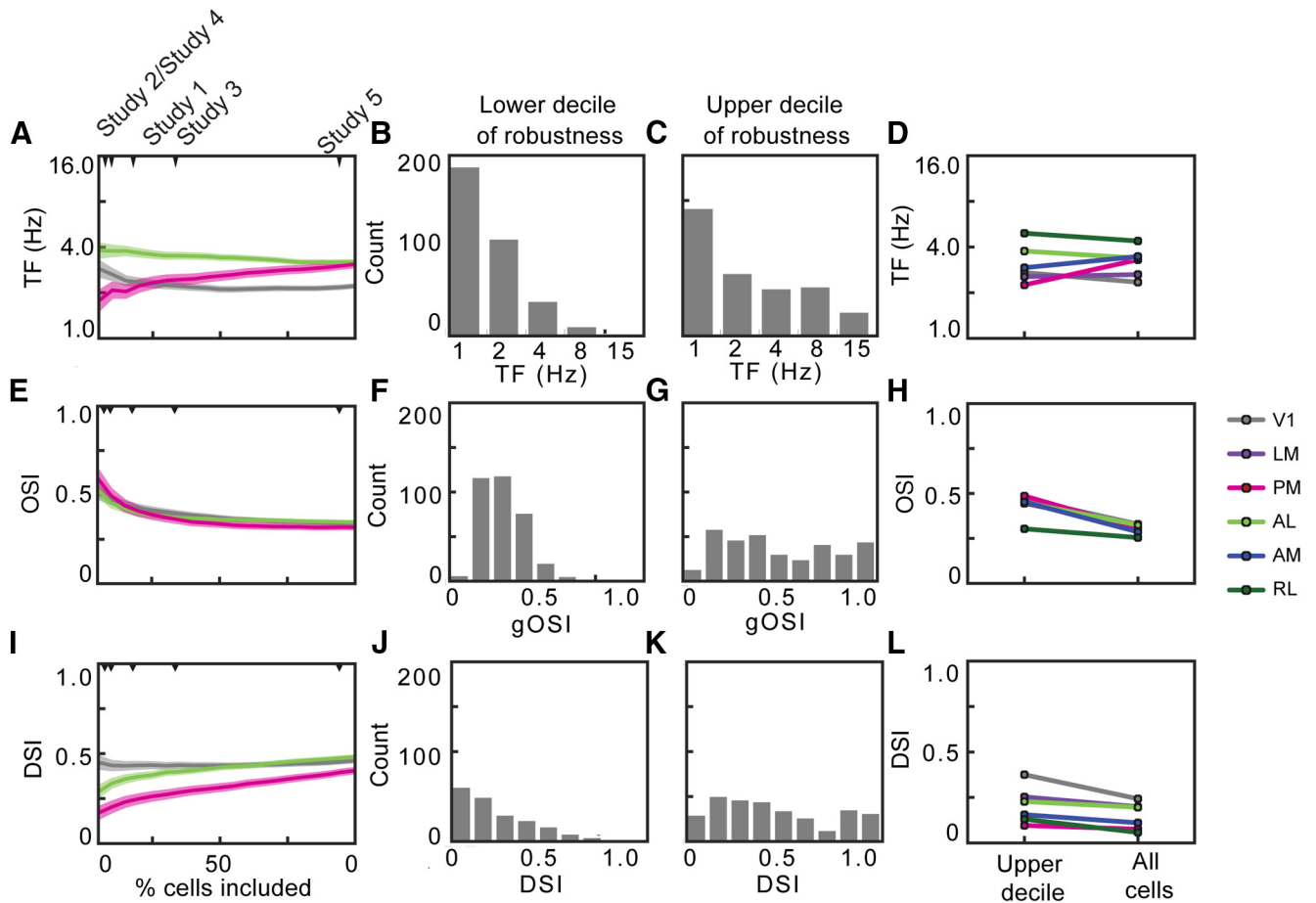


Figure 5. Tuning characteristics of neurons based on robustness with cross-validated metrics. **A, E, I,** Mean TF, OSI, and DSI tuning of neurons in V1, AL, and PM based on what percentage of neurons are included in the analysis, starting with the most robust neurons. Shaded regions indicate SEM. The minimum percentage most robust cells displayed is 5%. **B,** Distribution of TF tuning of 10% least robust neurons. **C,** Distribution of TF tuning of 10% most robust neurons. **F, G, J, K,** Same as in **B, C** but with OSI and DSI. **D, H, L,** Mean TF, OSI, and DSI tuning of neurons in all visual areas in 10% most robust neurons versus the entire population of neurons.

