Systems/Circuits

Visual Stimulus Detection Correlates with the Consistency of Temporal Sequences within Stereotyped Events of V1 Neuronal Population Activity

Jorrit S. Montijn,[1] Umberto Olcese,[1] and Cyriel M. A. Pennartz[1,2]

[1]Swammerdam Institute for Life Sciences, Center for Neuroscience, Faculty of Science, University of Amsterdam, Amsterdam 1098XH, The Netherlands, and [2]Research Priority Program Brain and Cognition, University of Amsterdam, Amsterdam 1098XH, The Netherlands

Sensory information about the world is translated into rate codes, such that modulations in mean spiking activity of neurons relate to differences in stimulus features. More recently, it has been proposed that also temporal properties of activity, such as assembly formation and sequential population activation, are important for understanding the relation between neuronal activity and behavioral output. These phenomena appear to be robust properties of neural circuits, but their relevance for perceptual judgments, such as the behavioral detection of stimuli, remains to be tested. Studying neuronal activity with two-photon calcium imaging in primary visual cortex of mice performing a go/no-go visual detection task, we found that assemblies (i.e., configurations of neuronal group activity) reliably recur, as defined using Ward-method clustering. However, population activation events with a recurring configuration of core neurons did not appear to serve a particular function in the coding of orientation or the detection of stimuli. Instead, we found that, regardless of whether the population event showed a recurring or nonrecurring configuration of neurons, the sequence of cluster activation was correlated with the detection of stimuli. Moreover, each neuron showed a preferred temporal position of activation within population events, which was robust despite varying neuronal participation. Furthermore, the timing of neuronal activity within such a sequence was more consistent when a stimulus was detected (hits) than when it remained unreported (misses). Our data indicate that neural processing of information related to visual detection behavior depends on the temporal positioning of individual and group-wise cell activity.

Key words: calcium imaging; mouse; population codes; stimulus detection; temporal sequences; vision

---

Significance Statement

Temporally coactive neurons have been hypothesized to form functional assemblies that might subserve different functions in the brain, but many of these proposed functions have not yet been experimentally tested. We used two-photon calcium imaging in V1 of mice performing a stimulus detection task to study the relation of assembly activity to the behavioral detection of visual stimuli. We found that the presence of recurring assemblies per se was not correlated with behavior, and these assemblies did not appear to serve a function in the coding of stimulus orientation. Instead, we found that activity in V1 is characterized by population events of varying membership, within which the consistency of the temporal sequence of neuronal activation is correlated with stimulus detection.

Introduction

Neurons in sensory cortex represent the outside world through modulations in firing rate that are dependent on the presence of particular features, such as the orientation of a visually presented grating (Hubel and Wiesel, 1959). Stimulus representations by single neurons have been studied over the past decades, but more recent studies have shown that the interplay between neurons is also of critical importance for understanding how sensory input is translated into behavioral output (Petersen et al., 2002; Averbeck et al., 2006; Beck et al., 2008; Gutnisky and Dragoi, 2008; Cafaro and Rieke, 2010; Churchland et al., 2010; Benucci et al., 2013; Adibi et al., 2014; Pennartz, 2015; Schölvinck et al., 2015; Sadacca et al., 2016). The behavioral detection of near-threshold

stimuli depends on a population's average activity but is more strongly related to neuronal population phenomena, such as instantaneous correlations in neuronal activity (Romo et al., 2003; Luna et al., 2005; Cohen and Maunsell 2009, 2011; Mitchell et al., 2009; Carnevale et al., 2013; T. W. Chen et al., 2013; Safaai et al., 2013; Buran et al., 2014; Montijn et al., 2015). Although the combination of these rate-based phenomena can predict an animal's choice probability to a large degree, also considering temporal interdependencies may lead to a deeper understanding of neural coding and its relation to stimulus detection.

It has been proposed that the temporal co-occurrence of spiking activity in sensory cortex plays a role in information processing, and especially in boosting the probability that representations of stimuli with low salience are transmitted to postsynaptic neuronal populations (Hebb, 1949; König et al., 1995; Harris, 2005; Pennartz, 2015; Zandvakili and Kohn, 2015). This has given rise to several hypotheses, of which we will focus on two in particular. First, it has been proposed that in mouse primary visual cortex (V1) groups of neurons are coactive in recurring configurations (Miller et al., 2014; Carrillo-Reid et al., 2015). These "ensembles" are hypothesized to show similar configurations during spontaneous activity and in response to visual stimulation. On top of this core ensemble configuration, additional participating neurons can signal stimulus aspects, such as orientation and salience. The second hypothesis entails that activity in sensory cortex consists of discrete "packets" of neuronal firing, with a relatively rigid temporal structure, where each neuron has a preferred position relative to the packet's temporal center of mass (Luczak et al., 2007, 2013, 2015). This packet-like structure would allow the weak effects of individual spikes to efficiently summate and effect a strong postsynaptic response. Evidence for these two proposals has been raised in passively stimulated sensory cortex, but their relevance for behaviorally reported visual detection of stimuli has not been established.

Therefore, we investigated whether assembly formation and consistent temporal ordering are behaviorally relevant in a stimulus detection paradigm. We used two-photon calcium imaging of populations of ~100 neurons in V1 L2/L3 of mice performing a detection task. Superficial layers are easy to access with calcium imaging, show ensemble activity in V1 (Miller et al., 2014), and display neural correlates of stimulus detection (Ito and Gilbert, 1999; van der Togt et al., 2006; Glickfeld et al., 2013; Montijn et al., 2015). We show that (1) recurring assemblies are present in L2/L3 of mouse V1 but (2) do not encode stimulus orientation, and their presence *per se* is not correlated with stimulus detection. Instead, (3) V1 responses consist of stereotyped events with varying neuronal membership, and the sequential activation of these events is more similar during the detection of visual stimuli. Moreover, (4) within population events (PEs), the precision of temporal positioning of neuronal responses is correlated with visual detection. We conclude that the sequential structure of neuronal activation is more behaviorally relevant than the mere activation of these assemblies.

Materials and Methods

Note on present study. The current study presents analyses of datasets that have been previously described (Montijn et al., 2015). Although the data used for these studies are the same, the questions investigated (and analyses performed) are very different. For more information on the methods and data used, and for further analyses, see Montijn et al. (2015).

Animals and surgery. All experiments were approved by the animal ethics committee of the University of Amsterdam. Eight male C57BL/6J

mice (Harlan, 128–164 d old at the day of calcium imaging) were used in experiments. Before the imaging sessions, a head-bar implant was surgically fixed to animals, and the cranial window was sealed using a layer of glue, silicon elastomer on top, and a small cover glass (6 mm) to fix the silicone. Animals were then trained to perform a head-fixed visual go/no-go detection task (see Behavioral training). On the day of the calcium imaging recording, intrinsic signal imaging was performed to identify the retinotopic region of the primary visual cortex (V1) responsive to the employed visual stimulus. A small craniotomy (1.5–2.0 mm) was then performed in the identified region. Multicell bolus loading with Oregon Green BAPTA-1 AM (OGB) was used to be able to detect calcium transients. Sulforhodamine-101 was used to label astrocytes (Stosiek et al., 2003; Nimmerjahn et al., 2004).

Behavioral training. Animals were trained daily (~45 min/d) over the course of 10–12 weeks. Mice were water-deprived for 6 h preceding training and otherwise had *ad libitum* access to water. Animal weight was consistently monitored and never dropped below 90% of the *ad libitum* growth curve. Training was performed in dark, sound-attenuated chambers and took place in the active (dark) cycle. The first 5 training days were used to condition a licking response following visual stimulation, by pairing passive stimulation with reward delivery (~9 ml of water with 15% sucrose containing 1% vanilla extract). After this phase, full-contrast visual stimuli (for details, see Craniotomy and dye injection) were presented until mice made a licking response. Licking was monitored via a custom-made infrared lick detector. Upon detection of a licking response, visual stimuli terminated and reward was available for 5 s. This phase lasted for at most 5 d. The next phase consisted of training animals on a simplified version of the final task: stimulus presentation time was reduced to at most 5 s and a trial could start only after a random interval of 1–3 s without any licking response. Reward size was gradually reduced to ~3 μl per trial. Once animals could perform at least 80 trials within 45 min, task complexity was increased. First, we introduced 0% contrast probe trials to test for false alarm (FA) responses and calculate whether performance was statistically above chance. Also, the intertrial interval was increased to a 6–8 s random duration. Once a sufficient ratio of hit/miss trials was achieved, we increased the intertrial interval to 10–12 s and introduced mild air puffs as a negative reinforcer, to be applied whenever mice would lick outside the stimulus presentation or reward delivery period. To avoid explicitly teaching animals to make fast behavioral responses, mice were always allowed to respond within 3 s after stimulus onset. If the mice failed to respond, they received a time-out period. Finally, if mice performed consistently and significantly above chance during the previous phase ($n = 12$ of 21 animals, duration of the phase: 8–10 weeks), they were trained for 2 more weeks on the microscope setup, especially to habituate them to the auditory noise of the two-photon calcium imaging setup.

Craniotomy and dye injection. On the day of the imaging experiment, buprenorphine (0.05 mg/kg) was injected subcutaneously before induction of isoflurane anesthesia (4.0% induction, 0.8% maintenance during intrinsic signal imaging, 1.5%–2.5% maintenance during surgical procedures). After induction, the animal was placed in a custom-built headbar holder. Intrinsic signal imaging was performed to localize the precise location of the visual stimulus' receptive field location in V1, after removing the cover glass, silicon elastomer, and layer of glue covering the skull in the cranial window. We then performed a small (1.5–2 mm) craniotomy above the localized portion of V1. The dura was kept wet with ACSF (NaCl 125 mM, KCl 5.0 mM, MgSO$_4$ · 7 H$_2$O 2.0 mM, NaH$_2$PO$_4$ 2.0 mM, CaCl$_2$ · 2 H$_2$O 2.5 mM, glucose 10 mM) buffered with HEPES (10 mM, adjusted to pH 7.4). Multicell bolus loading with OGB and sulforhodamine 101 (SR101) was then performed 230–270 μm below the dura as previously described (Stosiek et al., 2003; Montijn et al., 2015). After injection of the dyes, the dura was covered with agarose (1.5% in ACSF) and sealed with a circular cover glass, which was fixed to the skull by cyanoacrylate glue. Animals were allowed to recover for at least 90 min before starting the imaging session. Of the 12 mice that learned the task, we excluded 2 animals due to insufficient imaging quality.

Visual stimulation. Visual stimuli were delivered via a 15-inch TFT screen (60 Hz refresh rate) positioned 16 cm from the mouse's eye. Stimuli were generated in MATLAB (The MathWorks) using the Psych-

Toolbox extension (Brainard, 1997; Pelli, 1997). Stimuli consisted of sequences of square-wave drifting gratings along eight different directions; stimuli were presented to a single eye in randomized order. During calcium imaging, visual stimuli were presented for maximally 3 s and were terminated at the onset of the first licking response. Stimuli were alternated by an intertrial interval of variable, random duration (between 10 and 12 s) during which an isoluminant gray screen was presented. Drifting gratings (diameter: 60 retinal degrees, spatial frequency: 0.05 cycles/degree, temporal frequency: 1 Hz) were presented within a circular cosine-ramped window to prevent edge effects at the border of the circular window. A field-programmable gate array (FPGA, OpalKelly XEM6001) was used to synchronize microscope, behavioral setup, and visual stimulus presentation computer.

Calcium imaging recordings and final task parameters. We performed dual-channel two-photon imaging recordings (filtered between 500 and 550 nm for OGB fluorescence and between 565 and 605 nm for SR101 fluorescence; see Fig. 1B) (Montijn et al., 2014) using a modified Leica SP5 confocal system with a Spectra-Physics Mai-Tai HP laser (set at a wavelength of 810 nm to simultaneously excite OGB and SR101). Images were acquired from cortical layer 2/3 (140–170 μm depth from the pia mater) at a sampling frequency of 25.4 Hz via a resonant scanner and had a 512 $\times$ 512 pixel frame size. During data acquisition, mice performed the behavioral task described above. Stimulus parameters were equal to those described above. The contrast of the drifting grating was varied across the 0%–100% range (specific values: 0%, 0.5%, 2%, 8%, 32%, and 100%) to elicit a wide range of hit/miss ratios. Responses to all contrasts, except 0% contrast probe trials, were rewarded. To avoid overtraining and associated habitual or automated responses (Balleine and Dickinson, 1998), we did not overtly train animals to reach a very high detection performance (i.e., high hit rates and low miss rates). The order of presentation of a complete set of visual stimuli (48 trials: 6 contrasts times 8 directions) was randomized independently in each repetition block. We tested for statistically significant stimulus detection performance by calculating the binomial 2.5th-97.5th percentile intervals (95% CI) of response proportion to the two probe trial types (100% and 0% contrast stimuli) using the Clopper-Pearson (CP) method. Of the 10 animals used in the imaging experiments, one was rejected because of insufficient discriminability between the two types of probe trials (i.e., overlapping CIs). Another was rejected because of high variability in neural responses resulting from brain movement. Eight animals thus contributed data for all subsequent analyses. Each stimulus type (unique orientation $\times$ contrast) was repeated between 6 and 16 times. For most analyses, we used spontaneous as well as stimulus-driven neuronal activity (see Figs. 1234-5, 8A–D). For all trial-based analyses, we aimed to prevent confounds from having higher signal-to-noise ratios for miss than hit trials. Therefore, for each trial, we always analyzed the neuronal response during the 2 s period after stimulus onset (see Figs. 7, 8 E,F, 9A–D), except for additional control analyses (i.e., see Fig. 9E–K), and when analyzing cluster sequence similarity (see Fig. 6). For analyses as a function of stimulus contrast, we averaged across all orientations (n = 4), so that each contrast was presented 24–64 times (see Fig. 7). During most recordings, the mouse's eye was monitored with a near-infrared light-sensitive camera (JAI CV-A50IR-C Monochrome 1/2 inch IT CCD Camera) with a large-aperture narrow-field lens (50 mm EFL, f/2.8). Images were acquired at 25 Hz, and pupil tracking was performed offline using custom-written MATLAB scripts. When trials where the mice were blinking or making saccades at any point during stimulus presentation (10.2 $\pm$ 4.6% of trials, mean $\pm$ SD) were removed, this did not qualitatively affect the results (see also Montijn et al., 2015).

