

1           **A large-scale standardized physiological survey reveals**  
2           **functional organization of the mouse visual cortex**

3  
4           Saskia E. J. de Vries<sup>1,#,\*</sup>, Jerome Lecoq<sup>1,#,\*</sup>, Michael A. Buice<sup>1,#,\*</sup>, Peter A. Groblewski<sup>1</sup>,  
5           Gabriel K. Ocker<sup>1</sup>, Michael Oliver<sup>1</sup>, David Feng<sup>1</sup>, Nicholas Cain<sup>1</sup>, Peter Ledochowitsch<sup>1</sup>, Daniel  
6           Millman<sup>1</sup>, Kate Roll<sup>1</sup>, Marina Garrett<sup>1</sup>, Tom Keenan<sup>1</sup>, Leonard Kuan<sup>1</sup>, Stefan Mihalas<sup>1</sup>, Shawn  
7           Olsen<sup>1</sup>, Carol Thompson<sup>1</sup>, Wayne Wakeman<sup>1</sup>, Jack Waters<sup>1</sup>, Derric Williams<sup>1</sup>, Chris Barber<sup>1</sup>,  
8           Nathan Berbesque<sup>1</sup>, Brandon Blanchard<sup>1</sup>, Nicholas Bowles<sup>1</sup>, Shiella Caldejon<sup>1</sup>, Linzy Casal<sup>1</sup>,  
9           Andrew Cho<sup>1</sup>, Sissy Cross<sup>1</sup>, Chinh Dang<sup>1</sup>, Tim Dolbeare<sup>1</sup>, Melise Edwards<sup>1</sup>, John Galbraith<sup>1</sup>,  
10          Nathalie Gaudreault<sup>1</sup>, Terri L. Gilbert<sup>1</sup>, Fiona Griffin<sup>1</sup>, Perry Hargrave<sup>1</sup>, Robert Howard<sup>1</sup>,  
11          Lawrence Huang<sup>1</sup>, Sean Jewell<sup>2</sup>, Nika Keller<sup>1</sup>, Ulf Knoblich<sup>1</sup>, Josh Larkin<sup>1</sup>, Rachael Larsen<sup>1</sup>,  
12          Chris Lau<sup>1</sup>, Eric Lee<sup>1</sup>, Felix Lee<sup>1</sup>, Arielle Leon<sup>1</sup>, Lu Li<sup>1</sup>, Fuhui Long<sup>1</sup>, Jennifer Luviano<sup>1</sup>, Kyla  
13          Mace<sup>1</sup>, Thuyanh Nguyen<sup>1</sup>, Jed Perkins<sup>1</sup>, Miranda Robertson<sup>1</sup>, Sam Seid<sup>1</sup>, Eric Shea-Brown<sup>1,2</sup>,  
14          Jianghong Shi<sup>2</sup>, Nathan Sjoquist<sup>1</sup>, Cliff Slaughterbeck<sup>1</sup>, David Sullivan<sup>1</sup>, Ryan Valenza<sup>1</sup>, Casey  
15          White<sup>1</sup>, Ali Williford<sup>1</sup>, Daniela Witten<sup>2</sup>, Jun Zhuang<sup>1</sup>, Hongkui Zeng<sup>1</sup>, Colin Farrell<sup>1</sup>, Lydia Ng<sup>1</sup>,  
16          Amy Bernard<sup>1</sup>, John W. Phillips<sup>1</sup>, R. Clay Reid<sup>1</sup>, Christof Koch<sup>1</sup>

17  
18          <sup>1</sup>Allen Institute for Brain Science, Seattle, WA, USA  
19

20          <sup>2</sup>University of Washington, Seattle, WA, USA  
21

22          \*These authors contributed equally to this work  
23

24          <sup>#</sup>correspondence should be addressed to S.E.J.dV. ([saskiad@alleninstitute.org](mailto:saskiad@alleninstitute.org)), J.L.  
25          ([jeromel@alleninstitute.org](mailto:jeromel@alleninstitute.org)), and M.A.B. ([michaelbu@alleninstitute.org](mailto:michaelbu@alleninstitute.org)).  
26

26    **Summary**

27    To understand how the brain processes sensory information to guide behavior, we must know  
28    how stimulus representations are transformed throughout the visual cortex. Here we report an  
29    open, large-scale physiological survey of neural activity in the awake mouse visual cortex: the  
30    Allen Brain Observatory Visual Coding dataset. This publicly available dataset includes cortical  
31    activity from nearly 60,000 neurons collected from 6 visual areas, 4 layers, and 12 transgenic  
32    mouse lines from 243 adult mice, in response to a systematic set of visual stimuli. Using this  
33    dataset, we classify neurons based on joint reliabilities to multiple stimuli and validate this  
34    functional classification with models of visual responses. While most of the neurons fall into  
35    classes characterized by responses to specific subsets of the stimuli, a large minority of  
36    neurons are not reliably responsive to any of the stimuli, and these neurons increase in  
37    proportion in higher visual areas. These classes reveal a functional organization across cortical  
38    areas wherein putative dorsal areas show specialization for visual motion signals.

39

40

41      **Introduction**

42           Traditional understanding, based on several decades of research, is that visual cortical  
43        activity can be largely characterized by responses to a specific set of local features (modeled  
44        with linear filters followed by certain nonlinearities) and that these features become more  
45        selective and specialized in higher cortical areas<sup>1–4</sup>. However, it remains unclear to what extent  
46        this traditional understanding can account for the whole of V1<sup>5–7</sup>, let alone the rest of visual  
47        cortex. A key challenge results from the fact that this understanding is based on many small  
48        studies, recording responses from different stages in the circuit, using different stimuli and  
49        different analyses<sup>5</sup>. The inherent experimental selection biases and lack of standardization of  
50        this approach introduce additional obstacles to creating a cohesive understanding of cortical  
51        function. On the basis of these issues, influential reviews have questioned the validity of this  
52        standard model<sup>5–7</sup>, and have argued that “What would be most helpful is to accumulate a  
53        database of single unit or multi-unit data that would allow modelers to test their best theory  
54        under ecological conditions.”<sup>5</sup> To address these issues, we conducted a survey of visual  
55        responses across multiple layers and areas in the awake mouse visual cortex, using a diverse  
56        set of visual stimuli. This survey was executed in pipeline fashion, with standardized equipment  
57        and protocols and with strict quality control measures not dependent upon stimulus-driven  
58        activity (see **Methods, Supplemental Figures 1–8**).

59           Previous work in mouse has revealed functional differences among cortical areas in  
60        layer 2/3 in terms of the spatial and temporal frequency tuning of the neurons in each area<sup>8,9</sup>.  
61        However, it is not clear how these differences extend across layers and across diverse neuron  
62        populations. Here we expand such functional studies to include 12 Cre- defined neuron  
63        populations, including excitatory populations across 4 cortical layers (from layer 2/3 to layer 6),  
64        and two inhibitory populations (Vip and Sst). Further, it is known that stimulus statistics affect  
65        visual responses, such that responses to natural scenes cannot be well predicted by responses

66 to noise or grating stimuli<sup>10-15</sup>. To examine the extent of this discrepancy and its variation across  
67 areas and layers, we designed a stimulus set that included both artificial (gratings and noise)  
68 and natural (scenes and movies) stimuli. While artificial stimuli can be easily parameterized and  
69 interpreted, natural stimuli are closer to what is ethologically relevant to the mouse. Finally, as  
70 recording modalities have enabled recordings from larger and larger populations of neurons, it  
71 has become clear that populations might code visual and behavioral activity in a way that is not  
72 apparent by considering single neurons alone<sup>16</sup>. Here we imaged populations of neurons (mean  
73  $173 \pm 115$ , st. dev, for excitatory populations,  $19 \pm 11$  neurons for inhibitory populations) to  
74 explore both single neuron and population coding properties.

75 We find that 77% of neurons in the mouse visual cortex respond to at least one of these  
76 visual stimuli, many showing classical tuning properties, such as orientation and direction  
77 selective responses to gratings. These tuning properties exhibit differences across cortical  
78 areas and Cre lines. While subtle differences do exist between the excitatory Cre lines, these  
79 populations are largely similar, and the more marked differences are among the inhibitory  
80 interneurons. The responses to all stimuli are highly sparse and highly variable. We find that the  
81 variability of responses is not strongly correlated across stimuli, in general, but it does reveal  
82 evidence of functional response classes. We validate these functional response classes with a  
83 model of neural activity that contains most of the basic features found in visual  
84 neurophysiological modeling (e.g. “simple” and “complex” components) as well as the influence  
85 of the mouse’s running speed. For one class of neurons, these models perform quite well,  
86 predicting responses to both artificial and natural stimuli equally well. However, for many  
87 neurons, the models provide a poor description, particularly those in our largest single class of  
88 neurons, those that respond reliably to none of our visual stimuli. The representation of these  
89 response classes across areas reveals a separation of motion processing from spatial

90 computations. These results demonstrate the importance of a large, unbiased survey for  
91 understanding neural computation.

92

93 **Results**

94 Using adult C57BL/6 mice (mean age  $108 \pm 16$  days st. dev) that expressed a  
95 genetically encoded calcium sensor (GCaMP6f) under the control of specific Cre-line drivers (10  
96 excitatory lines, 2 inhibitory lines, **Supplemental Figure 7**), we imaged the activity of neurons in  
97 response to a battery of diverse visual stimuli. Data was collected from 6 different cortical visual  
98 areas (V1, LM, AL, PM, AM, and RL) and 4 different cortical layers. Visual responses of  
99 neurons at the retinotopic center of gaze were recorded in response to drifting gratings, flashed  
100 static gratings, locally sparse noise, natural scenes, and natural movies (**Figure 1f**), while the  
101 mouse was awake and free to run on a rotating disc. In total, 59,610 neurons were imaged from  
102 432 experiments (**Table 1**). Each experiment consisted of three one-hour imaging sessions, and  
103 33.6% of neurons were matched across all three sessions, and the rest were present in either  
104 one or two sessions (see Methods).

105 In order to systematically collect physiological data on this scale, we built data collection  
106 and processing pipelines (**Figure 1, Supplemental Figures 1-5**). The data collection workflow  
107 progressed from surgical headpost implantation and craniotomy to retinotopic mapping of  
108 cortical areas using intrinsic signal imaging, *in vivo* two-photon calcium imaging of neuronal  
109 activity, brain fixation, and histology using serial two-photon tomography (**Figure 1a,b,c**). To  
110 maximize data standardization across experiments, we developed multiple hardware and  
111 software tools to systematically collect data (**Figure 1d**). One of the key components was the  
112 development of a registered coordinate system that allowed an animal to move from one data  
113 collection step to the next, on different experimental platforms, and maintain the same

114 experimental and brain coordinate geometry (see Methods, **Supplemental Figure 1**). In  
115 addition to such hardware instrumentation, formalized standard operating procedures and  
116 quality control metrics were crucial for the collection of these data over several years (**Figure**  
117 **1e**).

118 Following data collection, movies of fluorescence associated with calcium influx were  
119 motion corrected, normalized, and regions of interest (ROIs) were segmented using automated  
120 algorithms (see **Methods, Supplemental Figure 9**). Signals from overlapping ROIs were  
121 demixed, and contamination from surrounding neuropil was subtracted (**Supplemental Figure**  
122 **11**). Segmented ROIs were matched across imaging sessions and ROIs were filtered to remove  
123 apical dendrites and other processes, with the aim of only including somatic ROIs. For each  
124 ROI, events were detected from  $\Delta F/F$  using an L0 regularized deconvolution algorithm (see  
125 **Methods, Supplemental Figure 12**), which deconvolves pointwise events assuming a linear  
126 calcium response for each event and penalizes the total number of events included in the  
127 trace<sup>17,18</sup>.

128 For each neuron, we computed the mean response to each stimulus condition using the  
129 detected events, and parameterized its tuning properties. Many neurons showed robust  
130 responses, exhibiting orientation-selective responses to gratings, localized spatial receptive  
131 fields, and reliable responses to natural scenes and movies (**Figure 2a-f, Supplemental Figure**  
132 **14**). For each neuron and each categorical stimulus (i.e. drifting gratings, static gratings, and  
133 natural scenes), the preferred stimulus condition was identified as the condition that evoked the  
134 largest mean response for that stimulus (e.g. the orientation and temporal frequency with the  
135 largest mean response for drifting gratings). For each trial of the stimulus, the activity of the  
136 neuron was compared to a distribution of activity for that neuron taken during the epoch of  
137 spontaneous activity, and a p-value was computed. If at least 25% of the trials of the neuron's  
138 preferred condition had a significant difference from the distribution of spontaneous activities

139 (p<0.05), the neuron was deemed to be responsive to that stimulus (see Methods for  
140 responsiveness criteria for locally sparse noise and natural movies). Neurons that met this  
141 criteria showed a change in activity that had some degree of reproducibility across trials.