The neurons with the noisiest responses (greatest CV) commonly displayed DSI ~ 1 (Fig. 4J), which is inevitable when the response to the null direction is 0. The response to the preferred direction need not be large and could even result from a single trial having just a small amplitude fluorescence change. As the preferred TF is the TF at which the neuron has its largest response, regardless of amplitude or reliability, the TF tuning is similarly sensitive to small numbers of noisy events. In contrast, OSI is calculated from the responses to all eight directions of drifting gratings and is thus less sensitive to a small amplitude response in one condition.

Might a calculation that is more robust to trial-to-trial variability reduce the sensitivity of measurements to inclusion criteria or CV? We recalculated OSI, DSI and TF with cross-validation, using half of the trials to identify the stimulus condition that evoked the largest mean responses (grating direction and TF) and then calculated OSI, DSI and TF for these preferred conditions from the other half of the trials. The overall effect of including more neurons based on their CV on the cross-validated metrics across different areas was similar to that on the non-cross-validated metrics (Fig. 5). The notable difference is

that the noisy neurons in the lowest decile of robustness no longer have high DSI or OSI values, but are shifted to much lower values (Fig. 5F,J). This difference is also reflected in the fact that the overall curves are shifted to lower values (compare Figs. 5E,I and Fig. 4E,I). Thus, while more statistically robust metrics calculated through cross-validation likely better reflect the true values of the population, they do not reduce the impact of selection on those metrics.

Different studies presented each visual stimulus multiple times, with numbers of repetitions ranging from 4 to 24 trials (Table 1). Might the number of repetitions account for some of the differences between studies? We computed OSI, DSI and preferred TF using subsets of 4–14 trials. As expected, the variability of the responses decreased as the number of trials increased, resulting in a lower mean CV across the entire population (Fig. 6A). Visualizing the neurons by plotting response mean versus SD for $n=4$ trials (Fig. 6B) and $n=14$ trials (Fig. 6C), it is clear that the bulk of the data are shifted to more robust responses. Increasing the number of trials had a small effect on the cross-validated metrics (Fig. 6D–F), decreasing both the mean OSI and DSI across all areas (when

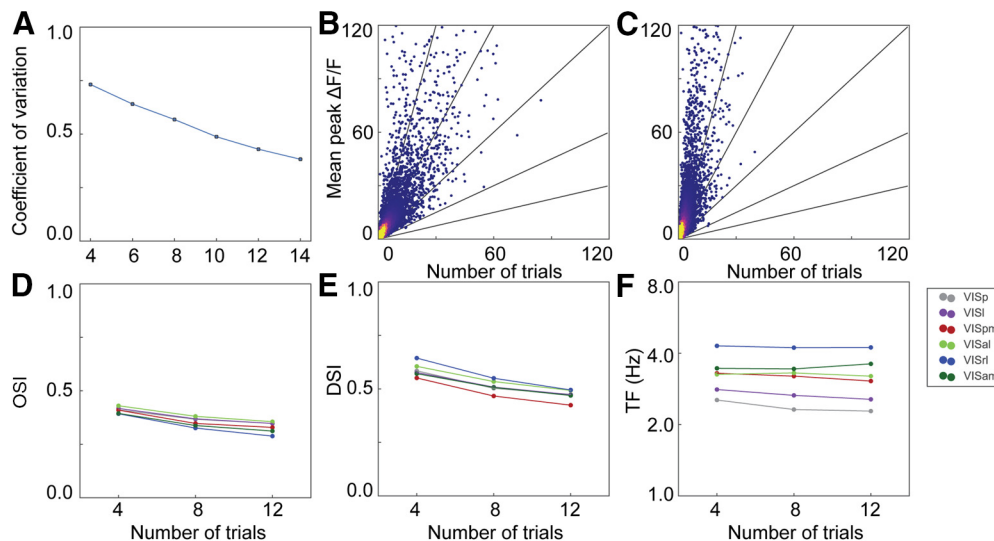


Figure 6. How trial number changes tuning metrics and CV. **A**, Mean CV calculated at the preferred condition using different numbers of trials and the cross-validation method. **B**, Mean peak response at the preferred condition versus SD at the preferred condition using only four trials and the cross validation method. **C**, Same as in **B** but using 14 trials. **D–F**, OSI, TF, and DSI calculated using the cross-validation method as a function of the number of trials used in the analysis.

including all neurons). The effect was consistent across all areas, however, thus the number of trials did not impact the rank order across areas. Thus, while more trials can reduce the variability of the response measurements, it is unlikely that these differences had a large effect on the differences observed between studies.