Data preprocessing. Small x-y drifts were corrected offline with an image registration algorithm (Guizar-Sicairos et al., 2008). Regions of interest (neurons, astrocytes, and blood vessels) were determined using a custom-made semiautomated MATLAB software separately on each repetition block (available at http://github.com/JorritMontijn/Preprocessing_Toolbox). dF/F_0 values were computed for each region of interest as previously described (Montijn et al., 2014). For each image frame i, a single dF_i/F_{0i} value was obtained for each neuron (regions of interest) by calculating the baseline fluorescence F_{0i} (computed as the mean of the lowest 50% fluorescence in a 30 s window around frame i) and the dF_i

value (computed as the difference between the neuron fluorescence in frame i and F_{0i}) (Montijn et al., 2014). We simultaneously recorded on average 92.6 neurons per imaging session (range: 68–130; SD: 19.0). To avoid any bias due to the selective inclusion of neurons, we did not specifically select only orientation-tuned neurons, but we instead analyzed data from all identified neurons. Unless specified otherwise, all analyses were based on across-animal meta statistics using one data point per animal (n = 8). All multiple-comparison t test p values were adjusted by the Benjamini and Hochberg False Discovery Rate (FDR) correction procedure. A value of $p < 0.05$ was considered significant.

Activation event (AE) detection. To be able to analyze temporal properties of neuronal population activity, we detected AEs in the dF/F0 data, thereby reducing the continually varying fluorescence levels to point events. To detect AEs, we used a previously published algorithm that fits exponential decay functions to dF/F0 traces (Greenberg et al., 2008). Detailed information can be found in the original publication, but in short, we performed the following steps per neuron. First, we marked all time points i (i.e., acquisition frames) that could potentially show AEs when they fulfilled all of the following requirements:

$$\text{dF/F0}_i > 6\% \tag{1}$$

$$\text{dF/F0}_i - \text{dF/F0}_{i-1} > 2\% \tag{2}$$

$$\text{dF/F0}_i - \text{dF/F0}_{i-2} > 0.8\% \tag{3}$$

$$\text{dF/F0}_{i+1} - \text{dF/F0}_{i-1} > -3\% \tag{4}$$

Finally, we fitted exponential decay functions to the dF/F0 traces that started at the detected onset frames for potential AEs. To avoid overfitting fluorescence noise, we removed all events that showed an insufficient amplitude (maximal dF/F0 < 10%). An AE was detected whenever:

$$\boldsymbol{F} * \boldsymbol{Y} / \|Y\| > 10\% \tag{5}$$

Here, $\boldsymbol{F}$ is the vector of single-cell dF/F0 responses at and after frame i, $\boldsymbol{Y}$ is an exponential kernel, and $*$ is the convolution operator as follows:

$$F = [\text{dF/F0}_i, \text{dF/F0}_{i+1}, \dots, dF/F0_{i+17}] \tag{6}$$

$$Y = [e^{-0 \cdot dt/\tau}, e^{-1 \cdot dt/\tau}, \dots, e^{-17 \cdot dt/\tau}] \tag{7}$$

The variables for the exponential kernel were $dt = 39.4$ ms (frame acquisition duration at 25.4 Hz) and $\tau = 500$ ms (Greenberg et al., 2008). All frames fulfilling Equation 5 were marked as time points where AEs were detected.

Decoding of stimulus orientation. To validate the quality of AE detection, we performed an orientation decoding procedure on the dF/F0 data as well as on the AE data (Greenberg et al., 2008) for a verification of the AE detection algorithm, including simultaneous electrophysiological spike recordings. A similar version of the leave-one-out cross-validated template matching decoding procedure we used has been described in detail previously (Montijn et al., 2014). In short, it calculates the Euclidian distance d in SDs of neuronal population response from any trial t's population activity $\boldsymbol{r}_t$ to the templates $\boldsymbol{r}_\theta$ obtained for the different stimulus orientations θ, and then selects as decoded orientation the class with the smallest distance as follows:

$$d(\boldsymbol{r}_\theta, \boldsymbol{r}_t) = \frac{\boldsymbol{r}_t - \boldsymbol{\mu}_\theta}{\boldsymbol{\sigma}_\theta} \tag{8}$$

Here, each element in $\boldsymbol{r}_\theta$, $\boldsymbol{r}_t$, $\boldsymbol{\mu}_\theta$, and $\boldsymbol{\sigma}_\theta$ corresponds to a single neuron, and the whole vector represents the entire neuronal population. Both the subtraction and division operations are performed element-wise. $\|.\|$ represents the vector norm. The template response $\boldsymbol{r}_\theta$ for stimulus orientation θ consists of the mean $\boldsymbol{\mu}_\theta$ and SD $\boldsymbol{\sigma}_\theta$ of neural activity (in number of AEs) for all neurons. For a given $\boldsymbol{r}_t$ of trial t, the decoded orientation is therefore the orientation θ that yields the smallest distance d between the template response ($\boldsymbol{r}_\theta$) and the actual neural response during that trial ($\boldsymbol{r}_t$). For the calculation of template responses, trials of all non-zero contrasts (0.5% up to 100%) and behavioral responses (hits and misses) are lumped together, explaining the relatively modest absolute accuracy (see

Fig. 2B). Splitting datasets per stimulus contrast yielded insufficient repetitions per orientation to reliably estimate template responses (results not shown). Although previous studies have shown a tight correspondence between detected AEs and single spikes (e.g., Greenberg et al., 2008), it should be stressed that we do not claim such a strong one-to-one relationship here. The aim of the procedure used here is to transform dF/F0 data into point events, and a precise estimate of the number of spikes occurring during a given calcium imaging frame is therefore not required to interpret our results.

PEs and clustering. To study ensemble-related neuronal activation patterns, we first grouped all AEs into PEs and isolated events (IEs). PEs were defined as epochs of at least 78.8 ms (two acquisition frames) with no enforced upper limit, where periods of nonactivity were never longer than 39.4 ms (i.e., the interval between subsequent frame acquisitions). Therefore, in every calcium imaging frame during these epochs, at least one neuron had to be active. If during such epochs less than three cells were participating, the constituent AEs were marked as IEs instead. All remaining epochs were defined as PEs (Figs. 2E, 3A).

To further specify different types of PEs, we detected groups of neurons that were repeatedly active in a similar configuration, so we constructed a hierarchical dendrogram of all PEs based on their participating neurons using Ward's method clustering (Fig. 3B,C) (Ward, 1963). To detect the optimal number of clusters, we calculated the mean silhouette distance between clusters based on the Ward dendrogram, for a number of clusters ranging from 2 up to 100 (Fig. 3D) (Rousseeuw, 1987). We next smoothed the resulting silhouette distance curve with a Gaussian kernel of SD 1.0 and defined the optimal number of clusters as the peak (local maximum) silhouette distance with the largest troughs on either side.

Core neurons and definition of recurring clusters (i.e., assemblies). Using the above method to define an optimal number of clusters, we reclustered all PEs using Ward's method and checked for each resulting cluster whether (1) it was not the rest group, and (2) it had more than one core neuron member. The rest group was defined as the cluster with the lowest within-cluster correlation of neuronal members participating in its constituent PEs (Fig. 3E). The rest group's PEs were marked as nonrecurring PEs (NPEs). A neuron was defined as being a core neuron of a cluster when it participated in more of the cluster's occurrences (i.e., its PEs) than would be expected by chance (Fig. 3F). For each cluster, the chance level participation was calculated with 1000 iterations of random resampling of PEs: each iteration, as many PEs were randomly selected as there were instances of the cluster's occurrence. We calculated the 99th percentile of a neuron's participation in a cluster across these 1000 iterations (Fig. 3F, right, black error bars). A neuron was defined as being a core neuron when its true participation rate in a cluster was higher than the upper 99th percentile of the random reshuffling (Fig. 3F, green bars). PEs were defined as recurring when at least more than one neuron was a core neuron of the PE's cluster, and the within-cluster correlation was higher than the rest group (i.e., cluster with lowest within-group correlation) (Fig. 3G,H). The PEs of these clusters were marked as recurring PEs (RPEs), and the clusters are referred to as "recurring clusters" (or "assemblies"). If only a single neuron was a core neuron, the cluster's constituent PEs were marked as single core-neuron PEs (SPEs). Because of the ambiguity of whether this last group of clusters represented a recurring configuration or not, they were not further analyzed when comparing recurring (RPEs) and NPEs (see, e.g., Figs. 4, 5, 7). When all clusters or PEs were analyzed regardless of recurrence characteristics, also SPEs were included (see, e.g., Figs. 6, 8, 9).

As the number of neurons is an important factor that could influence the results, we reperformed several analyses with the requirement of at least four (instead of three) neurons being active for an event to qualify as a PE. Increasing the required cell density to at least four neurons per PE reduces the number of PEs by 37.8%, a significant reduction in the size of the datasets. Because of this reduction, we found that the effect sizes were often reduced, but the direction of the effects was maintained without exception. For example, the cluster sequence similarity (the same analysis as presented in Fig. 6) was still significantly correlated with stimulus detection (hit vs miss, $p = 0.028$), and the reduction in latency variability within PEs showed a trend in the same direction as before (lower vari-

Table 1. t test p values for anatomical grouping, mean OSI, and mean difference in preferred orientation (dPO) as a function of different numbers of core neurons required for inclusion of the cluster[a]

No. of core neurons required for inclusion	Mean distance between core neurons; p value	Mean OSI of clusters; p value	Mean dPO of core neurons per cluster; p value	No. of recurring clusters
2 (original)	0.004**	0.309; ns	0.246; ns	71
3	0.019*	0.402; ns	0.852; ns	53
4	0.022*	0.418; ns	0.709; ns	36
5	0.050; ns	0.258; ns	0.428; ns	31
6	0.046*	0.123; ns	0.616; ns	23
7	0.028*	0.107; ns	0.433; ns	19
8	0.057; ns	0.352; ns	0.936; ns	16
9	0.183; ns	0.066; ns	0.756; ns	11
10	0.183; ns	0.066; ns	0.756; ns	11

[a]Only anatomical grouping was significant for a range of minimum core members; OSI and dPO were not.
**, $p < 0.005$; *, $p < 0.05$.

ability during hits than misses), but this effect was statistically no longer significant ($p = 0.063$). Together with our other results that show robustness to variation in most of the parameters of how the effects are quantified (e.g., time window, see Fig. 9; number of required neurons as shown in Table 1) and the consistency of temporal sequences when calculated differently (compare Fig. 8F with Fig. 9B,E), we conclude that the results we present in the manuscript are robust, albeit dependent on a particular cell density. This can be likely ascribed to a reduction in statistical power with higher minimum cell density.

Spatial grouping and orientation selectivity analysis. We hypothesized that temporally coactive neurons might be more anatomically grouped. To study this, we calculated the mean distance between pairs of neurons within real recurring clusters, and compared this with what would be expected if neurons were participating in the same recurring cluster randomly across the field of view. We ran 1000 iterations per cluster c where we randomly picked a number of neurons n from the pool of neurons that were a core neuron of any cluster, where n is the number of core neurons of cluster c. We calculated the mean pairwise anatomical distance between all randomly selected neurons per iteration. This shuffling procedure therefore yielded a vector $d_{shuffled}$ of 1000 mean pairwise distances that we used to calculate the mean and SD of a randomly assigned cluster of neurons. We transformed the actual mean pairwise distance d into a z-score normalized distance d_z based on the mean $\mu_{shuffled}$ and SD $\sigma_{shuffled}$ from the random distribution for each cluster c as follows:

$$d_z = \frac{d - \mu_{shuffled}}{\sigma_{shuffled}} \quad (9)$$

The index d_z therefore indicates the anatomical grouping of the core neurons of a cluster of coactive neurons relative to what would be expected by chance. A negative value indicates more anatomical grouping (i.e., the core neurons are closer together), and a positive value indicates less grouping (see Fig. 4B).