142 In total, 77% of neurons were responsive to at least one of the visual stimuli presented  
143 (**Figure 2g**). The percent of responsive neurons depended on area and stimulus, such that V1  
144 and LM showed the highest number of visually responsive neurons. This dropped in other  
145 higher visual areas and was lowest in RL where only 33% of neurons responded to any of the  
146 visual stimuli. Natural movies elicited responses from the most neurons, while static gratings  
147 elicited responses from the fewest (**Figure 2h**). In addition to varying by area, the percent of  
148 responsive neurons also varied by Cre lines and layers, suggesting functional differences  
149 across these dimensions (**Supplemental Figures 15-19**).

150 For responsive neurons, visual responses were parameterized by computing several  
151 metrics, including preferred spatial frequency, preferred temporal frequency, direction  
152 selectivity, receptive field size, and lifetime sparseness (see Methods). We mapped these  
153 properties across cortical areas, layers, and Cre lines to examine the functional differences  
154 across these dimensions (**Figure 3, Supplemental Figure 20, 21**).

155 Comparisons across areas and layers revealed that direction selectivity is highest in  
156 layer 4 of V1 (**Figure 3b**). While previous literature has found higher direction selectivity in layer  
157 4 within V1<sup>19</sup>, we find here that this result is significant across all layer 4 specific Cre lines, and  
158 extends to the higher visual areas as well. Comparisons across the higher visual areas reveals  
159 that in superficial layers, the lateral higher visual areas (LM and AL) show significantly higher  
160 direction selectivity than the medial ones (PM and AM), but this difference is not significant in  
161 the deeper layers. This erosion of the differences between higher visual areas in deeper layers

162 is found for all metrics reported here, wherein the population differences are less pronounced,  
163 and often not significant, in layers 5 and 6 (**Figure 3c,d,e, Supplemental Figure 21**).

164 Across all areas, layers, and stimuli, visual responses in mouse cortex were highly  
165 sparse (**Figure 3f**). Considering the responses to natural scenes, we found that most neurons  
166 responded to a very small number of scenes (see the examples in Figure 2d). The sparseness  
167 of individual neurons was measured using lifetime sparseness, which captures the selectivity of  
168 a neuron's mean response to different stimulus conditions<sup>20,21</sup> (**see Methods**). A neuron that  
169 responds strongly to only a few scenes will have a lifetime sparseness close to 1, (**see**  
170 **Supplemental Figure 22**), whereas a neuron that responds broadly to many scenes will have a  
171 lower lifetime sparseness. Outside of layer 2/3, there was lower lifetime sparseness in areas RL,  
172 AM, and PM than in area V1, LM, and AL. But lifetime sparseness did not increase outside of  
173 V1; Responses did not become more selective in the higher visual areas. (**Figure 3f,**  
174 **Supplemental Figures 22**).

175 The pattern in single neuron direction selectivity was reflected in our ability to decode the  
176 visual stimulus from single-trial population vector responses, using all neurons, responsive and  
177 unresponsive (**Figure 4a, Supplemental Figure 23**). We used a K-nearest-neighbors classifier  
178 to predict the grating direction. Matching the tuning properties, areas V1, AL, and LM showed  
179 higher decoding performance than areas AM, PM, and RL, and these differences were more  
180 pronounced in superficial layers than in deeper layers. Similarly, the population sparseness  
181 (**Suppl. Fig. 22**), a measure of the selectivity of each scene (i.e. how many neurons respond on  
182 a given trial), largely mirrors the high average lifetime sparseness of the underlying populations  
183 (**Figure 4b**). Such high sparseness suggests that neurons are active at different times and thus  
184 their activities are weakly correlated. Indeed, the noise correlations of the populations reflect the  
185 results on population sparsity where excitatory populations show weak correlations while the  
186 inhibitory populations show somewhat higher correlations (**Figure 4c**). Depending on the

187 structure of the correlations in each population, they may be serving to either help or hinder  
188 information processing<sup>16,22,23</sup>. To test this, we measured the decoding performance when  
189 stimulus trials were shuffled to break trial-wise correlations. This had variable effects on  
190 decoding performance (**Figure 4d**). Overall, removing correlations tended to aid decoding in  
191 excitatory cells, and strongly in V1, while SST populations were hurt by removing noise  
192 correlations. Thus, despite smaller amounts of noise correlations, and thus more independent  
193 activity, the excitatory populations of cells displayed patterns of correlation that interfered with  
194 information about the drifting grating stimulus, while the opposite was true for Sst.

195 For all stimuli, the visually-evoked responses throughout mouse cortex showed a large  
196 amount of trial-to-trial variability. Even when removing the neurons deemed unresponsive, the  
197 percent of responsive trials for most responsive neurons at their preferred conditions was low -  
198 the median is less than 50% (**Figure 5a, Supplemental Figure 24**). This means that the  
199 majority of neurons in the mouse visual cortex are usually unresponsive, even when presented  
200 with the stimulus condition that elicits their largest average response. This is true throughout the  
201 visual cortex, though V1 showed slightly more reliable responses than higher visual areas. Sst  
202 interneurons showed very reliable responses, in particular to drifting gratings. This variability is  
203 reflected in the high coefficient of variation, with median values for excitatory neurons above 2,  
204 indicating that these neurons are super-Poisson (**Figure 5b**). We sought to capture this  
205 variability with a simple categorical model for drifting grating responses that attempts to predict  
206 the trial response (defined as the integrated event magnitude during each trial) from knowledge  
207 of the stimulus condition (the direction and temporal frequency of the grating, or whether the trial  
208 was a blank sweep). This regression quantifies how well the average tuning curve predicts the  
209 response for each trial. Comparing the trial responses to the mean tuning curve shows the  
210 degree of variability even when the model is fairly successful (**Figure 5c**). Consistent with this  
211 variability in visual responses, this model does a poor job of predicting responses to drifting

212 gratings for most neurons (**Figure 5d**). Few neurons are well predicted by their average tuning  
213 curve alone (21% of responsive neurons have  $r>0.5$ , this becomes 11% when considering all  
214 neurons, where  $r$  is the cross-validated Pearson correlation between model prediction and  
215 actual response). As expected, the ability to predict the responses is correlated with the  
216 measured variability ( $r=0.8$ , Pearson correlation).

217 One possible source of this trial-to-trial variability could be the locomotor activity of the  
218 mouse. Previous studies have shown that not only is some neural activity in the mouse visual  
219 cortex associated with the animal's running activity, but that visual responses are also  
220 modulated by such activity<sup>24–28</sup>. The mice in our experiments were free to run on a disc during  
221 the imaging sessions, and animals showed a range of running behaviors (**Supplemental Figure**  
222 **25**). Ignoring the stimulus, we found that some neurons' activities were correlated with the  
223 running speed (**Figure 5e**). While layer 5 showed strong correlations in all visual areas, in the  
224 other layers V1 had stronger correlations than the higher visual areas, with some visual areas  
225 showing median negative correlations. Within V1, the inhibitory interneurons showed the  
226 strongest correlation with running, most notably Vip neurons in layer 2/3, while the excitatory  
227 neurons showed weaker correlations.

228 For experiments in which the animals spent enough time running such that there were  
229 sufficient stimulus trials for a neuron's preferred condition when the mouse was both stationary  
230 and running (at least 10% trials for each), we compared the responses in these two states.  
231 Consistent with other reports, many neurons show modulated responses, but the effect was  
232 modest (**Figure 5f**). The majority of neurons showed enhanced responses. Considering the  
233 entire population, there was a 1.9 fold increase in the median evoked response. The effect on  
234 individual neurons, however, was varied such that only 13% of neurons showed significant  
235 modulation in these conditions ( $p<0.05$ , KS test).

236 To test whether running activity could account for the variability in trial-wise responses to  
237 visual stimuli, we included a binary running state as a condition dependent gain into the  
238 categorical regression described above (i.e. computing separate tuning curves for the running  
239 and stationary conditions, **Figure 5g**). This did not consistently and significantly improve the  
240 response prediction. Comparing the model performance when the running state is included to  
241 the performance based solely on the stimulus, we found that the distribution is largely centered  
242 along unity, with a slight asymmetry in favor of the running dependent model for the better  
243 performing models (**Figure 5h**, 28% of responsive neurons have  $r>0.5$  for stimulus x running  
244 state; 21% when considering all neurons). Thus, while the running activity of the mouse does  
245 modulate visual responses, it does not do so in a predictable way. This was further corroborated  
246 by a simpler model that predicts neural response based on the running speed (rather than a  
247 binary condition, and without stimulus information) (**Supplemental Figure 26**).

248 One of the unique aspects of this dataset is the broad range of stimuli, allowing for a  
249 comparison of response characteristics and model predictions across stimuli. Surprisingly,  
250 knowing whether a neuron responded to one stimulus type (e.g. natural scenes, drifting  
251 gratings, etc.) was largely uninformative of whether it responded to another stimulus type.  
252 Unlike the examples shown in Figure 2, which were chosen to highlight responses to all stimuli,  
253 most neurons were responsive to only a subset of the stimuli presented (**Figure 6a**). To explore  
254 the relationships between neural responses to different types of stimuli, we computed the  
255 Pearson correlation between the percent of responsive trials for each stimulus. This comparison  
256 removes the threshold of “responsiveness” and examines underlying patterns of activity. We  
257 found that most stimulus combinations were weakly correlated (**Figure 6b**), demonstrating that  
258 knowing that a neuron responds reliability to drifting gratings, for example, carries little to no  
259 information about how reliably that neuron responds to one of the natural movies. There is a  
260 higher correlation between the reliability of the responses to the natural movie that is repeated

261 across all three sessions (natural movies 1A, 1B, 1C), providing an estimate of the variability  
262 introduced by imaging across days and thus a ceiling for the overall correlations across stimuli.  
263 Very few of the cross stimulus correlations approach this ceiling, with the exception of the  
264 correlation between static gratings and natural scenes.

265 We characterize the variability by clustering the reliability, defined by the percentage of  
266 significant responses to repeated stimuli. We used a Gaussian mixture model to cluster the  
267 25,958 neurons that were imaged in both Sessions A and B (see **Figure 1f**) and thus excluded  
268 the Locally Sparse Noise stimulus due to the lack of a comparable definition of reliability. Using  
269 neurons imaged in all three sessions did not qualitatively change the results (see **Supplemental**  
270 **Figure 27**). The clusters are described by the mean percent responsive trials for each stimulus  
271 for each cluster (**Figure 6c**). Note that there is only a weak relationship between the percent  
272 responsive trials to one stimulus and any other. We grouped the clusters into “classes” by first  
273 defining a threshold for responsiveness by identifying the cluster with the lowest mean percent  
274 responsive trials across stimuli, then setting the threshold equal to the maximum value across  
275 stimuli plus one standard deviation for that cluster. This allowed us to identify each cluster as  
276 responsive (or not) to each of the stimuli. Clusters with the same profile (e.g. responsive to  
277 drifting gratings and natural movies, but not static gratings or natural scenes), were grouped into  
278 one of sixteen possible classes.

279 The clustering was performed 100 times with different initial conditions to evaluate  
280 robustness. The optimal number of clusters, evaluated with model comparison, as well as the  
281 class definition threshold were consistent across runs (**Supplemental Figure 27**). By far the  
282 largest single class revealed by this analysis is that of neurons that are largely unresponsive to  
283 all stimuli, termed “none,” which contains  $34\pm2\%$  of the neurons (**Figure 6d**). Other large  
284 classes include neurons that respond to drifting gratings and natural movies (“DG-NM”,  $14\pm3\%$

285 of neurons), to natural scenes and natural movies (“NS-NM”, 14±2% of neurons), and to all  
286 stimuli (“DG-SG-NS-NM”, 10±1% of neurons).

287 It is important to note that we do not observe all 16 possible stimulus response  
288 combinations. For instance, very few neurons are classified as responding to one stimulus  
289 alone, the most prominent exception being neurons that respond uniquely to natural movies. In  
290 many runs, the other singleton classes were empty. We never find clusters of neurons that  
291 respond only to static gratings. Thus, while the pairwise correlations between most stimuli are  
292 relatively weak, there is meaningful structure in the patterns of responses. Nevertheless, within  
293 each class there remains a great deal of heterogeneity. For example, within the class that  
294 responds to all stimuli, there is a cluster in which the neurons respond with roughly equal  
295 reliability to all four stimuli (cluster 27 in **Figure 6c**) as well as a cluster in which the neurons  
296 respond reliably to drifting and static gratings and only weakly to natural scenes and natural  
297 movies (clusters 25 and 28). This heterogeneity underlies the inability to predict whether a  
298 neuron responds to one stimulus given that it responds to another.