Discussion

We applied different inclusion criteria to the Allen Brain Observatory two-photon dataset to examine how these criteria impact the reported tuning properties across visual areas after experimental differences are eliminated. That different inclusion criteria selected different subsets of neurons might not be surprising, but the extent of the differences between selected neurons was substantial. One key difference was in the numbers of neurons selected. To examine how including more, or fewer, neurons could impact the tuning properties, we used CV as a metric of robustness and shifted our threshold for inclusion. Mean TF, OSI, and DSI changed differently with the robustness of the responses of the underlying neurons. The preferred TF was the most sensitive, OSI the least sensitive.

Our results offer one possible explanation why published studies comparing TF, OSI, and DSI across mouse visual areas have produced different results for TF and more similar results for OSI and DSI. Mean TF tuning is more sensitive than OSI and DSI to the neurons selected. As a result, comparison across studies is difficult and there remains considerable uncertainty regarding the mean TF and the rank order of TF tuning across mouse visual areas.

We used CV to examine how including more neurons can impact the reported results, as one of the big differences of the criteria is the number of neurons they select from our dataset. But this is not the only difference

between these criteria. The Venn diagram (Fig. 3E) reveals that the cells selected by different criteria are not described by a set of concentric circles, and neurons with mean CV matched to those selected by an inclusion criterion have different tuning (Fig. 4O,P), revealing that the inclusion criteria use features of the neural responses in addition to the size and reliability of neurons' responses to their preferred condition. For instance, the statistical tests employed in Studies 3 and 4 also depend on the size and reliability of the neurons' responses to the blank sweep.

Cross-validating metrics and increasing the number of trials can each improve the accuracy of the measured responses. Cross-validation can mitigate the impact of particularly noisy responses, reducing the impact of small numbers of outlier trials. This is most evident in the effect of cross-validation on the DSI distribution for the neurons in the lowest decile of robustness (Fig. 5J). It is possible that inclusion criteria based on the reliability of metrics across iterations of cross-validation might be more effective for identifying neurons with truly robust responses.

Our results illustrate how inclusion criteria can play a role in determining the tuning properties of visual areas. The choice of inclusion criteria is unlikely to account for all of the differences observed between the original studies, indicating that other experimental factors are important. Other factors likely include anesthesia state, the type of anesthesia used, the calcium indicator, image brightness, as well as visual stimulus parameters. Brain state can modulate neural responses in visual cortex, and anesthesia in particular can impact both the spontaneous and evoked responses. The type of anesthesia can also be a factor, with urethane impacting spontaneous and evoked firing rates but not OSI (Niell and Stryker, 2010) and atropine affecting OSI but not spontaneous firing rate, evoked firing rate, DSI, preferred TF, or preferred SF (Durand et

al., 2016). Stimulus parameters, such as the size or contrast of the drifting gratings or the precise SFs and TFs, do also impact the evoked responses and could account for some of the differences observed between the original studies.

Calcium indicators have different sensitivities and signal-to-noise properties (Hendel et al., 2008; Chen et al., 2013), such that thresholds in mean DF/F appropriate for one indicator might not be appropriate for another. Most of the inclusion criteria selected ~40–50% of neurons when applied to their own data, but when applied to the Allen Brain Observatory data the percentage of neuron included often differed substantially, presumably because experimental conditions such as indicator brightness differed across studies. For example, simple thresholds on peak DF/F cannot be applied uniformly across different calcium indicators. Thus, it is unlikely that a single set of inclusion criteria would be appropriate across a wide range of experimental conditions, and that these criteria must be chosen and validated by experimenters, including, for instance, an analysis of how metrics change based on how restrictive criteria are (Kim et al., 2018).