Similar to the analysis described above, where we studied the anatomical grouping of clusters, we also investigated the orientation tuning of these clusters of coactive neurons. We calculated the orientation selectivity index (OSI) of clusters and of neurons as follows:

$$OSI = (A_{pref} - A_{orth})/(A_{pref} + A_{orth}) \quad (10)$$

Here, A_{pref} is the mean number of AEs during trials of the preferred stimulus orientation, and A_{orth} of the orientation orthogonal to the preferred orientation (separated by 90 degrees). In addition to using OSI, we also investigated the common orientation tuning within recurring clusters by calculating whether core neurons shared a more similar preferred orientation than expected by chance. We performed both analyses with a shuffle-control as described above for anatomical grouping. Instead of taking the pairwise anatomical distance between somata, we took the recurring cluster's OSI (OSI_c), or the pairwise angular difference in preferred orientation of the core neurons (θ_c). All other steps were identical (see above), which yielded for each cluster an index OSI_z, a normalized

metric of OSI, or θ_z, a normalized metric of the mean difference in preferred orientation relative to chance as follows:

$$OSI_z = \frac{OSI_c - \mu_{shuffled}}{\sigma_{shuffled}} \qquad (11)$$

$$\theta_z = \frac{\theta_c - \mu_{shuffled}}{\sigma_{shuffled}} \qquad (12)$$

Positive values of OSI_z and θ_z indicate that the recurring cluster is more orientation tuned and that the cluster's core neurons have a more dissimilar preferred orientation, respectively (see Fig. 4C,D). To calculate the presence of clusters during hit and miss trials shown in Figure 6F, we used the same z-scoring procedure, except that we calculated the ratio of trials in which a cluster was present (separately for hit and miss trials), relative to the expected distribution of trial presence ratios.

Temporal interactions of occurrence probability. We investigated whether the occurrence of a particular cluster's PE would influence the occurrence probability of PEs from that same cluster as well as from other clusters (see Fig. 5). To do so, we aligned for each PE belonging to a particular cluster all PEs from that same cluster over a window of −60 to 60 s around the time of the PE's occurrence. We summed the number of PEs at each time bin (size: 39.4 ms; 1/25.4 Hz) and divided by the overall probability of a PE from that cluster occurring at a single time bin. This way, we obtained a normalized occurrence probability for each cluster as, for example, shown in Figure 5A. We repeated this procedure for PEs from all other clusters (see Fig. 5A, blue trace). To quantify statistical significance, we calculated a FDR-corrected p value using a one-sample t test versus 100% chance level at each time bin, pooling the occurrence probabilities over all clusters from all animals (Fig. 5D). A similar procedure was performed as function of time after stimulus onset, by aligning the occurrence of PEs to stimulus onset, rather than the time of occurrence of the cluster's PEs (Fig. 5B,C,E,F). For these stimulus-related analyses, a bin was defined as being significantly above chance when the mean occurrence probability across recurring clusters was higher than the mean plus 2 SDs (i.e., $\alpha = 0.05$). The results were similar for a threshold of 3 SDs (i.e., $\alpha = 0.003$), and when the duration of PEs was artificially set to a single frame duration (39.4 ms).

Cluster sequence similarity. We analyzed whether clusters occurred in specific sequences during certain trial types, by calculating a cluster sequence similarity for same orientation trials versus different orientation trials, as well as for miss trials versus hit trials. We defined this similarity as follows. First, we considered only the order of clusters by itself: that is, a trial where cluster 1 was present for 1000 ms, followed by 100 ms of no PEs (denoted by "0") and cluster 2 for 80 ms, has a similarity of 1 with a trial where cluster 1 was present for 120 ms, followed by 800 ms of silence and cluster 2 for 800 ms. Mathematically, this entails that when $v = [1, 1, 0, 0, 2]$ and $w = [1, 0, 2, 2, 0]$, then $v \Leftrightarrow w$, as both vectors reduce to $[1, 2]$ after the removal of silent periods (i.e., "0") and durations are ignored. For reduced vectors that correspond to the sequential presence of clusters during a trial, the cluster sequence similarity Φ between two trials t and u with their corresponding vectors v and w, is defined as follows:

$$\Phi_{t,u} = \frac{\sum_{i=1}^{n} (\delta_{v_i w_i})}{n} \qquad (13)$$

Here, δ corresponds to the Kronecker δ: that is, it is 1 when $v_i = w_i$, and 0 when $v_i \neq w_i$. n represents the number of cluster occurrences in whichever vector of v and w has the least number of elements (i being a single element). Φ therefore has a maximum value of 1, when v and w have the same elements at the same locations, and Φ has a minimum value of 0, when the elements in v do not have the same value at the same position in w. For example, a trial with a reduced cluster sequence of $[4, 3, 2, 1]$ will have a similarity of 0 with a trial where the reduced cluster sequence was $[1, 2, 3, 4]$.

Dependence of recurring cluster properties on contrast. To ensure identical sampling durations per trial, our standard procedure was to take all PEs from stimulus onset until 2 s after onset for all stimulus-related analyses that were based on PEs. Per contrast, we split all trials into hits

and misses (FAs) and correct rejections (CRs) for 0% contrast, respectively. For each RPE, we calculated its size (number of participating neurons), duration (number of consecutive frames where at least one neuron was active), and number of AEs per second. Per recorded population of neurons, we averaged these three properties, as well as the number of occurrences, across all trials of a given type (response type × contrast). This yielded two contrast-dependent curves (hit/FA and miss/CR) for each property. To compare these curves to a baseline level, we also calculated these properties during the 2 s window preceding stimulus onset. The graphs in Figure 7 show these curves as the mean ± SEM across animals.

Temporal sequence analysis. We hypothesized that stimulus detection might be correlated with the sequential pattern of neuronal activation within a PE. To study this phenomenon, we based our analysis on a previously published procedure that has been applied to somatosensory and auditory cortex (Luczak et al., 2007, 2013). Because each neuron n could be active multiple times per PE p, we first calculated the neuron's average activation time $\bar{t}_{p,n}$ over all AEs during that PE. Next, for each neuron in each PE, we calculated the neuron's latency in milliseconds L to the center of mass μ_p (i.e., mean) of the PE's constituent AEs (see Fig. 8A). To better compare the ordering across PEs with different durations, we normalized the latency by the SD σ_p of AEs of all neurons (all lumped) within that PE as follows:

$$L_{p,n} = \frac{\bar{t}_{p,n} - \mu_p}{\sigma_p} \qquad (14)$$

The matrix L therefore contains the normalized latencies of all neurons and PEs. We asked whether temporal locations of neuronal activity in PEs were consistent across clusters. We investigated this by calculating the mean latency $\bar{L}_n$ for each neuron across all PEs and compared this value with the mean latency $\bar{l}_{c,n}$ per neuron within all PEs that were grouped within each cluster c. The consistency values r plotted in Figure 8B are Pearson correlations of each cluster's neuronal latency vector $\bar{l}_c$ (each element being a single neuron) with the overall neuronal latency vector $\bar{L}$ calculated across all PEs as follows:

$$r = \text{corr}(\bar{L}, \bar{l}_c) \qquad (15)$$

We assessed the statistical significance of these values by comparing them with consistency values obtained from a random shuffling procedure we reiterated 1000 times. Each iteration, we randomly shuffled the times of occurrence of all AEs from all neurons and reassigned these shuffled AE times to all neurons. This way, each neuron keeps the same number of AEs, and the mean population activity is not altered, as all AEs occur at the same points in time. However, the neuronal identities that belong to the AEs are now completely random. Therefore, the temporal structure of the population response is preserved, but the temporal structure of AEs per neuron is destroyed. For each iteration, we recalculated the temporal sequence consistencies (Eq. 15), and the shuffled values plotted in Figure 8C show the mean ± 95% CI. The boxplots in Figure 8D show the distribution of mean consistency values across animals ($n = 8$).

To study whether temporal sequence consistency was correlated with stimulus detection, we selected from the latency matrix L the subset of PEs that occurred during either hit trials or during miss trials (see Fig. 8E). For each individual PE, we calculated the Pearson correlation of latencies with the overall latency calculated across all PEs, similar to Equation 15, except for $\bar{l}$ being a vector of latencies during a single PE. This yielded a consistency value (Pearson's r) for every PE that occurred during a hit or miss trial (for two example PEs during miss trials and two example PEs during hit trials, see Fig. 8E). The mean sequence consistency was then statistically quantified by comparing the means across animals between hits and misses (see Fig. 8F).

The difference in single-trial sequence consistency between hits and misses was significant. However, considering the importance of this difference, we additionally examined this result using an alternative approach. Again, we selected the same subset of PEs during either hit trials or during miss trials as above, but now calculated the variability in latency v for each neuron across PEs and averaged across neurons, rather than the mean consistency in latency of each PE as follows:
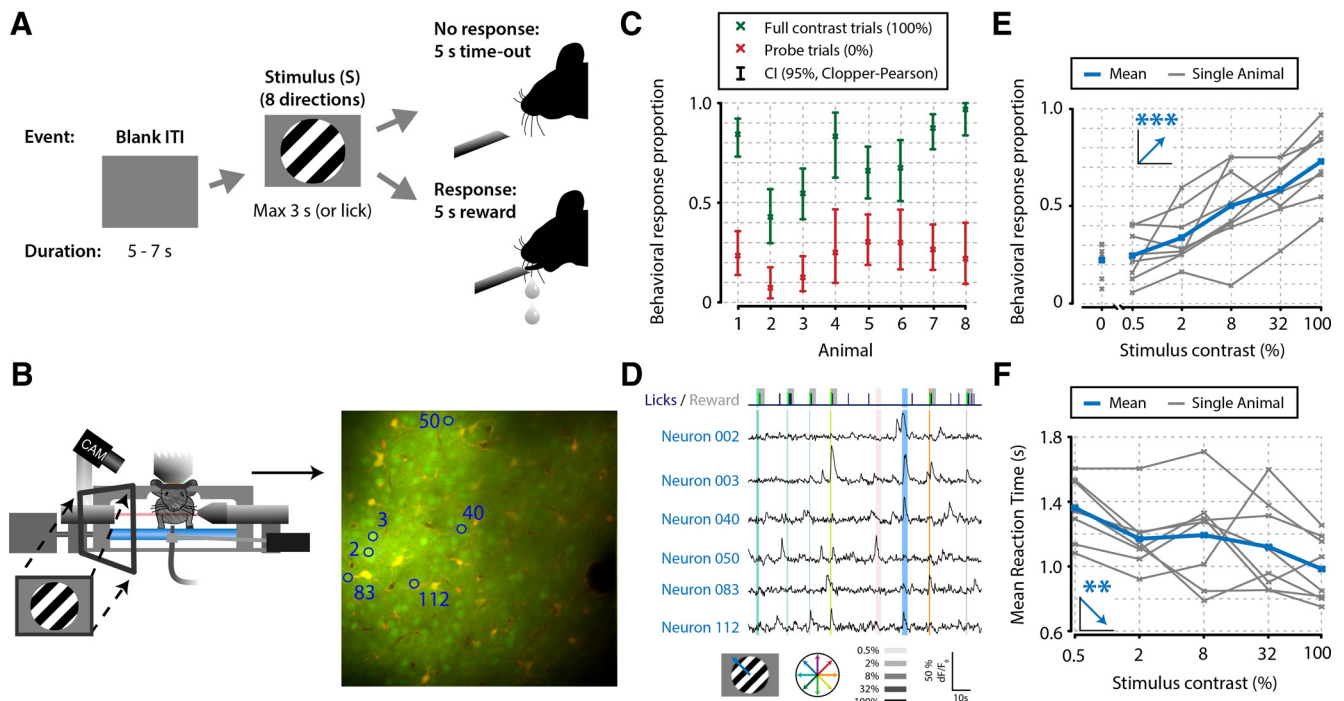
Figure 1. Mice perform a go/no-go stimulus detection task during *in vivo* calcium imaging of V1 populations. ***A***, Task schematic showing a single trial. In each trial, one of a combination of eight different directions and six contrasts was presented. Two of these contrasts were probe trials: a 0% contrast isoluminant gray blank screen was used to infer the FA rate, and a 100% full contrast stimulus was used to infer the lapse rate. The order of presentation of stimulus contrast and direction was randomized per block of 48 trials (6 × 8). At the first licking response during visual presentation, the stimulus was turned off and mice received a reward (sugar water). ***B***, Schematic of the behavioral setup with example recording field of view. The mouse's pupil was monitored, and licking responses and running on a treadmill were recorded. ***C***, The eight animals included in this study showed statistically significant stimulus detection during calcium imaging, quantified by nonoverlapping 2.5–97.5th CP percentile CIs (95% CI) (*p* < 0.05) of behavioral response proportions for 0% (red) and 100% (green) contrast probe trials. ***D***, Example of simultaneously recorded licking responses and traces of neurons labeled in ***B***. Blue or green ticks and gray shaded areas in the top row represent licking and reward presentation, respectively. Vertical colored bars represent stimulus presentations; width, color, and saturation represent duration, orientation, and contrast respectively. ***E***, ***F***, Mice showed higher behavioral response proportions (linear regression analysis, *******p* = 1.07 × 10^{-7}; ***E***) and faster reaction times (********p* = 0.004; ***F***) with higher stimulus contrasts. Figure elements adapted from Montijn et al. (2015).

$$v = \frac{\sum_{n=1}^{N} s_n}{N} \qquad (16)$$

Here, s_n is the variability (i.e., SD) in latency across PEs for neuron n, and N is the number of neurons. For this calculation, we discarded all neurons that participated in less than five PEs during hits or during misses (20.12 ± 5.67% of neurons were discarded, mean ± SEM across animals). The variability in latency values in Figure 8*B* is shown as mean ± SEM across animals.