299 Classes are not equally represented in all visual areas (**Figure 6e**). The “unresponsive”  
300 class is larger in the higher visual areas than in V1, and is largest in RL (see also **Figure 2g**).  
301 Classes related to moving stimuli, including “NM”, “DG-NM”, and “DG”, have relatively flat  
302 distributions across the visual areas, excluding RL. The natural classes, including “NS-NM”,  
303 “DG-NS-NM”, “SG-NS-NM”, and “DG-SG-NS-NM”, are most numerous in V1 and LM, with lower  
304 representation in the other visual areas. This divergence in representation of the motion classes  
305 from the natural classes in areas AL, PM, and AM is consistent with the putative dorsal and  
306 ventral stream segregation in the visual cortex<sup>32</sup>.

307 In addition to differential representation across cortical areas, the response classes are  
308 also differentially represented among the Cre lines (**Figure 6f**). Notably, Sst interneurons in V1

309 have the fewest "none" neurons and the most neurons responsive to all stimuli. Meanwhile, the  
310 plurality of Vip interneurons are in the classes responsive to natural stimuli or specifically natural  
311 movies.

312 Having characterized neurons by their joint reliability to multiple stimuli, we next ask to  
313 what extent we can predict neural responses, not on a trial-by-trial basis but including the  
314 temporal response dynamics, given the stimulus and knowledge of the animal's running  
315 condition. We use a model class that remains in widespread use for predicting visual  
316 physiological responses and that captures both "simple" and "complex" cell behaviors. The  
317 model structure uses a dense wavelet basis (sufficiently dense to capture spatial and temporal  
318 features at the level of the mouse visual acuity and temporal response) and computes from this  
319 both linear and quadratic features, each of which are summed, along with the binary running  
320 trace convolved with a learned temporal filter, and sent through a soft rectification (**Figure**  
321 **7a**). We train these models on either the collective natural stimuli or the artificial stimuli to  
322 predict the extracted event trace. Whereas we find example neurons for which this model works  
323 extremely well (**Figure 7b**), across the population only 2% of neurons are well fit by this model  
324 ( $r>0.5$ ; 2% for natural stimuli; 1% for artificial stimuli, **Figure 7c**), with the median being ~0.2 (for  
325 natural stimuli). Model performance was slightly higher in V1 than in the higher visual areas, and  
326 showed little difference across Cre lines. It is also worth noting that there is a great deal of  
327 visually responsive activity that is not being captured by these models (**Supplementary Figure**  
328 **23**). Comparing the models' performances across stimulus categories, we found that the overall  
329 distribution of performance for models trained and tested with natural stimuli was higher than  
330 the corresponding models for artificial stimuli (**Figure 7d**), consistent with previous reports<sup>10–</sup>  
331 <sup>15</sup>. The running speed of the mouse did not add significant predictive power to the model, as  
332 most regression weights were near zero, with the exception of Vip neurons in V1  
333 (**Supplemental Figure 28**). Including the pupil area of the viewing eye, thought to be correlated

334 with the animal's arousal state, as a regressor had similarly small effects (**Supplemental Figure**  
335 **28**). Adjusting the stimulus according the pupil position on the monitor also had little effect on  
336 the model performance at the population level (**Supplemental Figure 28**). Finally, retraining  
337 without the quadratic weights, thereby only using the linear components of the model, did not  
338 appreciably affect the model performance at the population level (**Supplemental Figure 28**).

339 When comparing the model performance for the neurons in each of the classes defined  
340 through the clustering analysis, we found that these classes occupy spaces of model  
341 performance consistent with their definitions (**Figures 7e-h**). The "none" neurons formed a  
342 relatively tight cluster and constituted the bulk of the density close to the origin (**Figure 7e**). By  
343 definition, these neurons had the least response reliability for all stimuli (**Figure 6c**), and were  
344 likewise the least predictable. Neurons in the "NS-NM" class showed high model performance  
345 for natural stimuli and low performance for artificial stimuli (**Figure 7f**). And finally, neurons that  
346 reliably respond to all stimuli ("DG-SG-NS-NM"), showed a broad distribution of model  
347 performance, with the highest median performance, equally predicted by both artificial and  
348 natural stimuli (**Figure 7g**). As running has been shown to influence neural activity in these data  
349 independent of visual stimuli (**Figure 5e**), one might expect that the "none" class is composed  
350 largely of neurons that are strongly driven by running activity rather than visual stimuli. Instead,  
351 we found that the "none" class has one of the smallest median correlations, overall, with the  
352 running speed of the mouse, while the "DG-SG-NS-NM" class had the largest (**Figure 7i**).

353

354 **Discussion**

355 Data standardization and experimental reproducibility are both a challenge and an  
356 opportunity for systems neuroscience. *In vivo* neuronal recordings are notoriously difficult  
357 experiments that require an in-depth expertise in many scientific fields and multiple years of

358 training. As such, these experiments are difficult to perform on a large scale. Here we combined  
359 standardized operating procedures with integrated engineering tools to address these long-  
360 standing difficulties. Here we report the results from a dataset that contains over an order of  
361 magnitude more animals (243 mice) than is typically reported in the field and maintained a tight  
362 standardization across three years of continuous data collection.

363 Our pipeline reduced critical experimental biases by separating quality control of data  
364 collection from response characterization. Historically, the field has been dominated by single-  
365 neuron electrophysiological recordings in which electrodes were advanced until a neuron was  
366 found that responded to a test stimulus. The stimulus was then optimized to elicit the strongest  
367 reliable response from that neuron. The experiment proceeded using manipulations around this  
368 stimulus condition hand-tuned to drive the strongest response. Such studies discovered many  
369 characteristic response properties, but may have failed to capture the variability of responses,  
370 the breadth of features that will elicit a neural response, and the breadth of features that do not  
371 elicit a response. Recently, calcium imaging and denser electrophysiological recordings<sup>29</sup> have  
372 enabled large populations of neurons to be recorded simultaneously. By combining calcium  
373 imaging with strong quality control and standardization, we created an unprecedented survey of  
374 59,610 neurons in mouse visual cortex using a standard and well-studied but diverse set of  
375 stimuli while limiting the selection bias towards those stimuli.

376 Such a survey is crucial for assessing the successes and shortcomings of contemporary  
377 models of visual cortex, namely that visual responses can be reasonably characterized by  
378 combinations of linear filters with nonlinearities such as half-wave rectification, squaring, and  
379 response normalization<sup>6</sup>, or that neurons (in V1 at least) largely cluster into "simple" and  
380 "complex" cells under this characterization. A complete understanding of the visual system, and  
381 in particular V1, however, is limited by systematic bias in the measurement of neurons and in  
382 choices of stimuli that lead to experiments that reinforce assumptions of the standard model.

383 These challenges are exacerbated by difficulties in comparing models derived from different  
384 experimental data, that were collected under different assumptions with different stimuli, and  
385 were designed to predict different properties (e.g. responses on different timescales).<sup>5-7</sup> The  
386 Allen Brain Observatory provides an unprecedented and much needed benchmark for testing  
387 models of cortical computation by systematically surveying responses throughout the visual  
388 cortex of freely-running mice with both rich natural stimuli as well as the standard artificial stimuli  
389 most often used by the field to characterize visual responses.

390 Using standard noise and grating stimuli we find many of the standard visual response  
391 features, including orientation selectivity, direction selectivity, and spatial receptive fields with  
392 opponent on and off subfields (**Figures 2, 3**). Based on responses to these stimuli, we observed  
393 functional differences in visual responses across cortical areas, cortical layers, and transgenic  
394 Cre lines. In a novel analysis of overall reliabilities to both artificial and naturalistic stimuli, we  
395 found classes of neurons responsive to different constellations of stimuli (**Figure 6**). The  
396 different classes are largely intermingled, and found in all of the cortical areas recorded here,  
397 suggesting a largely parallel organization<sup>30</sup>. At the same time, the overrepresentation of classes  
398 responsive to natural movies and motion stimuli in areas AL, PM, and AM relative to the other  
399 classes (which are responsive to more spatial stimuli) is consistent with the assignment of these  
400 areas to the putative “dorsal” or “motion” stream<sup>31</sup>. The lack of an inverse relationship, wherein  
401 spatial information is overrepresented relative to motion in a putative “ventral” stream, likely  
402 reflects the fact that we were unable to image the putative ventral areas LI, POR, or P within our  
403 cranial window. Area LM has previous been loosely associated with the ventral stream, but with  
404 evidence that it is more similar to V1 than other higher order ventral areas<sup>9,31</sup>, and our results  
405 appear consistent with the latter. Area RL has the largest proportion of neurons in the “none”  
406 class, over 85%, also consistent with the very low percent of responsive neurons (**Figure 2**). It  
407 is possible that neurons in this area are specialized for visual features not probed here, or that

408 they show a greater degree of multi-modality than in the other visual areas, integrating  
409 somatosensory and visual features<sup>32</sup>.

410 One of the unique features of this dataset is that it includes a large number of different  
411 transgenic Cre lines for characterization that label specific populations of excitatory and  
412 inhibitory neurons. On a coarse scale, excitatory populations behave similarly, however closer  
413 examination reveals distinct functional properties across Cre lines. For instance, Rorb, Scnn1a-  
414 Tg3, and Nr5a1, which label distinct layer 4 populations in V1, exhibit distinct spatial and  
415 temporal tuning properties (**Figure 3, Supplemental Figure 20, 21**), different degrees of  
416 running correlation (**Figure 5**), and subtle differences in their class distribution (**Figure 6**).  
417 These differences suggest that there are separate channels of feedforward information.  
418 Similarly, comparing Fezf2 and Ntsr1 in V1, which label two distinct populations of corticofugal  
419 neurons found in layers 5 and 6 respectively, reveals that Fezf2 has larger spatial receptive  
420 fields and encodes lower temporal frequencies than Ntsr1. These Cre lines show further  
421 differences in the functional class distribution, where Ntsr1 has the most neurons in the “none”  
422 class while Fezf2 has the fewest in “none” and the most neurons in the “DG-SG-NS-NM” class,  
423 among the excitatory neurons. These differences indicate distinct feedback channels from V1.

424 The Brain Observatory data also provide the first broad survey of visually evoked  
425 responses of both Vip and Sst inhibitory Cre lines. Sst neurons are strongly driven by all visual  
426 stimuli used here, with the plurality belonging to the “DG-SG-NS-NM” class (**Figure 5f**). Their  
427 responses to drifting gratings are particularly robust in that 94% of Sst neurons in V1 are  
428 responsive to drifting gratings, and respond quite reliably across trials, far more than the other  
429 Cre lines (**Supplemental Figure 15, 24**). Vip neurons, on the other hand, are largely  
430 unresponsive, even suppressed, by drifting gratings, with only 9% of Vip neurons in V1  
431 responding to the stimulus. This extreme difference in responses to drifting gratings between  
432 these two populations is consistent with previous literature examining the size tuning of these

433 interneurons, and supports the disinhibitory circuit between them<sup>24,33,34</sup>. Vip neurons, however,  
434 are very responsive to both natural scenes and natural movies, the majority falling in the “NS-  
435 NM” class (**Figure 6f**, **Supplemental Figure 18, 19, 27**), but show little selectivity to these  
436 stimuli, as their median lifetime sparseness is lower than both the Sst and the excitatory  
437 neurons (not shown). Interestingly, receptive field mapping using locally sparse noise revealed  
438 that Vip neurons in V1 have remarkably large receptive field areas, larger than both Sst and  
439 excitatory neurons (**Figure 3f**), in contrast to the smaller summation area for Vip neurons,  
440 previously measured using windowed drifting gratings<sup>24,35</sup>. This suggests that Vip neurons  
441 respond to small features over a large region of space. Both inhibitory interneuron Cre lines  
442 have higher noise correlations (**Figure 4c**), suggesting that these populations might function  
443 more coherently. In addition, both populations show strong running modulation: they both  
444 correlate stronger with the mouse’s running speed than the excitatory populations (**Figure 5e**),  
445 and a model based solely on the mouse’s running speed does a better job at predicting their  
446 activity than for the excitatory populations (**Supplemental Figure 26**). The widespread  
447 suppression of Vip neuron responses to drifting gratings is not observed in area PM  
448 (**Supplemental Figure 15**), suggesting that this inhibitory circuit likely functions differently in  
449 different cortical areas.

450 Beyond coarse characterizations of response properties, the true test of a model of  
451 cortical activity is the ability to predict novel responses, both to the types of stimuli used for  
452 characterizing the neuron and, more importantly, to arbitrary stimuli. Using the trial-averaged  
453 tuning curve from grating responses to predict neural activity on a trial-by-trial basis only  
454 predicts the responses of a minority of neurons, and is not significantly improved by including  
455 running (**Figure 7**). This discrepancy is even more extreme for wavelet models that attempt to  
456 predict the time course of response to a stimulus.