Functional specialization of the higher visual areas in mouse cortex has been interpreted as evidence of parallel streams (Andermann et al., 2011; Marshel et al., 2011). For example, V1 is thought to transfer low TF, high SF information to PM, the putative gateway to the dorsomedial stream (López-Aranda et al., 2009; Polack and Contreras, 2012; Glickfeld et al., 2013). However, in some studies, neurons in V1 and PM have similar mean TF tuning (with PM's being 1.3–2× that of V1; Marshel et al., 2011; Roth et al., 2012), while others show that mean TF tuning in PM neurons that is 1/3 that of V1 neurons (Andermann et al., 2011). Our results indicate that in the most robust neurons, V1 has a higher TF tuning than PM, but in the least robust neurons, PM has a higher TF tuning than V1, potentially explaining some of the difference between studies. Since TF is sensitive enough to inclusion criteria to change the relative order of TF tuning, it is difficult to interpret the relative TF tuning between visual areas currently. The most appropriate inclusion criteria would take into account how downstream targets filter or weight inputs and how robustness factors into that weighting. Since we do not know what this weighting is, we must be cautious in drawing conclusions about functional organization from these analyses.

References

- Andermann ML, Kerlin AM, Roumis DK, Glickfeld LL, Reid RC (2011) Functional specialization of mouse higher visual cortical areas. *Neuron* 72:1025–1039.
- Chen TW, Wardill TJ, Sun Y, Pulver SR, Renninger SL, Baohan A, Schreiter ER, Kerr RA, Orger MB, Jayaraman V, Looger LL, Svoboda K, Kim DS (2013) Ultrasensitive fluorescent proteins for imaging neuronal activity. *Nature* 499:295–300.
- de Vries SEJ, Lecoq JA, Buice MA, Groblewski PA, Ocker GK, Oliver M, Feng D, Cain N, Ledochowitsch P, Millman D, Roll K, Garrett M, Keenan T, Kuan L, Mihalas S, Olsen S, Thompson C, Wakeman W, Waters J, Williams D, et al. (2020) A large-scale standardized physiological survey reveals functional organization of the mouse visual cortex. *Nat Neurosci* 23:138–151.
- Durand S, Iyer R, Mizuseki K, de Vries S, Mihalas S, Clay Reid R (2016) A comparison of visual response properties in the lateral geniculate nucleus and primary visual cortex of awake and anesthetized mice. *J Neurosci* 36:12144–12156.
- Glickfeld LL, Andermann ML, Bonin V, Reid RC (2013) Cortico-cortical projections in mouse visual cortex are functionally target specific. *Nat Neurosci* 16:219–226.
- Hendel T, Mank M, Schnell B, Griesbeck O, Borst A, Reiff DF (2008) Fluorescence changes of genetic calcium indicators and OGB-1 correlated with neural activity and calcium in vivo and in vitro. *J Neurosci* 28:7399–7411.
- Kim MH, Znamenskiy P, Iacaruso MF, Mrsic-Flogel TD (2018) Segregated subnetworks of intracortical projection neurons in primary visual cortex. *Neuron* 100:1313–1321.e6.
- López-Aranda MF, López-Téllez JF, Navarro-Lobato I, Masmudi-Martín M, Gutiérrez A, Khan ZU (2009) Role of layer 6 of V2 visual cortex in object-recognition memory. *Science* 325:87–90.
- Marshel JH, Garrett ME, Nauhaus I, Callaway EM (2011) Functional specialization of seven mouse visual cortical areas. *Neuron* 72:1040–1054.
- Niell CM, Stryker MP (2008) Highly selective receptive fields in mouse visual cortex. *J Neurosci* 28:7520–7536.
- Niell CM, Stryker MP (2010) Modulation of visual responses by behavioral state in mouse visual cortex. *Neuron* 65:472–479.
- Polack PO, Contreras D (2012) Long-range parallel processing and local recurrent activity in the visual cortex of the mouse. *J Neurosci* 32:11120–11131.
- Ringach DL, Sapiro G, Shapley R (1997) A subspace reverse-correlation technique for the study of visual neurons. *Vision Res* 37:2455–2464.
- Roth MM, Helmchen F, Kampa BM (2012) Distinct functional properties of primary and posteromedial visual area of mouse neocortex. *J Neurosci* 32:9716–9726.
- Sun W, Tan Z, Mensh BD, Ji N (2016) Thalamus provides layer 4 of primary visual cortex with orientation- and direction-tuned inputs. *Nat Neurosci* 19:308–315.
- Tohmi M, Meguro R, Tsukano H, Hishida R, Shibuki K (2014) The extrageniculate visual pathway generates distinct response properties in the higher visual areas of mice. *Curr Biol* 24:587–597.