We also compared the interaction between stimulus detection, sequence consistency, and orientation decoding accuracy. We split all trials into hits and misses, or split all trials into the highest and lowest 50% sequence consistency trials and compared the decoding accuracy between these two groups for both ways of splitting the data (see Fig. 9*C*,*D*; for more information on the decoding procedure, see Eq. 8).

Results

We investigated which temporal characteristics of neuronal population activity in mouse primary visual cortex (V1) correlate with the detection of visual stimuli (Fig. 1). To this end, we recorded two-photon calcium imaging data of L2/L3 neurons (range: 68–130; mean ± SD: 92.6 ± 19.0 per mouse) in V1 of mice that were trained to perform a go/no-go stimulus detection task (Fig. 1*A*,*B*,*D*). During imaging, animals were awake, head-fixed and indicated by licking whether a square-wave drifting grating was presented. Stimulus duration was delimited by the onset of the first licking response. If licking did not occur within 3.0 s of stimulus onset, the trial was scored as "no response"; therefore, no licks occurred during presentation of the stimulus. To acquire a sufficient range of hit/miss ratios, we presented test stimuli with different luminance contrasts: 0.5%, 2%, 8%, and 32%. These test trials were interleaved with 0% no-contrast and 100% full-contrast probe trials to estimate the animals' ratio of FAs and omissions. For all analyses, we discarded trials where animals responded within 150 ms after stimulus onset (0.3%–3.5% of trials per animal) because such fast responses may be ascribed to spontaneous licking.

To quantify behavioral performance during execution of the task, we calculated the 2.5–97.5th percentile intervals (henceforth, 95% CIs) of response proportions to the two types of probe trials: no-contrast and full-contrast stimuli. All eight animals showed a significantly above-chance visual detection of square-wave drifting gratings during the acquisition of neural data (Fig. 1*C*) (nonoverlapping CP 95% CIs). Moreover, behavioral response proportions increased with higher stimulus contrasts (Fig. 1*E*) (group-level linear regression analysis, *p* = 1.07 × 10^{-7}) and mean reaction times decreased (Fig. 1*F*) (*p* = 0.004) (Montijn et al., 2015).

Detection of AEs in calcium data

To study temporal properties of neuronal population activity, we transformed our calcium imaging data of single cells from continuous dF/F0 signals to point events of neuronal activation (AEs). As the fluorescent somatic Ca²⁺ signal is strongly correlated with action potential firing, these point events can be allocated to acquisition frames where a neuron was likely firing one or more spikes (Stosiek et al., 2003; Kerr et al., 2005; Greenberg et al., 2008). To detect these AEs, we used a previously published procedure (Greenberg et al., 2008). In short (for more detailed
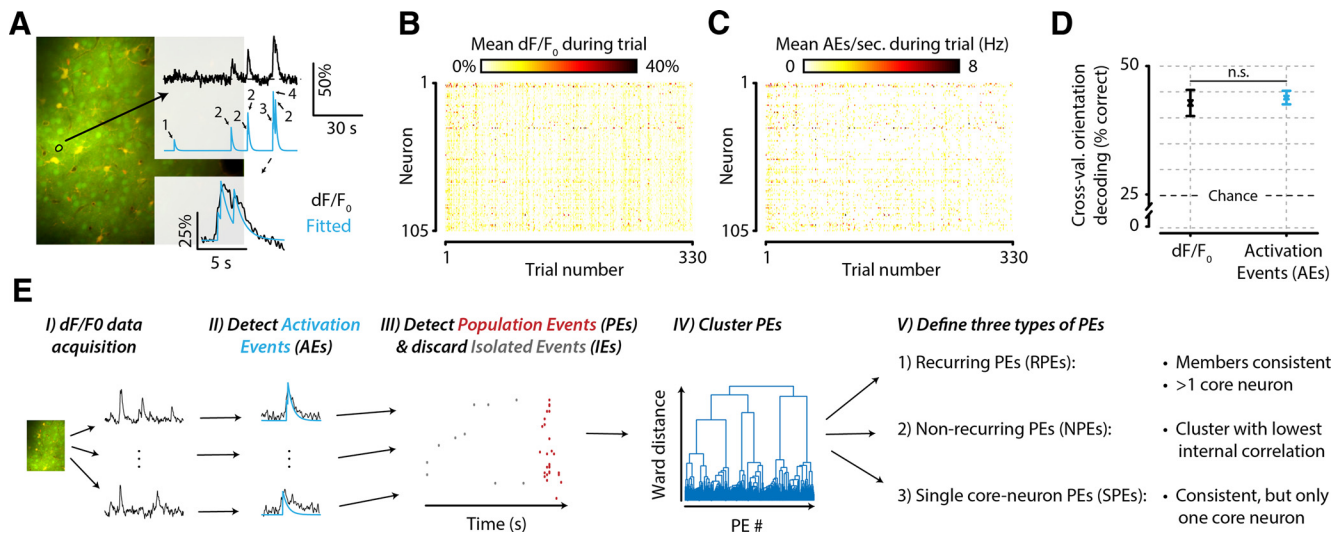
Figure 2. Schematic explaining and validating the procedure of detecting AEs and PEs in calcium imaging data. *A*, For each neuron, we detected putative spiking events with a previously published procedure (Greenberg et al., 2008), which is based on finding a summation of exponential decay functions that minimizes the error with the fluorescent dF/F0 calcium signal. *B*, *C*, Example dataset of one animal before (*B*) and after (*C*) AE detection. Mean neural responses per trial (*x*-axis) and neuron (*y*-axis) in dF/F0 are similar to those in number of AEs (element-wise Pearson correlation, mean ± SEM across animals: $r = 0.78 \pm 0.016$). *D*, A cross-validated orientation decoder performs equally well on raw dF/F0 signals as on AE-detected signals, which represent a drastically data-reduced version of the original signal (paired *t* test across animals, $n = 8$, $p = 0.739$, not significant). Equal decoding performance, despite severe data reduction, shows that the amount of stimulus information in the neural signal is not affected by the AE detection procedure. *E*, Flowchart summarizing the data analysis steps as follows: *EI*, Calcium-imaging data are acquired and preprocessed. *EII*, AEs are detected in each neuron's dF/F0 trace. *EIII*, AEs that are temporally concurrent are grouped into PEs. *EIV*, All PEs are clustered based on which neurons are participating. *EV*, Clustering yields three types of PEs: those that show a recurring configuration of neurons (RPEs), those that do not (NPEs), and those that show a recurring configuration but have only one core neuron (SPEs). n.s., Not significant.

information, see Materials and Methods), this procedure finds the best fit of a summation of exponential decay functions to the dF/F0 trace of a neuron (Fig. 2A–D). We used an information theoretical approach to verify that a similar amount of information on stimulus orientations was present before and after AE detection. We used a cross-validated orientation decoding algorithm (see Materials and Methods) to decode orientation from both the dF/F0 data and the AE detections and found no difference in decoding performance (Fig. 2D) (paired *t* test across animals, $n = 8$, $p = 0.739$). This shows that the information content on stimulus orientation was equally large for the raw dF/F0 signal and AE detected signals. Considering that AE detection represents a significant data reduction compared with dF/F0 signals, this suggests that this detection removed mostly noise, while most of the neural signal was preserved. Moreover, the mean neural activity across trials and neurons was highly correlated between the two quantification methods (dF/F0 and AEs; Pearson correlation, mean ± SEM across $n = 8$ animals: $r = 0.78 \pm 0.016$, *t* test vs 0, $p = 4.0 \times 10^{-10}$) (Fig. 2B,C). For the remainder of the manuscript, we will analyze the temporal behavior of AEs, in particular in relation to the temporal co-occurrence of the activity of multiple neurons (i.e., PEs), stimulus detection, and sequential activation patterns.

Defining PEs and recurring assemblies
Assemblies are often defined as groups of neurons that are repeatedly temporally coactive within a short time window (König et al., 1995; Harris, 2005; Miller et al., 2014). We investigated some of the fundamental properties of recurring population activity and studied whether the recurrence of specific neuronal configurations was correlated with the detection of visual stimuli. To identify such assemblies, we started by grouping all AEs into PEs. In short, a PE was defined as the epoch during which AEs (of any neuron) were continuously

present (i.e., cessation of activity in the population was never >39.4 ms, the interval between subsequent calcium imaging frame acquisitions). All AEs occurring during this period were grouped as belonging to the same PE. AEs that occurred in temporal isolation, or during PEs where fewer than three neurons were active, were discarded and marked as IEs (Figs. 2E, 33A; see Materials and Methods).

After grouping the data into PEs and discarding all AEs occurring outside PEs, we obtained thousands of PEs per animal (mean ± SEM across animals, number of PEs: 3920 ± 527; PEs per second: 1.01 ± 0.21; trials with at least one PE: $80.1 \pm 4.4\%$; $N = 8$ mice). To detect groups of neurons that were repeatedly temporally coactive, we performed a Ward-method clustering procedure of PEs based on the similarity of their constituent neuronal members, and defined the optimal number of clusters using an analysis of silhouette distance (Ward, 1963; Rousseeuw, 1987) (Fig. 3B–D). The optimal number of clusters per animal was 14.9 ± 3.1 (mean ± SEM across animals; also hereafter unless noted otherwise). Next, we calculated for each cluster its consistency (i.e., mean Pearson correlation between all PEs assigned to a given cluster) and defined the clustering procedure's rest group with the lowest mean correlation as nonrecurring PEs (NPEs, events per second: 0.28 ± 0.09) (Fig. 3E). To define which neurons were "core neurons" of a cluster, we calculated which cells were present in significantly more PEs than expected by chance (Fig. 3F,G; see Materials and Methods). All clusters that had more than one core neuron, and were not the rest group, were defined as recurring assemblies (Fig. 3H). We marked all PEs in such recurring clusters as RPEs (events per second: 0.57 ± 0.13). The number of recurring clusters per animal was 8.88 ± 1.95 (mean ± SEM across animals). PEs assigned to a cluster that had only one core neuron were marked as single core-neuron PEs (SPEs, events per second: 0.16 ± 0.02). The time the recorded V1 populations spent in different types of PEs, as a percentage of
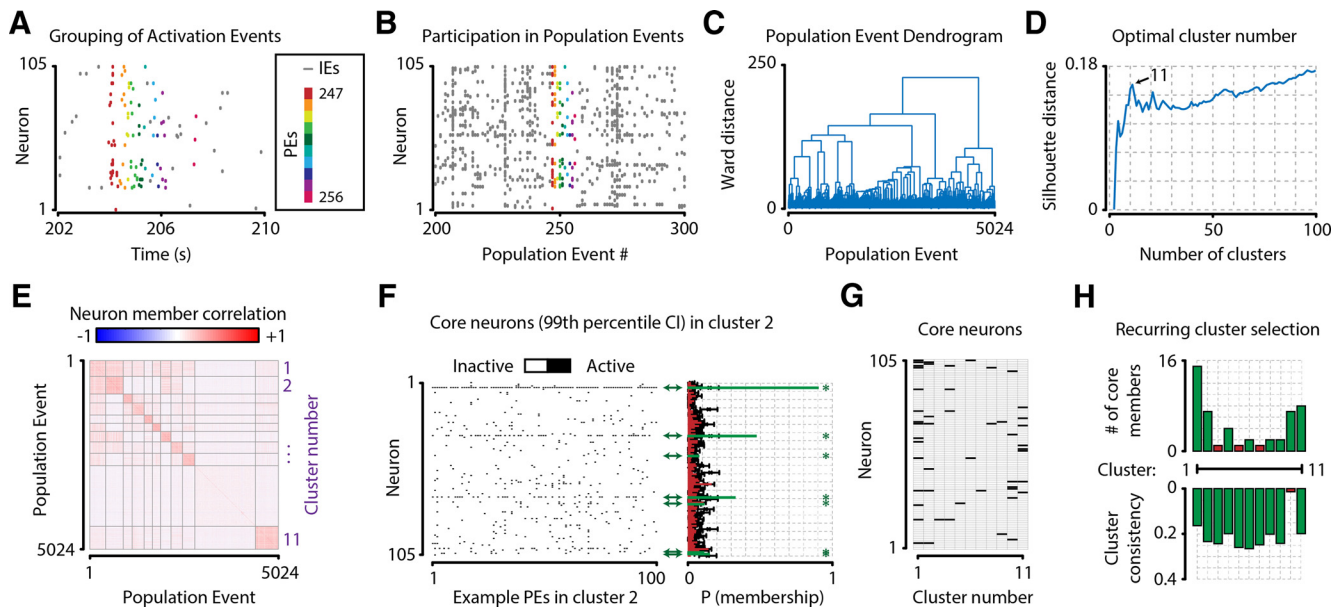
Figure 3. Grouping of AEs into PEs and subsequent Ward-method cluster analysis allow the determination of recurring neuronal clusters. ***A–H***, The procedure in one example animal. ***A***, AEs were grouped into PEs (dots of same color), which were initially defined as epochs where periods of nonactivity were never >39.4 ms (the duration of one acquisition frame). PEs were subsequently discarded if they were <78.8 ms (two acquisition frames) or when fewer than three neurons participated. Discarded events (i.e., IEs, gray dots) were not used for further analysis. ***B***, To determine the similarity in neuronal configuration between PEs, each PE was transformed into a Boolean vector, where each element represents a single neuron, and "True" indicated whether the neuron participated in that event. Colored PEs refer to the same events as those shown in ***A***. Gray PEs represent additional example events. ***C***, A Ward-method clustering procedure on PE membership vectors yielded a hierarchical dendrogram, where short distances between events indicate a high similarity in the configuration of participating neurons. ***D***, A silhouette distance analysis for different numbers of clusters provided an estimate of the optimal number of different configurations of neurons, which in this example animal was 11. ***E***, The correlation matrix of similarity in neuronal members between all PEs, grouped by cluster, shows a higher within-cluster correlation than between-cluster correlation. In this animal, cluster 10 is the "rest group" of NPEs (i.e., the cluster with lowest within-group correlation). ***F***, For each cluster (panel shows cluster 2 as example), we calculated which neurons were "core neurons" using a shuffling procedure that randomizes an AE's neuronal identity and estimates the chance level participation of all neurons in the cluster. Black error bars indicate mean ± 3 SDs. Significant core neurons (i.e., neurons with a participation above the 99th chance percentile) are shown in green (also marked by *) and nonsignificant neurons in red (right). ***G***, Matrix showing core neurons for all 11 clusters. Black represents the neuron is a core neuron. White represents it is not. Neurons could be core members of multiple clusters. ***H***, Clusters were defined as "recurring" when the cluster had at least two core neurons and did not have the lowest within-group correlation of all clusters in a single session. Green bar represents the cluster fulfills these requirements. Red bar represents it does not.