457           Neurons in the “DG-SG-NS-NM” class were well predicted, with values comparable to  
458   those found in primates<sup>6,11,36,37</sup>, for both natural and artificial stimuli (**Figure 7g**). Based on the  
459   way we chose our stimulus parameters, we expect that neurons with a strong “classical  
460   receptive field” would be most likely to appear in this class. However, this class constitutes only  
461   10% of the mouse visual cortex (**Figure 6d**). Neurons in the “NS-NM” class show equally high  
462   prediction for natural stimuli, but are poorly predicted of, and do not reliably respond to, artificial  
463   stimuli (**Figure 7f**). It is possible these neurons could be “classical” neurons as well, but are  
464   tuned for spatial or temporal frequencies that were not included in our stimulus set. As our  
465   stimulus parameters were chosen to match previous measurements of mouse acuity, this could  
466   suggest that the acuity of mouse has been underestimated<sup>38</sup>.

467           Remarkably, the largest class of neurons was the “none” class, constituting those  
468   neurons that did not respond reliably to any of the stimuli, which contained 34% of the neurons  
469   in mouse visual cortex. These neurons are the least likely to be described by “classical  
470   receptive fields,” as evidenced by their poor model performance for both natural and artificial  
471   stimuli (**Figure 7f**). What, then, do these neurons do? It is possible these neurons are in fact  
472   visually driven, but are responsive to highly sparse and specific natural features that may arise  
473   through hierarchical processing<sup>39</sup>. Indeed, the field has a growing body of evidence that the  
474   rodent visual system exhibits sophisticated computations. For instance, neurons as early as V1  
475   show visual responses to complex stimulus patterns<sup>40</sup>. Alternatively, these neurons could be  
476   involved in non-visual computation, including behavioral responses such as reward timing and  
477   sequence learning<sup>41</sup>, as well as modulation by multimodal sensory stimuli<sup>41,42</sup> and motor  
478   signals<sup>25,28,43–45</sup>. While we found little evidence that these neurons were correlated with the  
479   mouse’s running activity recent work has found running to be among the least predictive such  
480   motor signals<sup>45</sup>.

481        Given these results and the many open questions regarding visual representation and  
482 computation in the mouse visual cortex, we believe that the Allen Brain Observatory provides an  
483 important foundational resource for the community. This is an openly available dataset,  
484 accessible via a dedicated web portal (<http://observatory.brain-map.org/visualcoding>), with a  
485 custom Python-based Application Programming Interface (API), the AllenSDK  
486 (<http://alleninstitute.github.io/AllenSDK/>). In addition to providing a benchmark for future  
487 experimental results, these data can serve as a testbed for theories of cortical computation and  
488 for model development. Already, these data have been used by other researchers to develop  
489 image processing methods<sup>46,47</sup>, to examine stimulus encoding and decoding<sup>48–51</sup>, and to test  
490 models of cortical computations<sup>52</sup>. Ultimately, we expect these data will seed as many questions  
491 as they answer, fueling others to pursue both new analyses and further experiments to unravel  
492 how cortical circuits represent and transform sensory information.

493

494    **References**

- 495    1. Hubel, D. & Wiesel, T. Receptive fields of single neurones in the cat's striate cortex. *J.  
496       Physiol.* **148**, 574–591 (1959).
- 497    2. Hubel, D. H. & Wiesel, T. N. Receptive fields, binocular interaction and functional  
498       architecture in the cat's visual cortex. *J. Physiol.* **160**, 106–154.2 (1962).
- 499    3. Felleman, D. J. & Van Essen, D. C. Distributed Hierarchical Processing in the Primate  
500       Cerebral Cortex.
- 501    4. DiCarlo, J. J., Zoccolan, D. & Rust, N. C. How does the brain solve visual object  
502       recognition? *Neuron* (2012). doi:10.1016/j.neuron.2012.01.010
- 503    5. Olshausen, B. & Field, D. What is the other 85 % of V1 doing? *Probl. Syst.* ... 1–29  
504       (2004). doi:10.1093/acprof:oso/9780195148220.003.0010
- 505    6. Carandini, M. *et al.* Do we know what the early visual system does? *J. Neurosci.* **25**,  
506       10577–97 (2005).
- 507    7. Masland, R. H. & Martin, P. R. The unsolved mystery of vision. *Curr. Biol.* **17**, R577-82  
508       (2007).
- 509    8. Andermann, M. L., Kerlin, A. M., Roumis, D. K., Glickfeld, L. L. & Reid, R. C.  
510       Functional specialization of mouse higher visual cortical areas. *Neuron* **72**, 1025–1039  
511       (2011).
- 512    9. Marshel, J. H., Garrett, M. E., Nauhaus, I. & Callaway, E. M. Functional specialization of  
513       seven mouse visual cortical areas. *Neuron* **72**, 1040–1054 (2011).
- 514    10. Fournier, J., Monier, C., Pananceau, M. & Frégnac, Y. Adaptation of the simple or  
515       complex nature of V1 receptive fields to visual statistics. *Nat. Neurosci.* **14**, 1053–60  
516       (2011).
- 517    11. David, S., Vinje, W. & Gallant, J. L. Natural Stimulus Statistics Alter the Receptive Field  
518       Structure of V1 Neurons. *J. Neurosci.* **24**, 6991–7006 (2004).

- 519 12. Talebi, V. & Baker, C. L. Natural versus Synthetic Stimuli for Estimating Receptive Field  
520 Models: A Comparison of Predictive Robustness. *J. Neurosci.* **32**, 1560–1576 (2012).
- 521 13. Yeh, C.-I., Xing, D., Williams, P. & Shapley, R. Stimulus ensemble and cortical layer  
522 determine V1 spatial receptive fields. *Proc. Natl. Acad. Sci.* **106**, 14652–14657 (2009).
- 523 14. Sharpee, T. O. *et al.* Adaptive filtering enhances information transmission in visual cortex.  
524 **439**, 936–942 (2006).
- 525 15. Felsen, G., Touryan, J., Han, F. & Dan, Y. Cortical Sensitivity to Visual Features in  
526 Natural Scenes. **3**, (2005).
- 527 16. Averbeck, B. B., Latham, P. E. & Pouget, A. Neural correlations, population coding and  
528 computation. *Nature Reviews Neuroscience* (2006). doi:10.1038/nrn1888
- 529 17. Jewell, S. & Witten, D. Exact spike train inference via  $\ell_0$  optimization. *Ann. Appl. Stat.*  
530 **12**, 2457–2482 (2018).
- 531 18. Jewell, S., Hocking, T. D., Fearnhead, P. & Witten, D. Fast Nonconvex Deconvolution of  
532 Calcium Imaging Data. *Biostatistics* (2019).
- 533 19. Sun, W., Tan, Z., Mensh, B. D. & Ji, N. Thalamus provides layer 4 of primary visual  
534 cortex with orientation- and direction-tuned inputs. *Nat. Neurosci.* **1**, (2015).
- 535 20. Rolls, E. T. & Tovee, M. J. Sparseness of the neuronal representation of stimuli in the  
536 primate temporal visual cortex. *J. Neurophysiol.* **73**, 713–726 (1995).
- 537 21. Vinje, W. E. & Gallant, J. L. Sparse Coding and Decorrelation in Primary Visual Cortex  
538 During Natural Vision. *Science (80-.)*. **287**, 1273–1276 (2000).
- 539 22. Moreno-bote, R. *et al.* Information-limiting correlations. *Nat. Publ. Gr.* **17**, 1410–1417  
540 (2014).
- 541 23. Kohn, A., Coen-cagli, R., Kanitscheider, I. & Pouget, A. Correlations and Neuronal  
542 Population Information. 237–258 doi:10.1146/annurev-neuro-070815-013851
- 543 24. Dipoppa, M. *et al.* Vision and Locomotion Shape the Interactions between Neuron Types

- 544                  in Mouse Visual Cortex. *Neuron* **98**, 602–615.e8 (2018).
- 545    25. Niell, C. M. & Stryker, M. P. Modulation of visual responses by behavioral state in mouse  
546                  visual cortex. *Neuron* **65**, 472–9 (2010).
- 547    26. Dadarlat, M. C. & Stryker, M. P. Locomotion enhances neural encoding of visual stimuli  
548                  in mouse V1. **37**, 3764–3775 (2017).
- 549    27. Polack, P. O., Friedman, J. & Golshani, P. Cellular mechanisms of brain state-dependent  
550                  gain modulation in visual cortex. *Nat. Neurosci.* **16**, 1331–1339 (2013).
- 551    28. Saleem, A., Ayaz, A., Jeffery, K., Harris, K. & Carandini, M. Integration of visual motion  
552                  and locomotion in mouse visual cortex. *Nat. Neurosci.* **16**, 1864–1869 (2013).
- 553    29. Jun, J. J. *et al.* Fully integrated silicon probes for high-density recording of neural activity.  
554                  *Nature* **551**, 232–236 (2017).
- 555    30. Han, Y. *et al.* The logic of single-cell projections from visual cortex. *Nature* **556**, 51–56  
556                  (2018).
- 557    31. Wang, Q., Sporns, O. & Burkhalter, A. Network analysis of corticocortical connections  
558                  reveals ventral and dorsal processing streams in mouse visual cortex. *J. Neurosci.* **32**,  
559                  4386–99 (2012).
- 560    32. Olcese, U., Iurilli, G. & Medini, P. Cellular and synaptic architecture of multisensory  
561                  integration in the mouse neocortex. *Neuron* **79**, 579–593 (2013).
- 562    33. Pfeffer, C. K., Xue, M., He, M., Huang, Z. J. & Scanziani, M. Inhibition of inhibition in  
563                  visual cortex: the logic of connections between molecularly distinct interneurons. *Nat.*  
564                  *Neurosci.* **16**, 1068–1076 (2013).
- 565    34. Fu, Y. *et al.* A cortical circuit for gain control by behavioral state. *Cell* **156**, 1139–1152  
566                  (2014).
- 567    35. Adesnik, H., Bruns, W., Taniguchi, H., Huang, Z. J. & Scanziani, M. A neural circuit for  
568                  spatial summation in visual cortex. *Nature* **490**, 226–31 (2012).

- 569 36. Mcfarland, X. J. M., Cumming, B. G. & Butts, X. D. A. Variability and Correlations in  
570 Primary Visual Cortical Neurons Driven by Fixational Eye Movements. **36**, 6225–6241  
571 (2016).
- 572 37. Vintch, B., Movshon, J. A. & Simoncelli, E. P. A Convolutional Subunit Model for  
573 Neuronal Responses in Macaque V1. **35**, 14829–14841 (2015).
- 574 38. Dyballa, L., Hoseini, M. S., Dadarlat, M. C., Zucker, S. W. & Stryker, M. P. Flow stimuli  
575 reveal ecologically-appropriate responses in mouse visual cortex. *bioRxiv* 1–9 (2018).  
576 doi:10.1101/362459
- 577 39. Yosinski, J., Clune, J. & Fuchs, T. Understanding Neural Networks Through Deep  
578 Visualization. (2015).
- 579 40. Palagina, G., Meyer, J. F. & Smirnakis, S. M. Complex Visual Motion Representation in  
580 Mouse Area V1. *J. Neurosci.* **37**, 164–183 (2017).
- 581 41. Bieler, M. *et al.* Rate and Temporal Coding Convey Multisensory Information in Primary  
582 Sensory Cortices. *Eneuro* **4**, ENEURO.0037-17.2017 (2017).
- 583 42. Ibrahim, L. A. *et al.* Cross-Modality Sharpening of Visual Cortical Processing through  
584 Layer-1-Mediated Inhibition and Disinhibition. *Neuron* **89**, 1031–1045 (2016).
- 585 43. Keller, G. B., Bonhoeffer, T. & Hübener, M. Sensorimotor mismatch signals in primary  
586 visual cortex of the behaving mouse. *Neuron* **74**, 809–15 (2012).
- 587 44. Stringer, C. *et al.* Spontaneous behaviors drive multidimensional, brain-wide population  
588 activity. *bioRxiv* 306019 (2018). doi:10.1101/306019
- 589 45. Musall, S., Kaufman, M. T., Gluf, S. & Churchland, A. Movement-related activity  
590 dominates cortex during sensory-guided decision making. *bioRxiv* 308288 (2018).  
591 doi:10.1101/308288
- 592 46. Petersen, A., Simon, N. & Witten, D. SCALPEL: Extracting Neurons from Calcium  
593 Imaging Data. 1–31 (2017).
- 594 47. Sheintuch, L. *et al.* Tracking the Same Neurons across Multiple Days in Ca<sup>2+</sup>Imaging

- 595 Data. *Cell Rep.* **21**, 1102–1115 (2017).
- 596 48. Ellis, R. J. *et al.* High-accuracy Decoding of Complex Visual Scenes from Neuronal  
597 Calcium Responses. *bioRxiv* 1–32 (2018). doi:<https://doi.org/10.1101/271296>
- 598 49. Cai, L., Wu, B. & Ji, S. Neuronal Activities in the Mouse Visual Cortex Predict Patterns  
599 of Sensory Stimuli. (2018).
- 600 50. Zylberberg, J. Untuned but not irrelevant: A role for untuned neurons in sensory  
601 information coding. *bioRxiv* 1–18 (2017). doi:10.1101/134379
- 602 51. Christensen, A. J. & Pillow, J. W. Running reduces firing but improves coding in rodent  
603 higher- order visual cortex. *bioRxiv* 1–14 (2017). doi:<https://doi.org/10.1101/214007>
- 604 52. Sweeney, Y. & Clopath, C. Population coupling predicts the plasticity of stimulus  
605 responses in cortical circuits. (2018). doi:10.1101/265041
- 606

607 **Data and code availability.**

608 This is an openly available dataset, accessible via a dedicated web portal

609 (<http://observatory.brain-map.org/visualcoding>), with a custom Python-based Application

610 Programming Interface (API), the AllenSDK (<http://alleninstitute.github.io/AllenSDK/>). In

611 addition, code for analyses presented in this paper are available at

612 [https://github.com/alleninstitute/visual\\_coding\\_2p\\_analysis](https://github.com/alleninstitute/visual_coding_2p_analysis).