total recording time, was 15.9 ± 4.3% for PEs, 10.0 ± 2.9% for RPEs, 3.9 ± 1.4% for NPEs, and 2.1 ± 0.3% for SPEs. RPEs, NPEs, and SPEs add up to the total number of PEs. In Figures 4 and 5, we analyzed only RPEs; in Figure 7, RPEs and NPEs; and in Figures 6, 8, and 9, all PEs (RPEs, NPEs, and SPEs). Mean PE duration was 146.9 ± 7.8 ms, and the number of participating neurons was 6.1 ± 0.6 (6.7 ± 0.5%).

Members of recurring clusters are anatomically grouped, but not similarly tuned
Having defined which neurons are core neurons of recurring clusters; and when they occur, we addressed whether these clusters' core neurons are anatomically grouped, and share the same orientation preference. For each recurring cluster, we computed the mean pairwise distance in microns between its neuronal core neurons. We compared this to a distribution of expected random mean distances, given the anatomical positions of the neurons. This distribution was obtained by shuffling neuronal identities within the subset of neurons that were a core neuron of any cluster, and recomputing for each iteration the mean pairwise distance of a similarly sized group of neurons (Fig. 4A). The distribution was then used to z-score normalize the cluster's mean pairwise distance (Eq. 9). We repeated this procedure for all recurring clusters of all animals and tested whether the overall z-scored distances were different from 0 (t test vs 0, n = 71 assemblies, mean = −0.431 σ, p = 0.004; Fig. 4B). The distribution was significantly skewed

toward negative values, showing that core neurons of recurring clusters are anatomically grouped.

Next, we asked whether core neurons of recurring clusters were coactive because they shared a similar orientation preference. An analysis of the orientation selectivity index (Eq. 10) of recurring clusters and all recorded neurons showed that these clusters were on average less orientation-tuned than single neurons (two-sample t test of neuronal OSI, n = 891, mean OSI = 0.451) versus recurring cluster OSI (n = 71, mean OSI = 0.373, p = 0.009). Analyzing the OSIs of recurring clusters relative to a shuffle control where we randomized each cluster's core neurons showed a comparable result. When using a similar procedure as described above for investigating anatomical grouping (see Materials and Methods), we found that the recurring cluster OSIs were not significantly different than expected if their core neurons were selected randomly (t test vs 0, n = 71 recurring clusters, mean = −0.127 σ, p = 0.309, not significant; Fig. 4C). We corroborated these results by calculating the similarity of orientation preference of core neurons. This analysis revealed that core neurons of a recurring cluster were not more similarly tuned than chance (t test vs 0, n = 71 recurring clusters, mean = −0.157 σ, p = 0.246, not significant; Fig. 4D). Combining the results from these analyses, we conclude that the grouping of core neurons in recurring assemblies (i.e., recurring clusters) appears to be more related to these neurons being anatomically close to each other than to them preferring similar stimulus features. These

effects appeared to be robust to the minimum number of core neurons required to define a cluster as "recurring," as similar results were obtained with minimum core neuron numbers of up to seven: anatomical grouping was predominantly present, and orientation tuning effects were consistently absent (Table 1).

PE occurrence is independent of prior events at the timescale of seconds

Assembly occurrence in brain regions other than the visual cortex, such as hippocampus, has been associated with replay of past events and other memory-related phenomena (Wilson and McNaughton, 1994; Nádasdy et al., 1999; Lee and Wilson, 2002; Girardeau et al., 2009; Lansink et al., 2009), and it has been suggested that similar mechanisms may operate in rodent V1 (Xu et al., 2012). We therefore investigated whether the occurrence probability of a cluster's PEs was related to prior and/or future occurrences (Fig. 5A), and whether there was an increase of recurrence probability following stimulus presentation and hit responses (Fig. 5B,C). First, we calculated for each recurring cluster from all animals the self-recurrence probability as well as the non–self-recurrence probability as a function of time (Fig. 5D, red and blue traces, respectively). Although a short-term increase in self- or other-cluster occurrence probability was present around the occurrence of a PE (5–10 s), there was no long-term memory component: the occurrence of a PE has no effect on the probability of occurrence of other PEs after 5–10 s (Fig. 5D; FDR-corrected one-sample t test of normalized occurrence probability vs 100%, n = 71 clusters, p < 0.05 only between −2.1 and 2.3 s after PE occurrence for same-cluster PEs, and between −4.8 and 6.7 s for other-cluster PEs). This suggests that there are periods of several seconds within which the probability of a PE is higher than baseline, but longer time-scale interactions in occurrence of RPEs and NPEs were not observable in our dataset. We next investigated the probability of occurrence of PEs as a function of time after stimulus onset and found a stimulus induced, but nearly memoryless behavior: stimulus onset significantly enhanced the probability of a PE occurring, but a couple of seconds after stimulus onset this probability was back to chance level (Fig. 5E; occurrence probability >2 SDs above chance, n = 71 clusters, p < 0.05 only between 0.2 and 2.9 s after stimulus onset). A similar analysis centered at the first licking response during a hit trial showed a comparable result (Fig. 5F; p < 0.05 between −0.7 and 1.1 after hit response), and these effects were still present when PE durations were artificially set to a single frame duration (39.4 ms) (Fig. 5G–I). We therefore conclude that PE occurrences in L2/L3 of mouse V1 do not show experience-dependent replay-like features and show no evidence of a memory-component that persists longer than several seconds. However, these effects all pertain to timescales shorter than several hours and that long-term effects (e.g., learning across
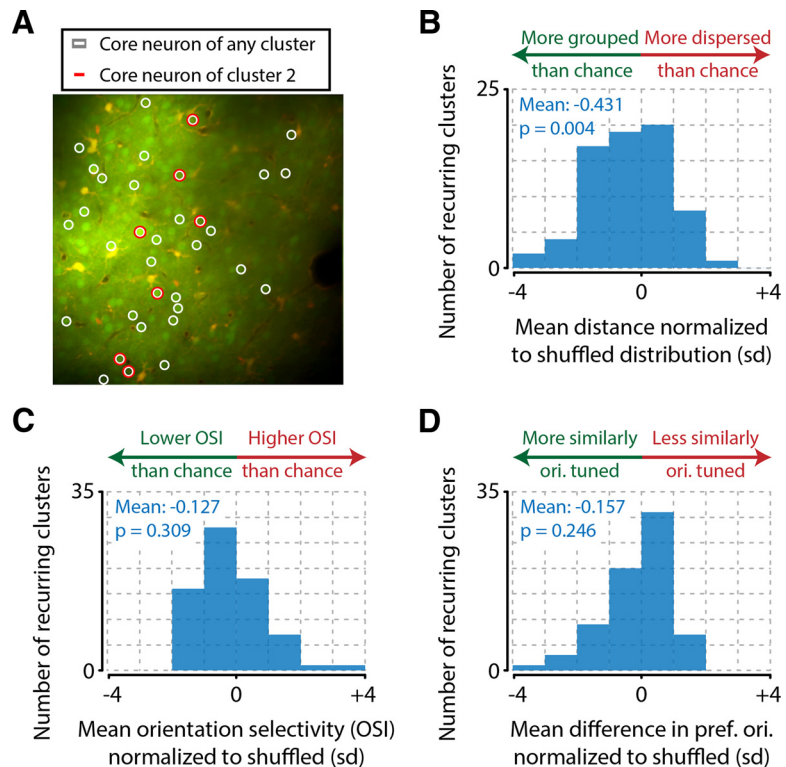

Figure 4. Anatomical proximity between neurons increases the probability of forming a recurring cluster, regardless of orientation tuning. ***A***, Example field of view (average across frames) for one example animal, showing all neurons (green blobs) that were a core neuron in at least one cluster (white circles) and neurons that were a core neuron in the example cluster 2 (red circles). ***B***, Core neurons' somata were spatially widespread across the recording plane but showed a significant tendency to be more spatially localized than chance (one-sample t test of normalized distances vs 0, n = 71 recurring clusters from 8 animals total, p = 0.004). ***C***, Recurring clusters were not more orientation selective than chance (one-sample t test of OSI normalized to chance, n = 71, p = 0.309, not significant). ***D***, This low orientation tuning for recurring clusters can be explained by the observation that core neurons of the same recurring cluster do not share a similar orientation preference. Core neurons were not more similarly tuned than expected from randomly combining neurons regardless of their orientation preference (one-sample t test of normalized difference in preferred orientation, n = 71, p = 0.246, not significant).

of a PE has no effect on the probability of occurrence of other PEs after 5–10 s (Fig. 5D; FDR-corrected one-sample t test of normalized occurrence probability vs 100%, n = 71 clusters, p < 0.05 only between −2.1 and 2.3 s after PE occurrence for same-cluster PEs, and between −4.8 and 6.7 s for other-cluster PEs). This suggests that there are periods of several seconds within which the probability of a PE is higher than baseline, but longer time-scale interactions in occurrence of RPEs and NPEs were not observable in our dataset. We next investigated the probability of occurrence of PEs as a function of time after stimulus onset and found a stimulus induced, but nearly memoryless behavior: stimulus onset significantly enhanced the probability of a PE occurring, but a couple of seconds after stimulus onset this probability was back to chance level (Fig. 5E; occurrence probability >2 SDs above chance, n = 71 clusters, p < 0.05 only between 0.2 and 2.9 s after stimulus onset). A similar analysis centered at the first licking response during a hit trial showed a comparable result (Fig. 5F; p < 0.05 between −0.7 and 1.1 after hit response), and these effects were still present when PE durations were artificially set to a single frame duration (39.4 ms) (Fig. 5G–I). We therefore conclude that PE occurrences in L2/L3 of mouse V1 do not show experience-dependent replay-like features and show no evidence of a memory-component that persists longer than several seconds. However, these effects all pertain to timescales shorter than several hours and that long-term effects (e.g., learning across

days) may very well exist and interact with the creation of recurring events and their occurrence probability (Xu et al., 2012).