613

614 **Acknowledgements**

615

616 We thank the Animal Care, Transgenic Colony Management and Lab Animal Services for mouse

617 husbandry. We thank Z. Josh Huang for the use of the Fezf2-CreER line. We thank Daniel

618 Denman, Josh Siegle, Yazan Billeh and Anton Arkhipov for critical feedback on the manuscript.

619 This work was supported by the Allen Institute, and in part by NIH 2R90DA033461, NSF DMS-

620 1514743, Falconwood Foundation, Center for Brains, Minds & Machines funded by NSF

621 Science and Technology Center Award CCF-1231216, Natural Sciences and Engineering

622 Research Council of Canada, NIH Grant DP5OD009145, and NSF CAREER Award DMS-

623 1252624. We thank Allan Jones for providing the critical environment that enabled our large

624 scale team effort. We thank the Allen Institute founder, Paul G Allen, for his vision,

625 encouragement, and support.

626

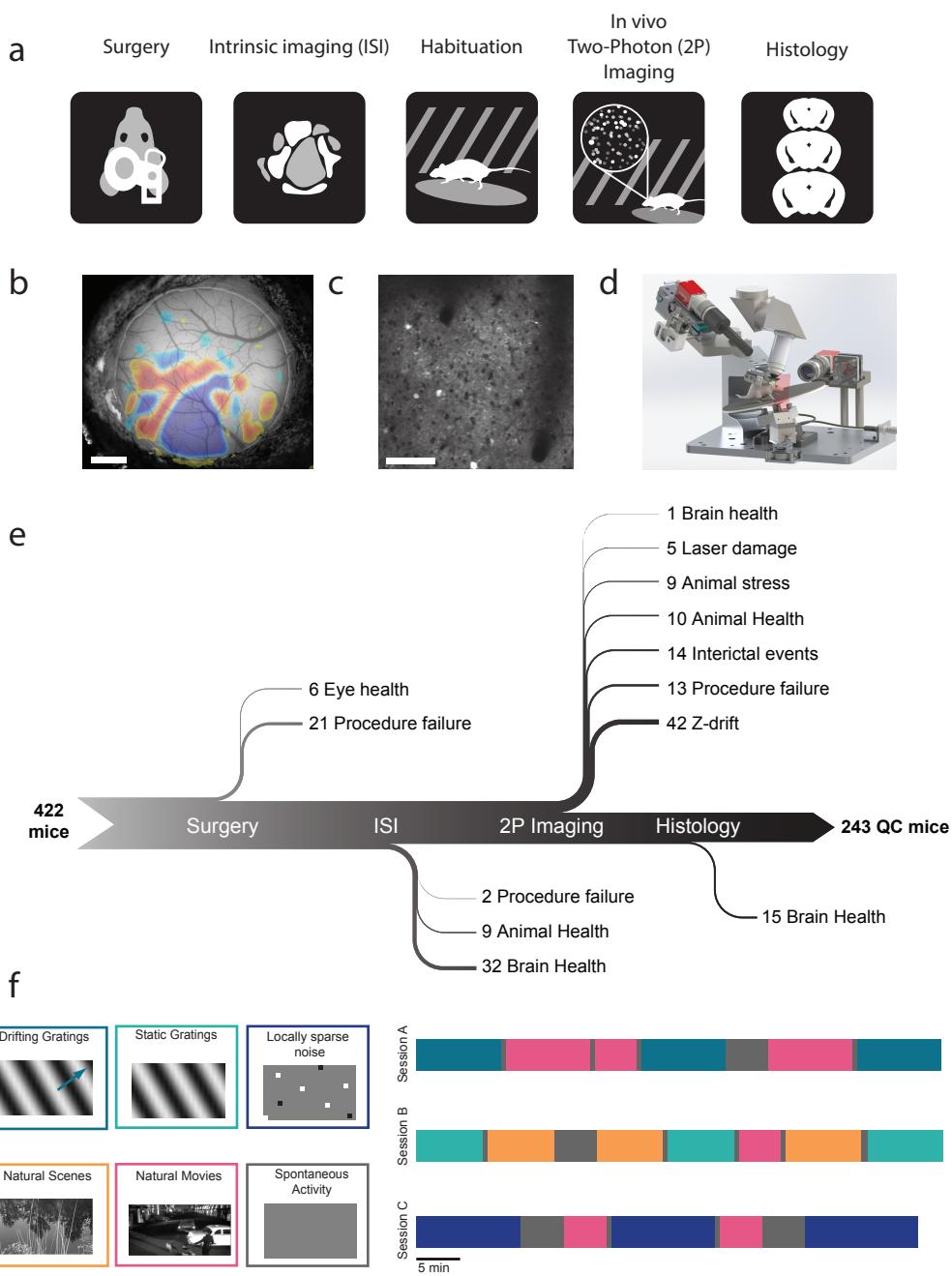
627 **Author Contributions**

628

629 SEJdV, MAB, KR, MG, TK, SM, SO, JW, CD, LN, AB, JWP, RCR, and CK conceived of and

630 designed the experiment. JL, TK, PH, AL, CS, DS, and CF built and maintained the hardware.

631 SEJdV, JL, MAB, GKO, DF, NC, LK, WW, DW, RV, CB, BB, TD, JG, SJ, NK, CL, FL, JP,  
632 NS, DW, JZ, and LN developed algorithms and software, including the SDK and website. KR,  
633 NB, NB, SC, LC, AC, SC, ME, NG, FG, RH, LH, UK, JL, RL, EL, LL, JL, KM, TN, MR, SS,  
634 CW, and AW collected data. JL and PAG managed the pipeline. SEJdV, JL, MAB, GKO, MO,  
635 NC, PL, DM, and RV analyzed data. SEJdV, JL, and MAB wrote the paper with input from  
636 PAG, GKO, MO, NC, PL, DM, RCR, and MG.  
637  
638



639

## Figure Legends

### 640 **Figure 1: A standardized systems neuroscience data pipeline to map visual responses**

641 (a) Schematic describing the workflow followed by each mouse going through our large scale  
642 data pipeline. (b) Examplar intrinsic imaging map labelling individual visual brain areas. Scale  
643 bar = 1mm. (c) Examplar averaged two photon imaging field of view (400  $\mu$ m x 400  $\mu$ m)  
644 showcasing neurons labeled with GCaMP6f. Scale bar = 100  $\mu$ m. (d) Custom design apparatus  
645 to standardize the handling of mice in two photon imaging. We engineered all steps of the  
646 pipeline to co-register data and tools, enabling reproducible data collection and a standardized  
647 experimental process (see **Supplementary Figure 1-4**). (e) Number of mice passing Quality  
648 Control (QC) criteria established by Standardized Operating Procedures (SOPs) at each step of  
649 the data collection pipeline with their recorded failure reason. The data collection pipeline is  
650 closely monitored to maintain consistently high data quality. (f) Standardized experimental  
651 design of sensory visual stimuli to map responses properties of neurons across the visual  
652 cortex. 6 blocks of different stimuli were presented to mice (left) and were distributed into 3  
653 separate imaging session called session A, session B and session C across different days  
654 (right).

655

656

**Table 1**

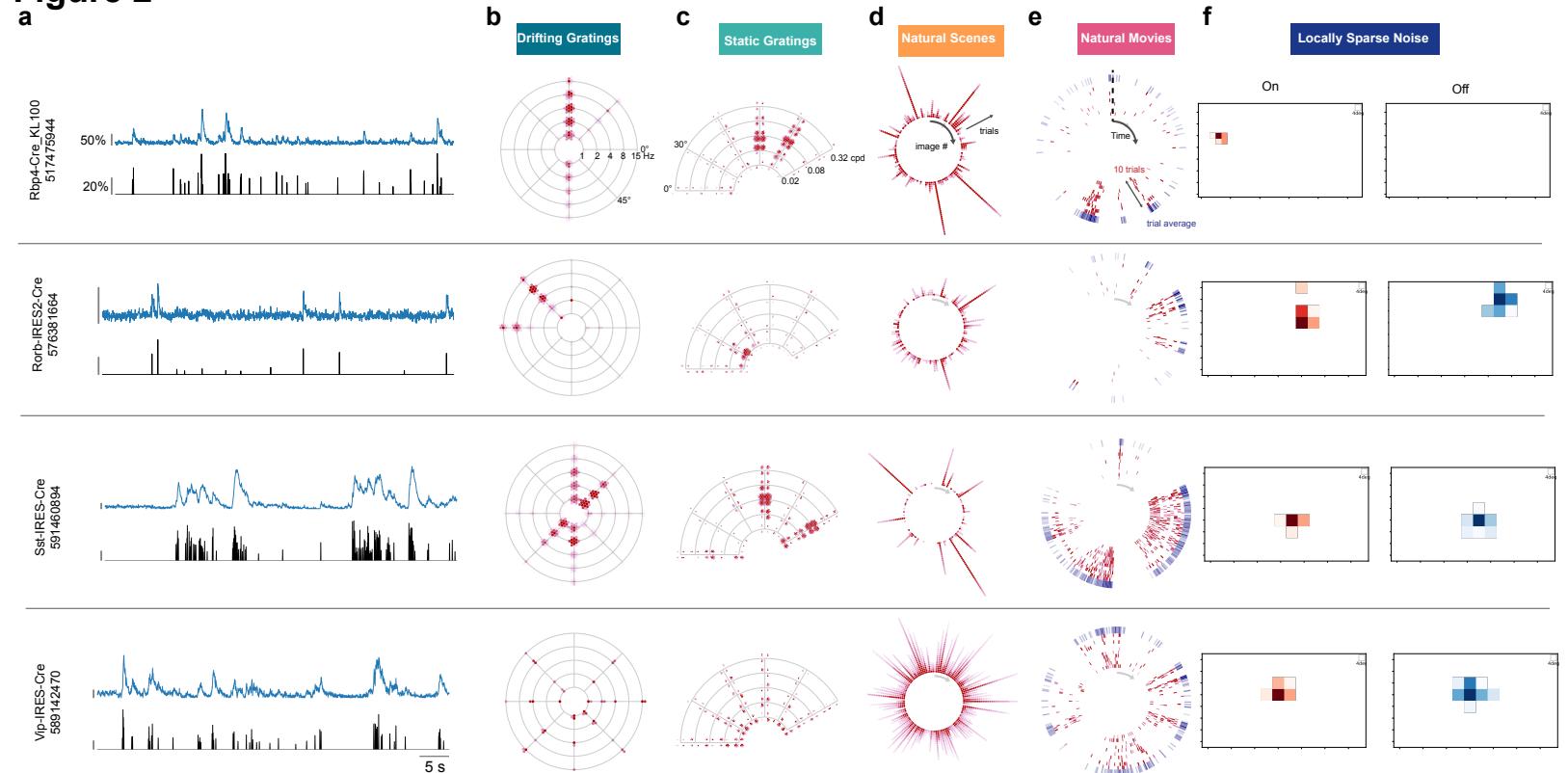
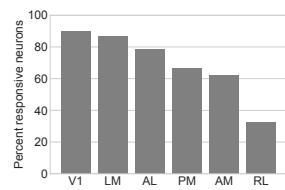
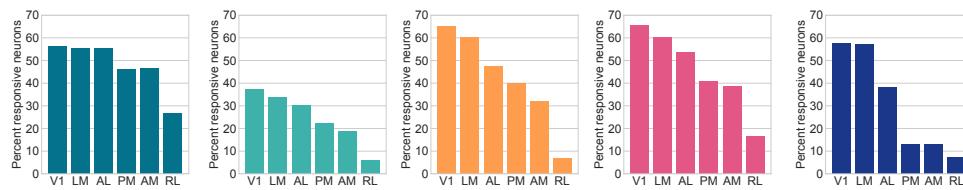
<b>Cre line</b>	<b>Layers</b>	<b>E/I</b>	<b>V1</b>	<b>LM</b>	<b>AL</b>	<b>PM</b>	<b>AM</b>	<b>RL</b>
Emx1-IRES-Cre; Camk2a-tTA; Ai93	2/3,4,5	E	3073 (10)	2098 (8)	1787 (7)	835 (4)	457 (3)	2152 (9)
Slc17a7-IRES2-Cre; Camk2a-tTA; Ai93	2/3,4,5	E	4840 (17)	3230 (16)	374 (2)	1970 (15)	235 (2)	137 (2)
Cux2-CreERT2; Camk2-tTA; Ai93	2/3, 4	E	5081 (16)	2792 (11)	3103 (13)	2361 (13)	1616 (11)	1578 (12)
Rorb-IRES2-Cre; Camk2a-tTA; Ai93	4	E	2218 (8)	1191 (6)	1242 (6)	764 (7)	735 (8)	1126 (5)
Scnn1a-Tg3-Cre; Camk2a-tTA; Ai93	4	E	1873 (9)					
Nr5a1-Cre; Camk2a-tTA; Ai93	4	E	578 (8)	421 (6)	220 (6)	331 (7)	171 (6)	1354 (6)
Rbp4-Cre_KL100; Camk2a-tTA; Ai93	5	E	458 (7)	485 (7)	441 (6)	509 (6)	355 (8)	93 (4)
Fezf2-CreER;Ai148 (corticothalamic)	5	E	407 (4)	981 (5)				
Tlx3-Cre_PL56;Ai148 (cortico-cortical)	5	E	1181 (6)	946 (3)				
Ntsr1-Cre_GN220;Ai148	6	E	573 (6)	719 (7)		581 (5)		
Sst-IRES-Cre;Ai148	4, 5	I	266 (17)	301 (15)	24 (1)	247 (14)		46 (2)
Vip-IRES-Cre;Ai148	2/3, 4	I	352 (17)	315 (17)		387 (16)		