Sequences of clusters are more similar during stimulus detection trials than nondetection trials

Having established that long timescale interactions in the occurrence probabilities of clusters did not appear to be present, we next addressed whether short timescale interactions between clusters may be functionally relevant. It has been reported before that particular sequences in states of activity in gustatory cortex can be used to distinguish different olfactory stimuli (Jones et al., 2007). We therefore analyzed whether clusters occurred in sequences specific for the presented orientation, or whether they distinguish miss trials versus hit trials (see Materials and Methods). Because probe trials (0% and 100% contrast) would dilute hit/miss differences, we included only test contrasts (0.5%, 2%, 8%, and 32%) and discarded probe trials for this analysis. We calculated a measure of sequence similarity between two trials, based on the order of cluster occurrence (Fig. 6A). We did not find that different stimulus orientations evoked different sequences of clusters (Fig. 6B,D; t test across animals, N = 8, p = 0.885), but cluster sequences did appear to be more consistent during the detection of visual stimuli (i.e., hits) than when stimuli were not detected (i.e., misses) (Fig. 6C,E; p = 0.002). Using a procedure where we shuffled the identities of clusters to estimate the chance level of sequence similarities, we found that, during hit
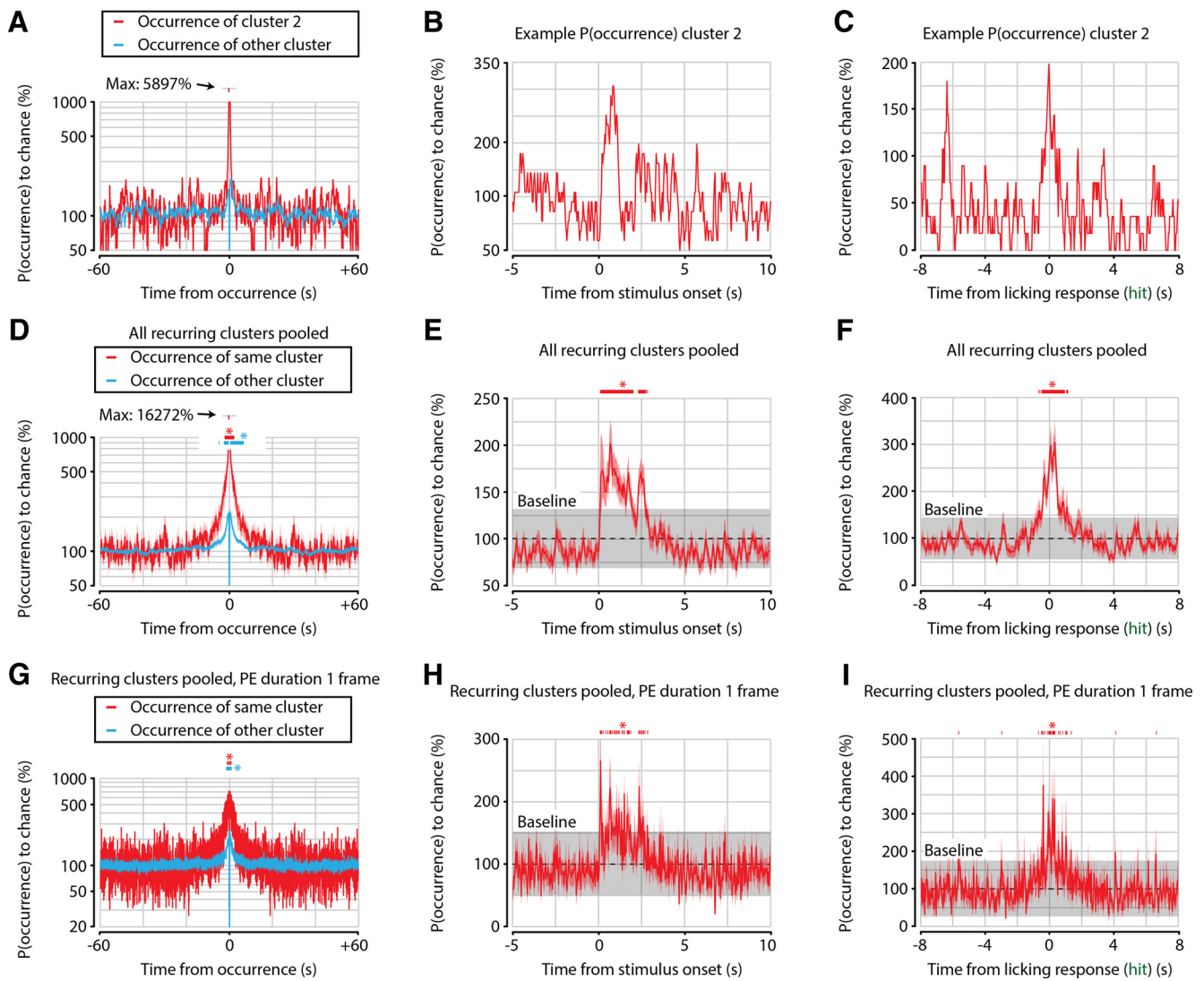
Figure 5. The occurrence of PEs is independent of prior occurrences at timescales longer than several seconds. **A**, Example occurrence probability of PEs of cluster 2 (red) and other clusters (blue) as a function of time after a cluster 2 PE, normalized to chance occurrence (100%). A small peak of increased PE probability for both cluster 2 and other clusters can be seen several seconds around the occurrence of a cluster PE. **B**, Example occurrence probability of cluster 2 PEs as a function of time around stimulus onset. An increased probability of cluster 2 PEs can be seen immediately following stimulus onset. **C**, Similar analysis as in **B**, but now centered on the first licking response during hit trials. **D**, Quantification as in **A**, now showing mean ± SEM across all clusters of occurrence probabilities. Increased probability of occurrence is fairly symmetrical around time of occurrence ($t = 0$ s) and extends for only several seconds backward as well as forward (FDR-corrected one-sample t test vs 100%, $n = 71$ clusters, $p < 0.05$ between −2.1 and 2.3 s for same-cluster PEs, and between −4.8 and 6.7 s for other-cluster PEs). **E**, Quantification of **B**, showing mean ± SEM across all clusters of occurrence probabilities as in **B**. Stimulus onset increases occurrence probability of PEs for several seconds, but occurrence probability quickly returns to chance level with no obvious long-term (>3 s) memory component (probability >2 SDs above baseline, gray shaded area; $n = 71$ clusters, $p < 0.05$ between 0.2 and 2.9 s after stimulus onset). **F**, Quantification of **C**, using same conventions as in **E**. Occurrence probability of RPEs is enhanced for approximately 1 s around the licking response during hit trials (probability >2 SDs above baseline, $p < 0.05$ between −0.7 and 1.1 s after hit response). **G–I**, Same as **D–F**, but now the duration of each PE is artificially set to one acquisition frame (39.4 ms). This control analysis shows that the short timescale temporal correlations in **A–F** are not due to the relatively long durations of PEs.

trials, sequences were more similar than expected by chance (hits: t test, $N = 8$ animals, real vs shuffled: $p = 5.7 \times 10^{-5}$), but no such effect was observable for miss trials (misses: real vs shuffled, $p = 0.202$). Importantly, this sequence effect was present without individual clusters preferentially occurring during hit or miss trials, as the distribution of the presence of clusters during hit and miss trials was not significantly different (Fig. 6F; t test of means, hit vs miss, $p = 0.079$; F test of variance, hit vs miss, $p = 0.106$). Moreover, this effect was not correlated with licking behavior itself, as a separate analysis of FAs and CRs showed no difference in sequence similarity (t test, FAs vs CRs, $p = 0.329$). This suggests that sequences of clusters are more similar when stimuli are detected, above and beyond individual clusters being correlated with either hit trials or miss trials.

PEs are stereotyped, but neither the identity nor occurrence frequency of clusters is correlated with detection

While assemblies in L2/L3 of mouse V1 do not appear to be involved in coding for grating orientation, they do appear to play a role in the visual detection of stimuli. We hypothesized that not only the sequence in which a cluster occurred might be relevant for stimulus detection, but that also the underlying properties (such as the duration, number of participating neurons, or level of neuronal activity) of recurring and NPEs may change between hits and misses. For the following analyses, we therefore split all trials into hits and misses (0.5%–100% contrast), and into CRs and FAs (0% contrast). We selected all PEs that occurred during the first 2 s after stimulus onset and split them into PEs of recurring clusters (RPEs) and PEs of nonrecurring clusters (NPEs).
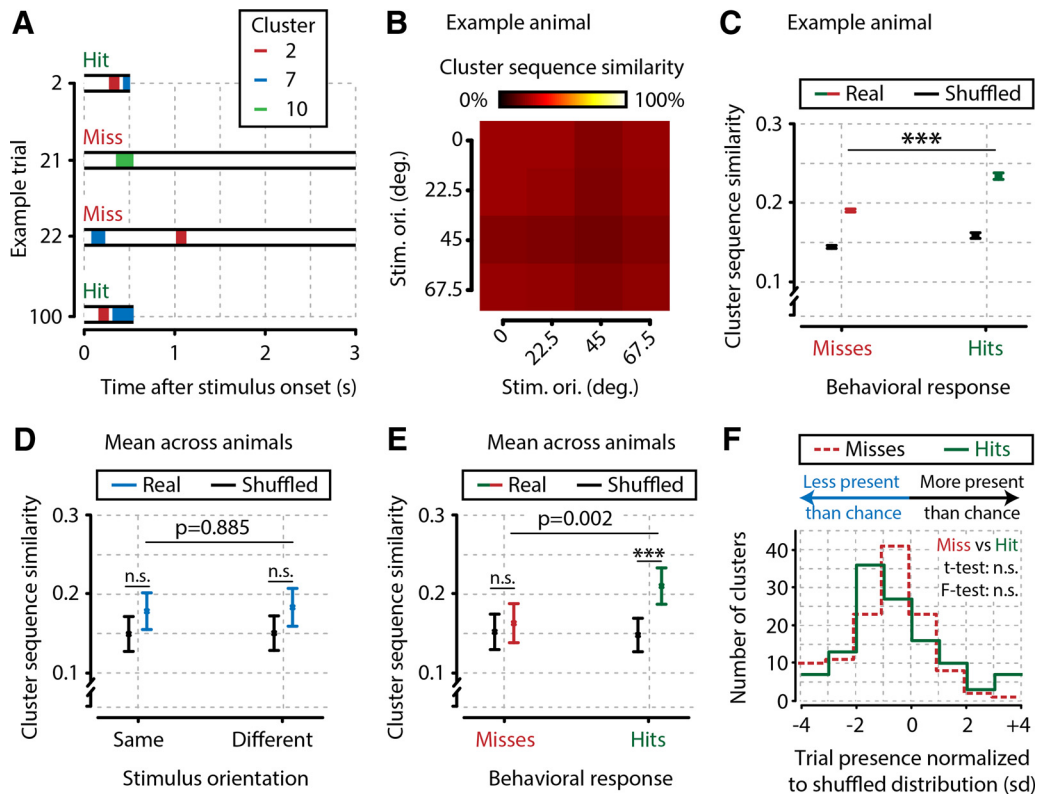
Figure 6. Clusters occur in more similar sequences during trials where the stimulus is detected than during trials where it remains undetected. **A**, Cluster occurrences during four example trials, including also nonrecurring clusters (i.e., cluster 10). Trials 2 and 100 show a sequence similarity of 1.0, whereas the other combinations of trials show a similarity of 0.0. **B**, Cluster sequence similarity for this example animal is not higher for trials with the same stimulus orientation (diagonal entries) than for trials with different orientation (off-diagonal entries), showing that stimulus orientation is not encoded in the sequence of cluster occurrences (all *t* tests between sequence similarities based on stimulus orientation, $p > 0.05$; not significant). **C**, Same example animal as in **B**, showing a higher cluster sequence similarity between hit trials than between miss trials ($p = 1.04 \times 10^{-28}$, $n = 9702$ hit trial pairs, $n = 26,406$ miss trial pairs). Black bars represent baseline (shuffled) similarity values expected by chance. **D**, Across animals, we found no evidence for stimulus orientation being encoded by sequences of clusters (same orientation trial pairs: *t* test of real vs shuffled, $p = 0.330$; different orientation trial pairs: *t* test of real vs shuffled, $p = 0.384$; *t* test of same vs different orientation trial pairs: $p = 0.885$; $n = 8$ animals). **E**, Cluster sequence similarity is higher between hit trials than between miss trials (*t* test, $p = 0.002$), and sequence similarity between hit trials is higher than expected by chance (hits: *t* test of real vs shuffled: $p = 5.7 \times 10^{-5}$), but not for miss trials (misses: *t* test of real vs shuffled, $p = 0.202$). **F**, The enhanced sequence similarity for hit trials cannot be explained by particular clusters occurring preferentially during hit or miss trials: the distribution of cluster presence during trials, normalized to chance, is not different for hits and misses in terms of mean and variance; hits (mean = −0.433, SD = 2.140) versus misses (mean = −0.890, SD = 1.843), *t* test, $p = 0.079$; *F* test, $p = 0.106$; $n = 119$ clusters (including nonrecurring clusters). ***, $p < 0.001$. n.s., Not significant.

For each contrast and response type, we calculated the number of RPEs and NPEs, their mean duration, the number of participating neurons per second, and the number of AEs per second (Fig. 7). Perhaps surprisingly, none of these properties differed between hits and misses or between FAs and CRs (for all four properties: FDR-paired *t* tests across animals, $n = 8$, hit vs miss, or FA vs CR, for each contrast, $p > 0.05$; not significant). Moreover, only the number of NPEs increased as a function of contrast (FDR-corrected paired *t* tests across animals, contrast occurrences vs baseline, 2 s period preceding stimulus onset, occurrences; FAs and hits: 0%, $p = 0.744$; 0.5%, $p = 0.184$; 2%, $p = 0.038$; 8%, $p = 0.014$; 32%, $p = 0.0011$; 100%, $p = 0.0007$; CRs and misses: 0%, $p = 0.895$; 0.5%, $p = 0.352$; 2%, $p = 0.045$; 8%, $p = 0.045$; 32%, $p = 0.036$; 100%, $p = 0.045$). All other comparisons versus baseline were nonsignificant ($p > 0.05$). Moreover, although we observed a wide range of hit/miss occurrence ratios across recurring clusters, the overall distribution was not significantly different from chance (permutation test, $n = 71$ clusters, $p = 0.552$, not significant). We therefore conclude that PEs in general occur in a relatively stereotyped manner; all PEs are approximately equally long, of the same size (in neurons), show the same number of AEs per second, and these properties are not related to the behavioral detection of visual stimuli.

Neurons in mouse V1 show a preferred temporal position in PEs

Previous studies investigating stereotyped population AEs (or "packets") in sensory cortex have mostly been performed in auditory and somatosensory areas (Luczak et al., 2007, 2013, 2015; but see also Carrillo-Reid et al., 2015). We therefore analyzed our data to check whether temporal sequences that characterize these packets may be present in visual cortex. In short (see also Materials and Methods), we calculated, for each PE in the entire dataset, the center of mass in time of its AEs. For each neuron, we then calculated the mean latency in milliseconds of its AEs to the PE's center of mass (Fig. 8A). Next, we computed for each cluster the average AE latency per neuron (Fig. 8B), and correlated this within-cluster ordering with the ordering obtained when computed across the entire dataset (Fig. 8C, blue points). We performed a shuffling control to correct for potential biases in our analysis of sequence consistency (i.e., the Pearson correlation with the mean overall sequence). Per shuffle iteration, the neuronal identity of all AEs was randomized; therefore, the temporal structure of the population response remains intact, but the timing per neuron is random (see Materials and Methods). As expected, these shuffled distributions yielded consistency values close to 0 (Fig. 8C). Comparing the consistency of neuronal acti-
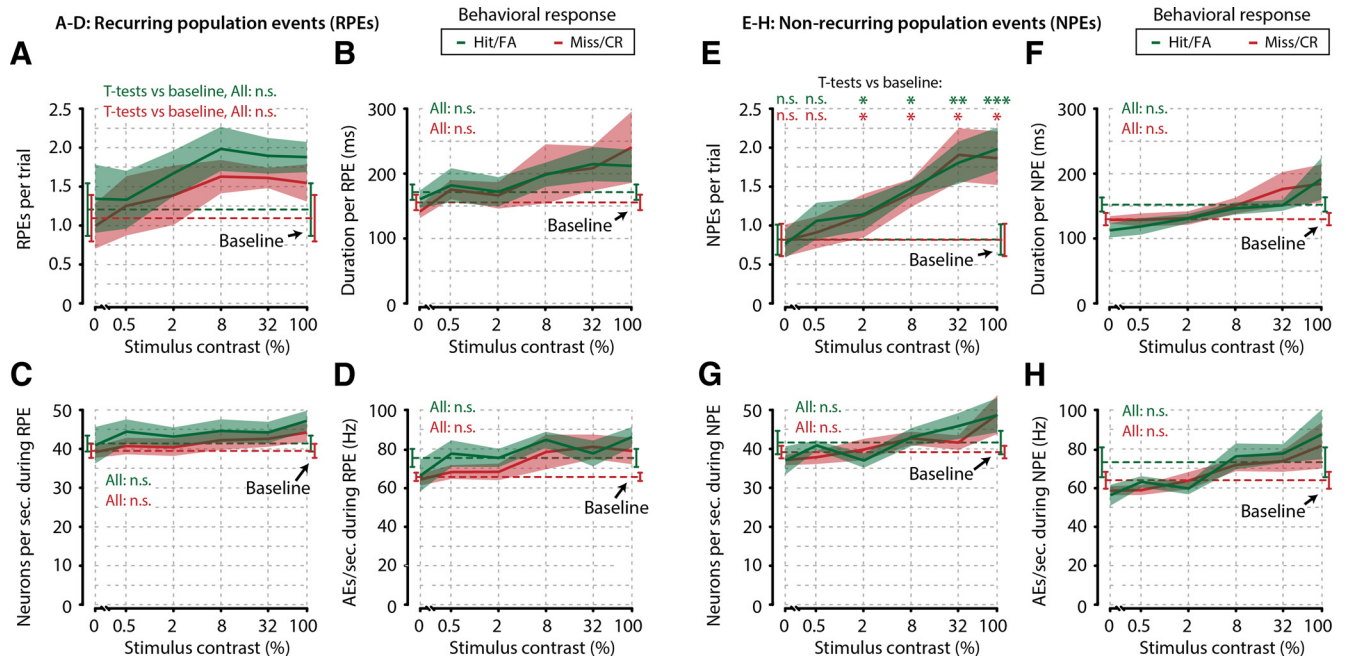
Figure 7. Several parameters of recurring and NPEs show no correlation with stimulus detection, and all PEs are stereotyped. ***A–D***, Number of RPEs per trial (***A***), duration per RPE (***B***), number of neurons participating per second during RPEs (***C***), and number of AEs per second during RPEs (***D***), as a function of stimulus contrast, were identical when the animals made a response (green, Hit/FA) compared with when they made no response (red, Miss/CR) (FDR-corrected paired *t* tests, *n* = 8 animals, 6 tests per property, *p* > 0.05 for all comparisons, not significant). Moreover, these properties were not different during spontaneous PEs (Baseline, i.e., occurrences during 2 s preceding stimulus onset) compared with stimulus-evoked PEs (FDR-corrected paired *t* tests, *n* = 8 animals, *p* > 0.05 for all comparisons, not significant). ***E–H***, Same as in ***A–D***, but for NPEs. As for recurring events, there was no correlation of NPE properties with stimulus detection (FDR-corrected paired *t* tests, *n* = 8 animals, 6 tests per property, *p* > 0.05 for all comparisons, not significant). Interestingly, the number of NPEs did show a dependence on stimulus contrast (FDR-corrected paired *t* tests, stimulus-driven occurrences vs baseline, hits/FAs: 0%, *p* = 0.744; 0.5%, *p* = 0.184; 2%, *p* = 0.038; 8%, *p* = 0.014; 32%, *p* = 0.0011; 100%, *p* = 0.0007; misses/correct rejections: 0%, *p* = 0.895; 0.5%, *p* = 0.352; 2%, *p* = 0.045; 8%, *p* = 0.045; 32%, *p* = 0.036; 100%, *p* = 0.045), but several properties of NPEs did not (f-h, FDR-corrected paired *t* tests, *n* = 8 animals, *p* > 0.05 for all comparisons, not significant). These results show that some basic properties of PEs are statistically stereotyped and are not related to stimulus detection. *, *p* < 0.05; **, *p* < 0.005; ***, *p* < 0.001 n.s., Not significant.

vation sequences across animals (*n* = 8 animals, one mean consistency value per animal), we found that RPEs (mean ± SEM, *r* = 0.505 ± 0.06, *t* test vs 0, *n* = 8, *p* = 5.37 × 10⁻⁵), but also NPEs and SPEs (mean ± SEM, *r* = 0.497 ± 0.05, *p* = 1.71 × 10⁻⁵), showed significant temporal sequence consistencies (difference between RPEs and NPEs/SPEs, *p* = 0.924, not significant) (Fig. 8D). Nonrecurring clusters (and to a lesser degree also recurring clusters) show varying neuronal members with each occurrence. The reported temporal consistency should therefore not be taken as an indication of a fixed neuronal sequence, but rather as an indication that each neuron shows a consistent latency with respect to the center of mass of a PE, regardless of the identity of other neurons participating in the event. We therefore conclude that, even with the relatively modest temporal resolution of calcium imaging (25.4 Hz), neurons have a clear preferred temporal position within PEs. However, this does not rule out the possibility of finding different effects, or an even stronger temporal structure in population activity, when an analysis at shorter timescales would have been possible. Considering that neurons show a consistent temporal latency, and core neurons of recurring clusters tend to be anatomically grouped, we hypothesized that these phenomena might represent aspects of traveling waves of neuronal activity. We therefore investigated whether, for pairs of neurons, there was a correlation of the difference in latency with the anatomical distance between somata. We found no such consistent effect across animals (*t* test of Pearson correlations across animals, *n* = 8; *r* = 0.006 ± 0.019, mean ± SEM, *p* = 0.751, not significant), suggesting that anatomical grouping and preferred temporal positions of neuronal activation are likely not related to traveling waves.

Temporal position of activation is more reliable during stimulus detection

We hypothesized that the temporal order of neuronal activation within PEs might be more important for visual stimulus detection than the features of PEs examined above, such as their rate of occurrence, duration, size, and number of AEs. We therefore investigated the consistency of neuronal sequences in PEs during miss and hit trials. As before, we included only test contrasts (0.5%, 2%, 8%, and 32%) and discarded probe trials for this analysis. To ensure equal durations of time windows, we considered only PEs that occurred during the first 2 s after stimulus onset. Given our previous results that PE properties changed little as a function of contrast or orientation, we pooled the data over all orientations and contrasts. For each PE that occurred during a miss or hit trial, we calculated its consistency with the overall sequence (i.e., Pearson correlation of neuronal latencies; Fig. 8E; see Materials and Methods). We observed a consistent effect across animals (*n* = 8), where neuronal population activity during hits showed a more consistent temporal sequence than during misses (Fig. 8F) (paired *t* test, hit vs miss sequence consistency, *n* = 8 animals, *p* = 0.022).

Given the potential importance of this result, we examined this difference using an additional, alternative analysis of temporal sequence stability. We investigated the variability of neuronal AE latencies in PEs during miss and hit trials. As before, we calculated AE latencies during single PEs that occurred during miss and hit trials and normalized all latency values by the duration of the PE (Fig. 9A). Next, we calculated the variability of the preferred temporal position of each neuron across PEs and found that this "jitter" was lower during hits than misses (Fig. 9B)
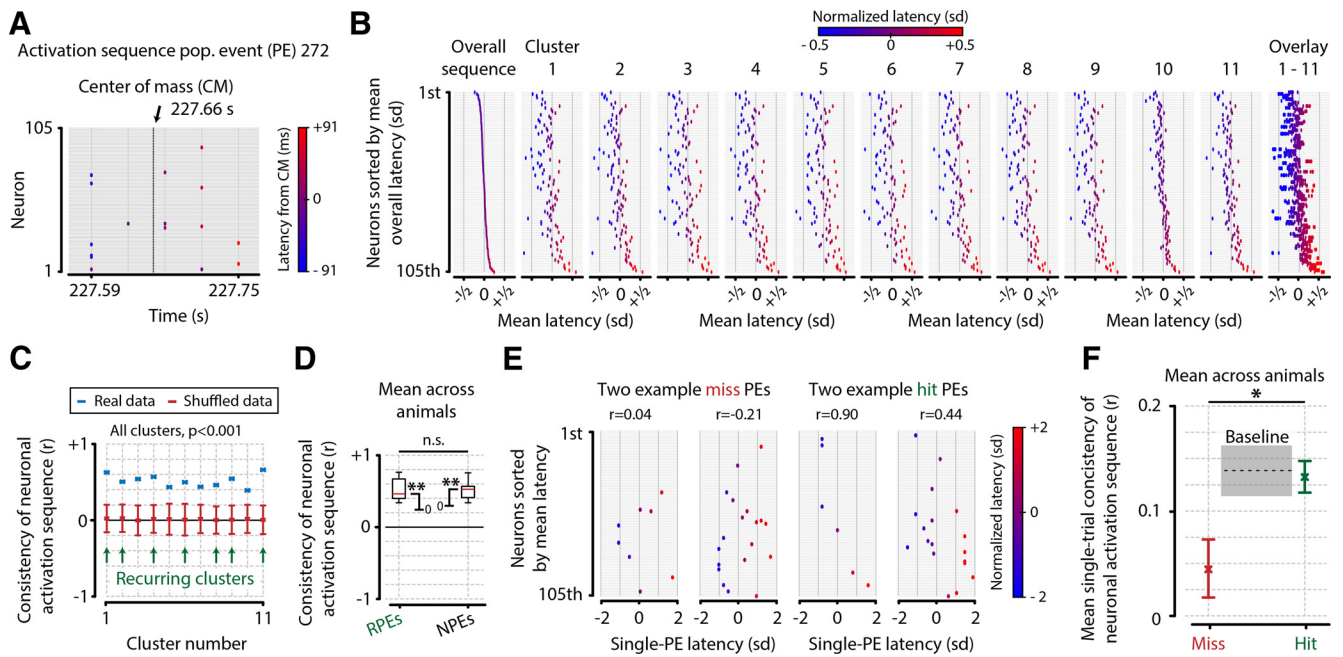
Figure 8. Neurons show consistent, preferred temporal locations of activity within recurring and NPEs, and this temporal positioning is more precise during stimulus detection. **A**, For each PE (panel shows an example event), we calculated the mean activation latency (in milliseconds) per neuron to the center of mass of the event. **B**, Example sequences of activation for all clusters averaged across the cluster's PEs, ordered by position in the overall sequence. **C**, Quantification of **B**, showing that, across all clusters, including both recurring and nonrecurring events, the temporal sequence of activation is consistent, and significantly more so than expected by chance (permutation test vs 0, $n = 11$, for all clusters, $p < 0.001$). **D**, Mean across animals ($n = 8$) of sequence consistency was significantly >0, for both recurring clusters (mean consistency $\pm$ SEM, $r = 0.505 \pm 0.06$, t test vs 0, $n = 8$, $p = 5.37 \times 10^{-5}$) and nonrecurring clusters (mean $\pm$ SEM, $r = 0.497 \pm 0.05$, $p = 1.71 \times 10^{-5}$). The difference in consistency between recurring and nonrecurring events was not significant (paired t test, $n = 8$, $p = 0.924$, not significant). This suggests that the preferred temporal position of a neuron's activation is not related to the specific configuration of active neurons but is rather a universal property of all PEs. **E**, **F**, Single PE, rather than cluster-based sequence analyses. **E**, Four example PEs during stimulus presentation of test contrasts (0.5%–32%), showing correlation values with the overall sequence. The displayed events were chosen to illustrate the difference between high and low sequence consistency and are therefore not representative of the actual hit/miss effect size. **F**, Quantification of average single-trial sequence consistency during misses (red), hits (green), and preceding stimulus presentation (gray), showing that population sequences of neuronal activity are more temporally consistent when the visual stimulus is behaviorally reported (paired t test of miss vs hit consistency, $n = 8$ animals, $p = 0.022$). Colored error bars and gray baseline indicate mean $\pm$ SEM across animals. *, $p < 0.05$; **, $p < 10^{-4}$ n.s., Not significant.