657 **Table 1: Visual coding dataset.**

658 The number of cells (and experiments) imaged for each Cre line in each cortical visual area. In  
659 total, 59,610 cells imaged in 432 experiments in 243 mice are included in this dataset.

660

661

**Figure 2****g****h**

662 **Figure 2: Neurons exhibit diverse responses to visual stimuli.**

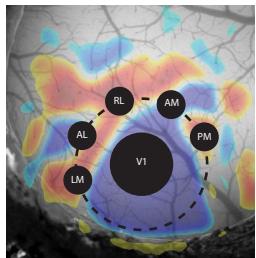
663 (a) Activity for four example neurons, two excitatory neurons (Rorb, layer 4, Rbp4, layer 5) and  
664 two inhibitory neurons (Sst layer 4, and Vip layer 2/3).  $\Delta F/F$  (top, blue) and extracted events  
665 (bottom, black) for each cell. (b) “Star” plot summarizing orientation and temporal frequency  
666 tuning for responses to the drifting gratings stimulus (For details on response visualizations see  
667 **Supplemental Figure 14**). (c) “Fan” plot summarizing orientation and spatial frequency tuning  
668 for responses to static gratings. (d) “Corona” plot summarizing responses to natural scenes.  
669 Each arm represents the response to an image, with individual trials being represented by  
670 circles whose color represents the strength of the response on that trial. (e) “Track” plot  
671 summarizing responses to natural movies. The response is represented as a raster plot moving  
672 clockwise around the circle. Ten trials are represented in red, along with the mean PSTH in the  
673 outer ring, in blue. (f) Receptive field subfields mapped using locally sparse noise. (g) Percent of  
674 neurons that responded to at least one stimulus across cortical areas. (h) Percent of neurons  
675 that responded to each stimulus across cortical areas. Colors correspond to the labels at the top  
676 of the figure.

677

678

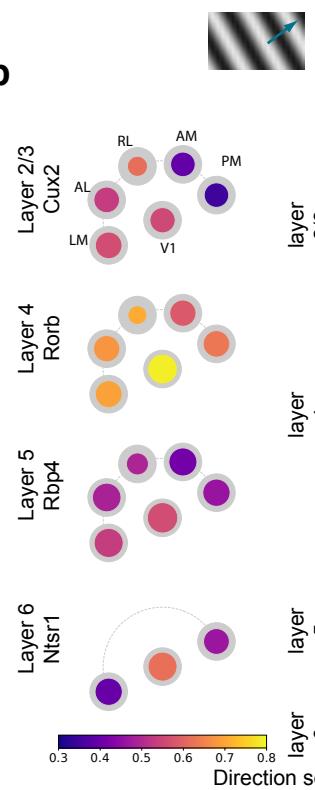
# Figure 3

a



area = fraction of responsive cells  
color = median value

b



layer  
2/3

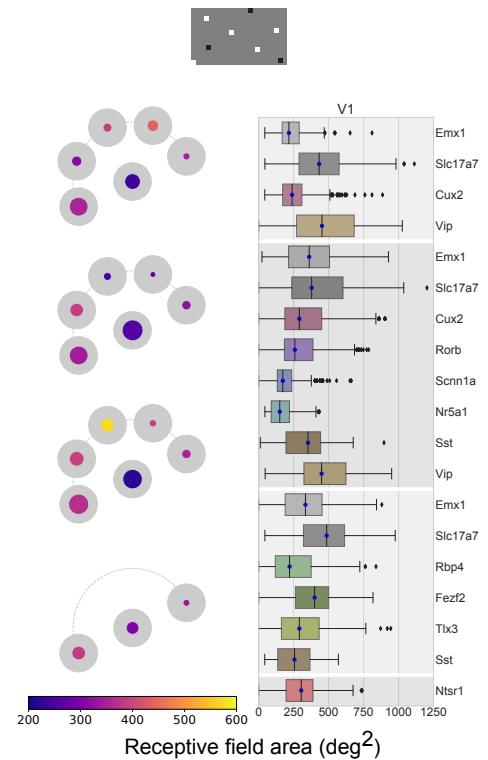
layer  
4

layer  
5

layer  
6

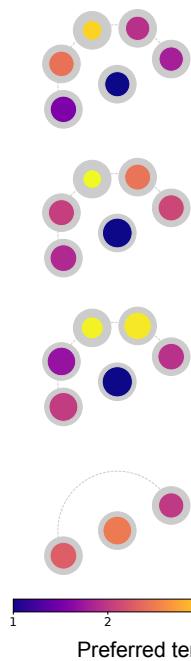
Direction selectivity  
0.3 0.4 0.5 0.6 0.7 0.8

c



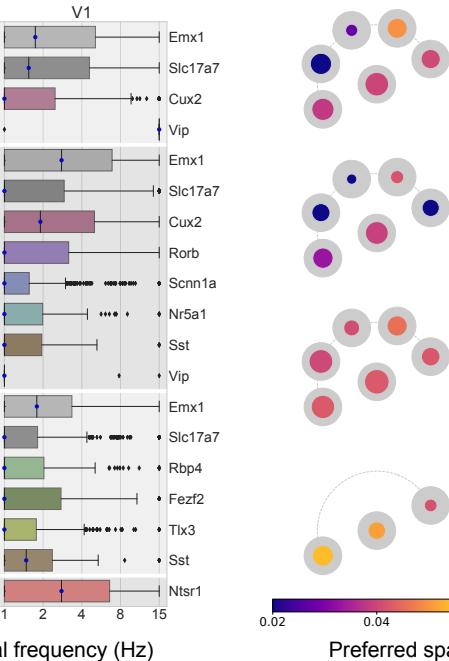
Receptive field area (deg<sup>2</sup>)  
200 300 400 500 600

d



Preferred temporal frequency (Hz)  
1 2 4

e

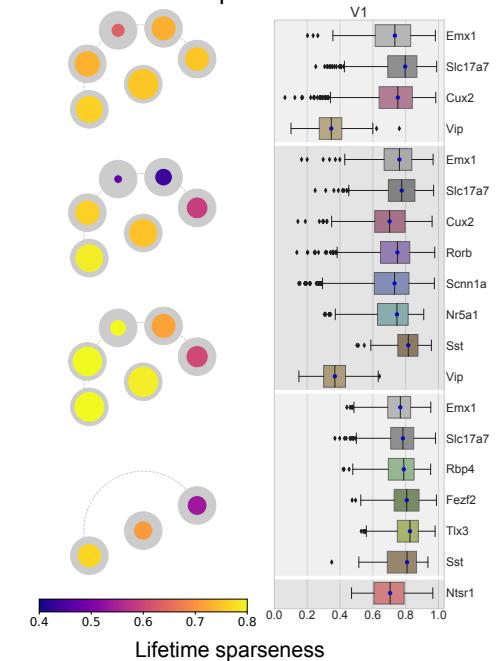


Preferred spatial frequency (cpd)  
0.02 0.04 0.08 0.16 0.32

f



Lifetime sparseness



Lifetime sparseness  
0.4 0.5 0.6 0.7 0.8

679 **Figure 3: Tuning properties reveal functional differences across areas and Cre lines.**

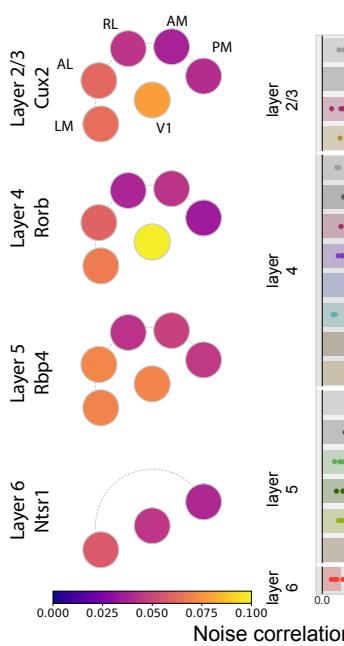
680 (a) “Pawplot” visualization summarizes median value of a tuning metric across visual areas.  
681 Top, each visual area is represented as a circle, with V1 in the center and the higher visual  
682 areas surrounding it according to their location on the surface of the cortex. Bottom, each paw-  
683 pad (visual area) has two concentric circles. The area of the inner, colored, circle relative to the  
684 outer circle represents the proportion of responsive cells for that layer and area. The color of the  
685 inner circle reflects the median value of the metric for the responsive cells, indicated by the  
686 colorscale at the bottom of the plot. For a metric’s summary plot, four pawplots are shown, one  
687 for each layer. Only data from one Cre line is shown for each layer. For each panel, a pawplot is  
688 paired with a box plot or a strip plot (for single cell and population metrics respectively) showing  
689 the full distribution for each Cre line and layer in V1. Data is assigned to cortical layers based on  
690 both the Cre line and the imaging depth. Data collected above 250um from the surface is  
691 considered to be in layer 2/3. Data collected between 250µm and 365µm is considered to be in  
692 layer 4. Data collected between 375µm and 500µm is considered to be in layer 5. Data collected  
693 at 550µm is considered to be in layer 6. The box shows the quartiles of the data, and the  
694 whiskers extend to 1.5 times the interquartile range. Points outside this range are shown as  
695 outliers. For other cortical areas, see **Supplemental Figure 20**. (b) Pawplot and box plot  
696 summarizing direction selectivity. (c) Pawplot and box plot summarizing receptive field area. (d)  
697 Pawplot and box plot summarizing preferred temporal frequencies. (e) Pawplot and box plot  
698 summarizing preferred spatial frequencies. (f) Pawplot and box plot summarizing lifetime  
699 sparseness of responses to natural scenes.

700

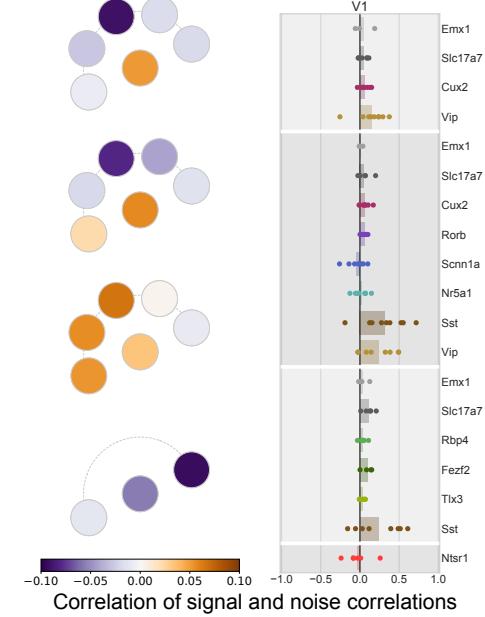
701

**Figure 4**

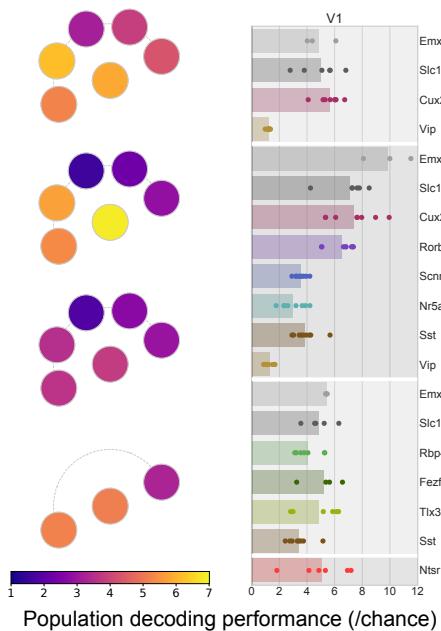
**a**



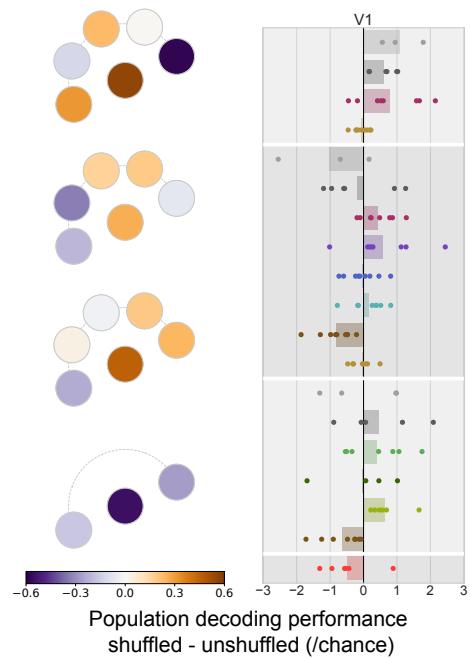
**b**



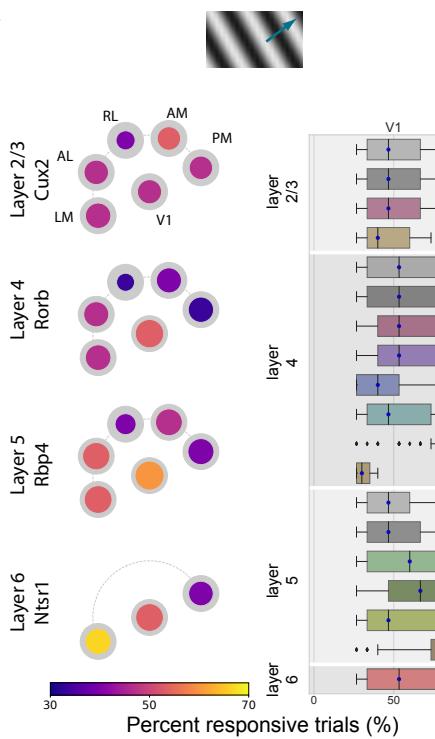
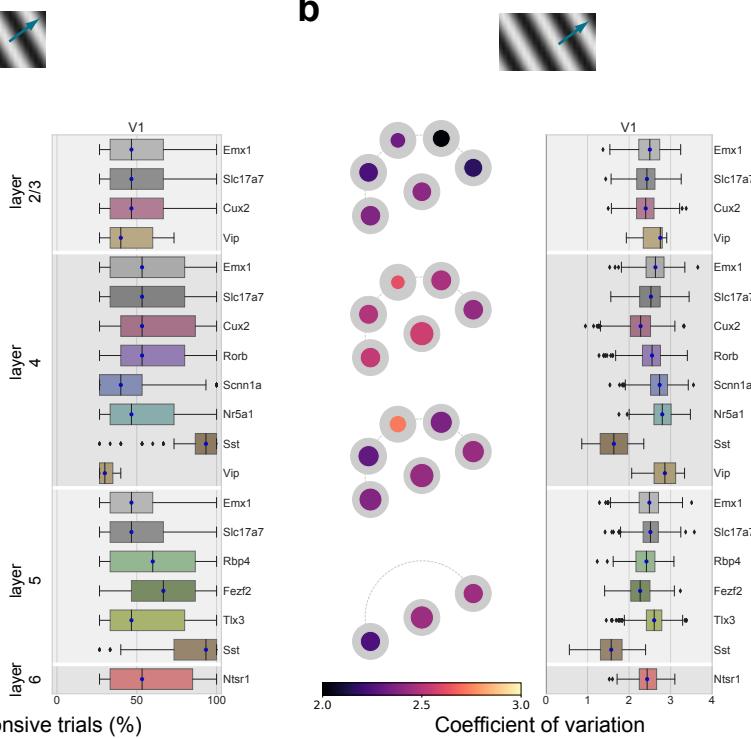
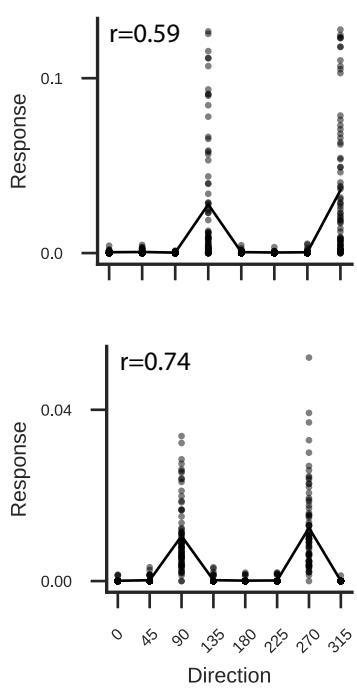
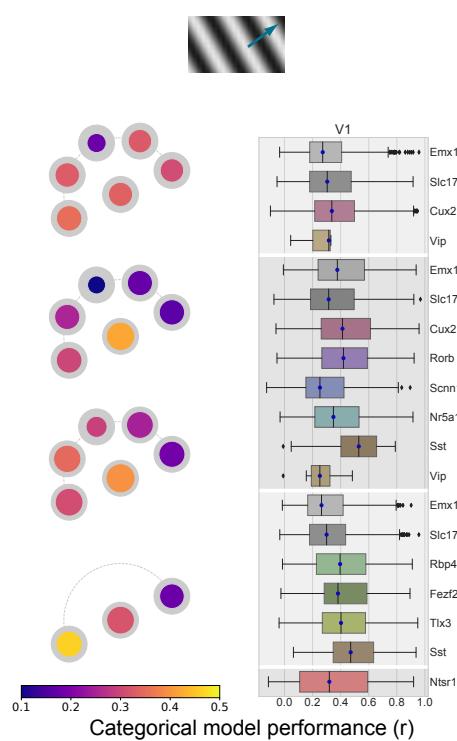
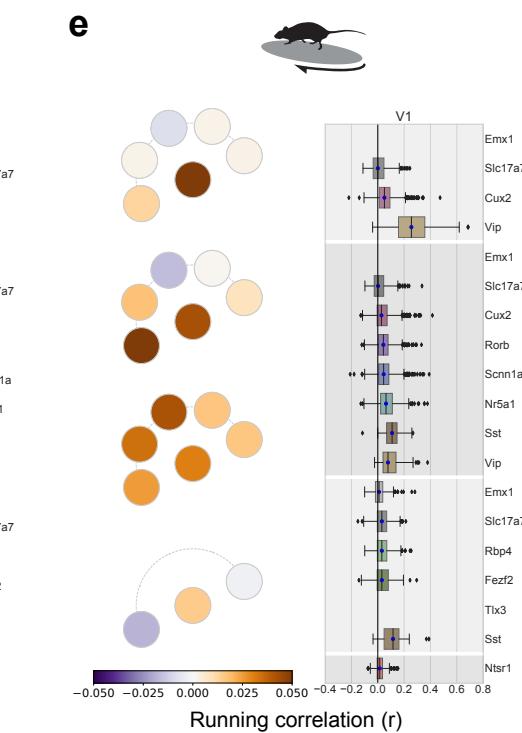
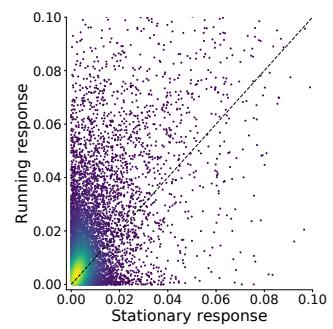
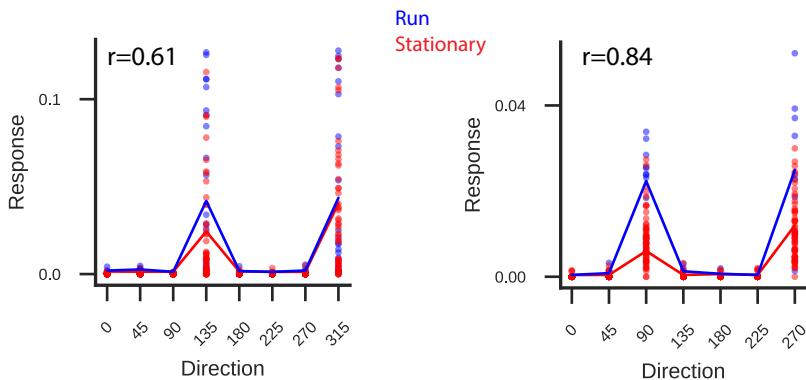
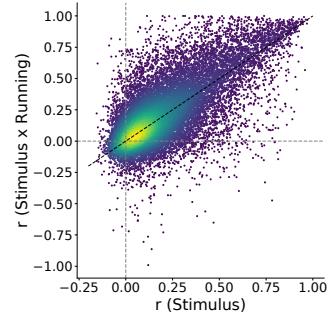
**c**



**d**



702 **Figure 4: Population correlations have heterogenous impact on decoding performance**  
703 (a) Pawplot and strip plot summarizing decoding performance for drifting grating direction using  
704 K-nearest neighbors. Each dot represents the mean five-fold cross-validated decoding  
705 performance of a single experiment, with the median performance for a Cre-line/layer  
706 represented by bar. (b) Pawplot and strip plot summarizing the population sparseness of  
707 responses to natural scenes. (c) Pawplot and strip plot summarizing noise correlations in the  
708 responses to drifting gratings. (d) Pawplot and strip plot summarizing the impact of shuffling on  
709 decoding performance for drifting grating direction.

**Figure 5****a****b****c****d****e****f****g****h**

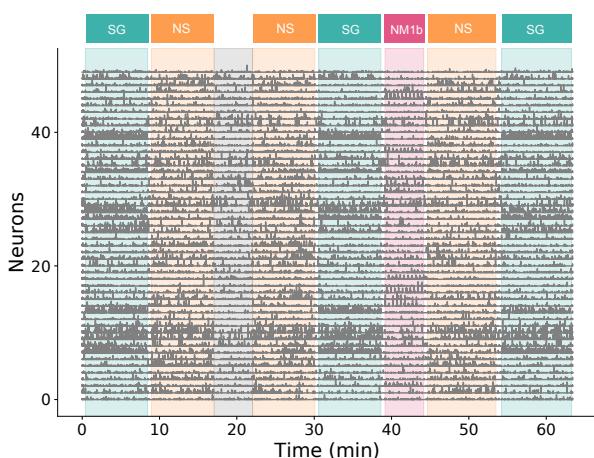
710 **Figure 5: Neural activity is extremely variable and is not accounted for by running**  
711 **behavior.**

712 (a) Pawplot and box plot summarizing the percent of responsive trials that have a significant  
713 response for each neuron's preferred drifting gratings condition. The responsiveness criteria is  
714 that a neuron responded to 25% of the trials, hence the low end is capped at 25%. For box plots  
715 for other cortical areas see **Supplemental Figure 24**. (b) Pawplot and box plot summarizing the  
716 coefficient of variation for each neuron's response to its preferred drifting grating condition. (c)  
717 Two example neurons showing individual trial responses along with mean tuning curve. (d)  
718 Pawplot and box plot summarizing the categorical regression, where  $r$  is the cross-validated  
719 Pearson correlation between model prediction and actual response. Only neurons that are  
720 responsive to drifting gratings using our criteria are included (**Supplemental Figure 15**). (e)  
721 Pawplot and box plot summarizing the correlation of neural activity with running speed. Only  
722 neurons in imaging sessions where the running fraction is between 20 and 80% are included,  
723  $n=30,159$  (**Supplemental Figure 25**). For neurons present in multiple session that met the  
724 running criteria, mean of their running correlation across those sessions is used here. (f) Density  
725 plot of the evoked response to a neuron's preferred drifting grating condition when the mouse is  
726 running (running speed  $> 1$  cm/s) compared to when it is stationary (running speed  $< 1$  cm/s).  
727 Only neurons that are responsive to drifting gratings, and have sufficient number of running and  
728 stationary trials for their preferred condition are included,  $n=20,794$ . (g) Categorical model for  
729 two example neurons (same as in c) in which the running (blue) and stationary (red) trials have  
730 been segregated. (h) Density plot of  $r$  for the categorical regression for drifting gratings using  
731 only the stimulus condition (horizontal axis) and stimulus condition X running state (vertical  
732 axis). Only neurons that are responsive to drifting gratings, and have sufficient number of  
733 running and stationary trials across stimulus conditions, are included.  $n=11,799$ .

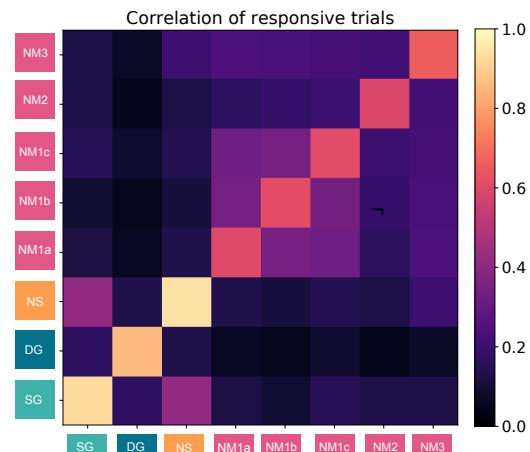
734

**Figure 6**

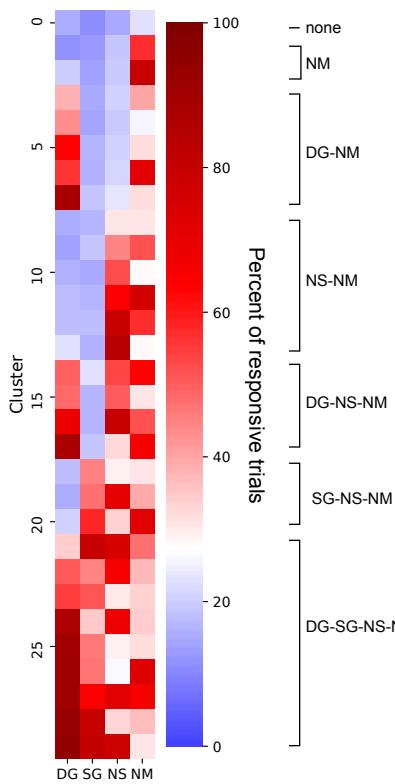
**a**



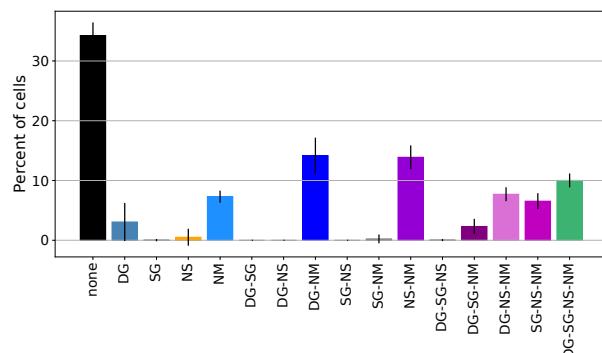
**b**



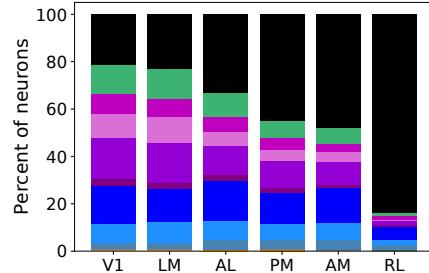
**c**



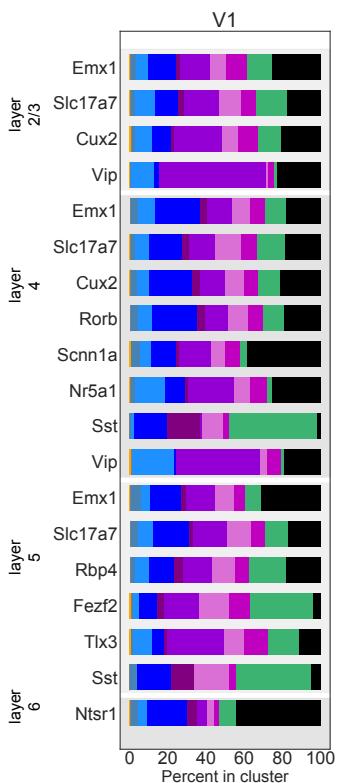
**d**



**e**



**f**



735 **Figure 6: Correlated response variability reveals functional classes of neurons**

736 (a) Responses of 50 neurons during one imaging session (Cux2, layer 2/3 in V1) with stimulus  
737 epochs shaded using stimulus colors from Figure 1f. (b) Heatmap of the correlation of the  
738 percent of responsive trials for neurons' responses to each pair of stimuli. The diagonal is the  
739 mean correlation between bootstrapped samples of the percent responsive trials for the given  
740 stimulus. (c) Mean percent responsive trials for each cluster per stimulus for one example  
741 clustering from the Gaussian mixture model. On the right, classes are identified according to the  
742 response profile of each cluster. (d) Percent of neurons belonging to each class predicted by the  
743 model, mean and st. dev. over 100 repeats. Clustering was performed on 25,958 neurons  
744 imaged in sessions A and B. (e) The percent of cells belonging to each class per cortical area.  
745 Colors correspond to panel d. (f) The percent of cells belong to each class for each transgenic  
746 Cre line within V1. Colors correspond to panel d. For other cortical areas see **Supplemental**

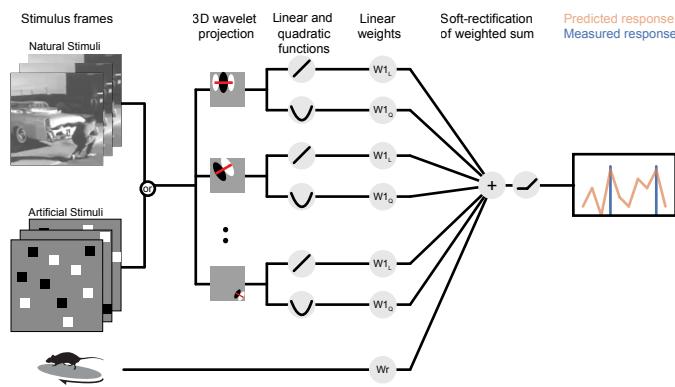
747 **Figure 27.**

748

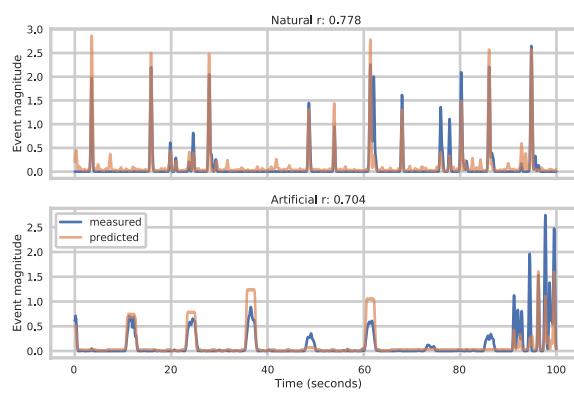
749

**Figure 7**

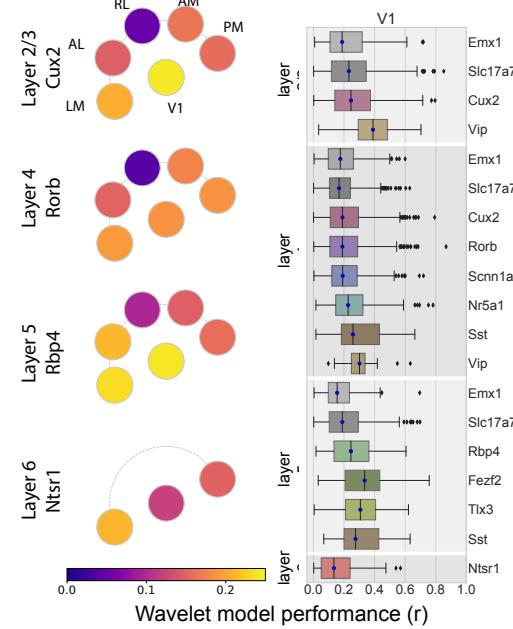
**a**



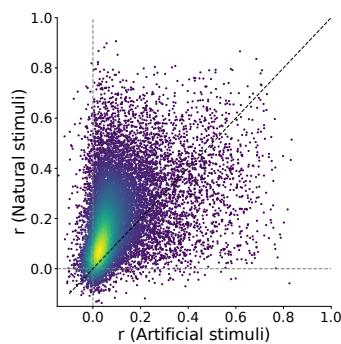
**b**



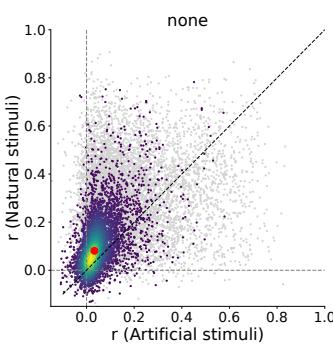
**c**



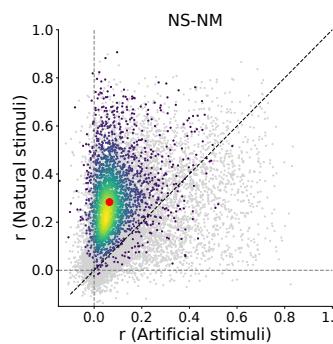
**d**



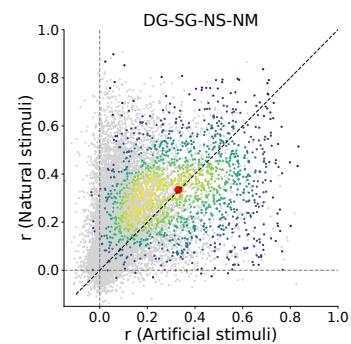
**e**



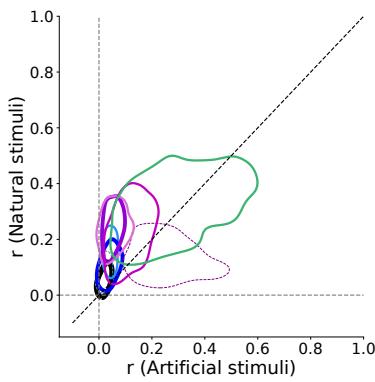
**f**



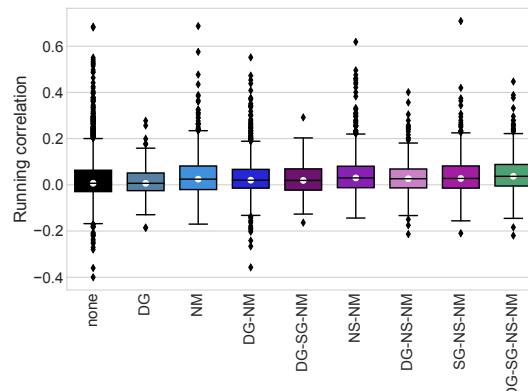
**g**



**h**



**i**



750 **Figure 7: Class labels are validated by model performance.**

751 (a) Schematic for the wavelet models. (b) Example model performance for one neuron for both  
752 natural (top) and artificial (bottom) stimuli. (c) Pawplot and box plot of model performance, r, for  
753 wavelet models trained on natural stimuli. Only neurons imaged in all three sessions are  
754 included, n=15,921. For other cortical areas, see **Supplemental Figure 28**. (d) Density plot  
755 comparing the r values for model trained and tested on natural stimuli to the r values for model  
756 trained and tested on artificial stimuli, n=15,921. (e) Same as d. Only neurons in the “none”  
757 class are shown in the density (n=5,566), all other neurons are in gray. The red dot marks the  
758 median model performance for neurons in this class. (f) Same as e for the “NS-NM” class  
759 (n=2,412). (g) Same as e for the “DG-SG-NS-NM” class (n=1,451). (h) Contours for the density  
760 of model performance, as in e-g, for all classes. The countours mark the boundaries within  
761 which 66% of datapoints lie. Linewidths reflect the number of neurons in each class. Colors are  
762 in i. (i) Box plot of running correlation for each class.

763