(paired t test, hit vs miss variability in latency, $n = 8$ animals, $p = 0.035$). We therefore conclude, after having shown the same effect with two different analyses, that a neuron's activity within a PE is more temporally consistent during the detection of visual stimuli.

Orientation coding and sequence consistency are uncorrelated

It has been reported that neuronal population representations of stimulus features are more accurate during the detection of stimuli (Montijn et al., 2015; Yang et al., 2016). To investigate whether temporal sequence consistency was also correlated with an enhanced orientation decoding accuracy, we decoded stimulus orientation as before for each trial (Fig. 2D; see Materials and Methods), and split all trials into hits or misses (Fig. 9C), or into trials belonging to the lowest 50% and highest 50% of sequence consistency (Fig. 9D). Interestingly, although both sequence consistency (Figs. 8F, 9B) and decoding accuracy were correlated with stimulus detection (Fig. 9C; paired t test, $n = 8$ animals, $p = 0.026$), decoding accuracy was not correlated with sequence consistency (Fig. 9D, $p = 0.660$, not significant). This suggests that population code accuracy and temporal sequence consistency are two separate neural phenomena independently correlated with the detection of visual stimuli.

Finally, we addressed the potential concern that our 2-s-long analysis window was excessively long. We therefore repeated the analyses of sequence variability, stimulus orientation decoding, and their interaction, but now with a 1 s analysis window. As this entails a significant reduction in amount of data, we expected it would be more difficult to detect small differences due to decreased statistical power. Indeed, the sequence variability effect was now bordering on significance (Fig. 9E; paired t test, $n = 8$ animals, $p = 0.050$). On the one hand, converging evidence (Figs. 8F, 9B) shows that sequence consistency is correlated with stimulus detection regardless of the specifics of the underlying analysis, but on the other hand our results also suggest that this correlation is fairly subtle and only reaches statistical significance when working with large datasets (Fig. 9E). Supporting our previous results, we found that with a 1 s analysis window stimulus orientation decoding was significantly higher for hit than miss trials (Fig. 9F; paired t test, $n = 8$ animals, $p = 0.004$), and as before there was no correlation between sequence consistency and orientation decoding performance (Fig. 9G; paired t test, $n = 8$ animals, $p = 0.990$). To further query the reliability of these results, we also performed the same analyses as above with a variable analysis window. Now, the duration of the window was delimited by the licking response during hit trials, and miss trial durations were randomly assigned from the distribution of hit trial durations. This analysis showed that our results were robust to changes in the specific size of the analysis window, as the variability in latency was lower for hits than misses ($p = 0.045$; Fig. 9I), decoding accuracy higher ($p = 0.008$; Fig. 9J), and latency variability and decoding accuracy were not correlated on a trial-by-trial basis (Fig. 9K).
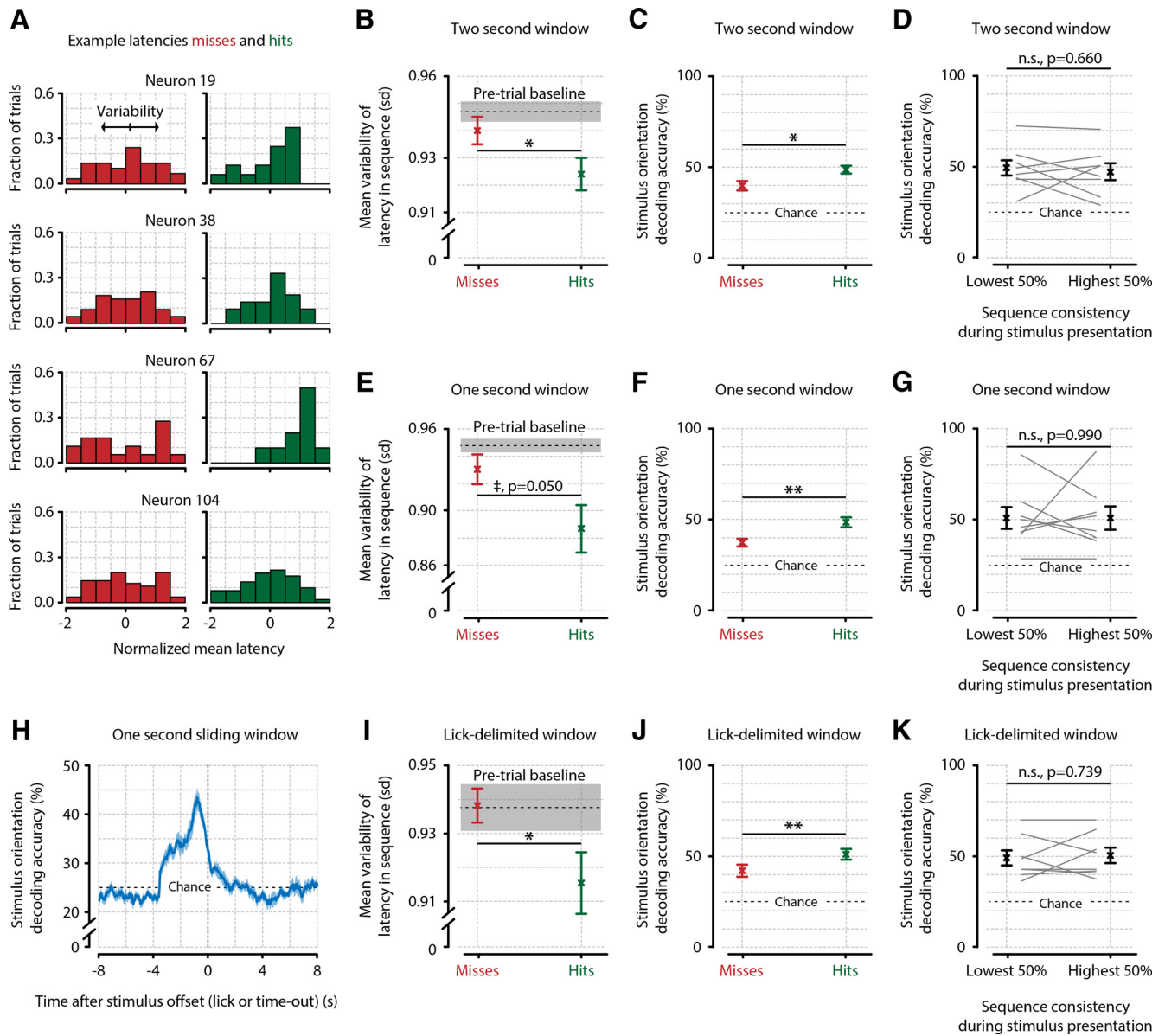
Figure 9. Temporal sequence consistency and orientation decoding accuracy correlate with the detection of stimuli, but not with each other. **A**, Four example neurons, showing the distribution of latencies across miss trials (red; left) and hit trials (green; right). **B–G**, The variability is calculated by taking the SD of latencies across PEs per neuron, and averaging across neurons. **B**, Quantification of the variability in latency across animals ($n = 8$) shows that the temporal sequence of neuronal activation is less variable during the detection of a stimulus (hit) than when it remains unreported (miss) (paired t test, $n = 8$, $p = 0.035$), confirming the results shown in Figure 8F. **C**, Decoding of stimulus orientation shows that neuronal population activity represents stimulus orientation more accurately when it is behaviorally reported (miss vs hit decoding accuracy, paired t test, $n = 8$ animals, $p = 0.026$). **D**, Sequence consistency and decoding accuracy are not correlated across trials, suggesting they represent two neural phenomena that independently correlate with stimulus detection (paired t test, $p = 0.660$, not significant). **E–G**, Same as **B–D**, but with a shorter analysis window of 1 s after stimulus onset. A shorter analysis window leads to the latency variability correlate of stimulus detection showing borderline significance (**E**, paired t test, $n = 8$ animals, $p = 0.050$) but does not influence the correlation of decoding accuracy with stimulus detection (**F**, $p = 0.004$) or the lack of interaction between the two (**G**, $p = 0.990$). **H**, Sliding-window (1 s) orientation decoding, centered on stimulus offset. Shortly after stimulus offset, orientation decoding performance returns to chance level (25%, horizontal dotted line). **I–K**, Same as **B–D**, but now for analysis windows that are delimited by the licking response for hit trials, and by random assignment of hit-trial durations for miss trials. The results are quantitatively and qualitatively similar to taking a 2 s window, showing that the specific parameters of the time windows chosen do not strongly affect the results (hit vs miss t tests; **I**: latency variability, $p = 0.045$; **J**: stimulus decoding accuracy, $p = 0.008$; **K**: correlation between orientation decoding and sequence consistency, $p = 0.739$). *, $p < 0.05$; **, $p < 0.01$ n.s., Not significant.

Discussion

Our data show that V1 population activity of mice performing a stimulus detection task is characterized by the presence of recurring, coactive groups of neurons (i.e., recurring clusters, or assemblies), intermixed with other configurations of coactive neurons. These recurring clusters are most likely formed through processes dependent on anatomical proximity of neuronal members rather than their stimulus selectivity. Our results further suggest that these events do not show experience-dependent replay-like behavior, but leaves open a potential role in short-term or iconic-like memory. Together, our results suggest that neural processing of information related to detection behavior depends more on the accuracy of temporal positioning of the activation by individual neurons, and on the sequence in which these clusters of neurons are active, than on the simple activation of particular neurons or configurations of neurons.

Potential confounds

A potential confound of our trial-related results is the use of the fixed 2 s window in our analyses, as the onset of licking responses varies. We therefore also performed some important trial-related analyses with windows of 1 s, and with lick-delimited windows. These analyses mostly reproduced our previous findings (Fig. 9E–K), but the correlation between stimulus detection and V1 temporal sequence consistency was occasionally of modest effect size (e.g., Fig. 9E) and likely reaches statistical significance only when averaging over large datasets. One explanation for this modest effect size is that the definition of a PE we use in our analyses cannot disentangle sequences that might be overlapping in time, and could have occasionally grouped multiple PEs into a single PE.

A second caveat could be that the use of two-photon calcium imaging is ill suited to investigate temporal properties of neuronal population activity. Although the low temporal resolution of calcium imaging (39.4 ms, or 25.4 Hz) is suboptimal, the technique offers the advantage of recording large populations of unambiguously isolated neurons simultaneously, with accurate micro-anatomical location. Especially when studying population-related phenomena, having a sufficient number of well-isolated neurons is essential. We note, however, that other assembly characteristics and temporal aspects of population activity might have been uncovered if a high temporal resolution had been available. The choice for two-photon calcium imaging was therefore a trade-off of temporal accuracy in favor of robust population sizes. Moreover, we used the synthetic indicator OGB rather than GCaMP6f, as OGB has faster dynamics and therefore allows a more temporally precise estimate of spiking activity (J. L. Chen et al., 2013). This diminishes, but does not negate, the fact that our relatively low temporal resolution may have added uncertainty when estimating the temporal position of a neuron's activity within a PE. The fact that we find significant effects despite this uncertainty suggests that the underlying neural phenomena are robust.

Furthermore, the statistically significant, but modest behavioral performance of our animals could present a potential confound. Relatively mild water restriction applied during training led to a detection performance at the 100% probe trials that is lower compared with other studies using similar tasks (Glickfeld et al., 2013). This could mean that the observed differences between hit and miss trials are diluted, as the animal may have in fact detected the stimulus in a significant fraction of "miss" trials. As we have also discussed previously in more detail, this would mean that the effect sizes we report in the current study are likely lower than when the behavioral performance had been high (Montijn et al., 2015).

Relation to other studies investigating assembly activity in visual cortex

Previous studies have reported that neuronal population activity in passively stimulated mouse visual cortex is characterized by the presence of recurring ensembles of neurons (Miller et al., 2014; Carrillo-Reid et al., 2015). It has also been reported by the Yuste laboratory that population activity patterns show consistent sequential activity of neurons within ensemble events (Carrillo-Reid et al., 2015). We confirm these observations using a paradigm where the animals are actively involved in a visual detection task. Moreover, we report the novel findings that cluster sequence and within-cluster sequence precision are correlated with stimulus detection, and that V1 PEs have stereotyped characteristics, such as duration and number of participating neurons. However, we also note differences with the aforementioned

studies. First, it has been suggested that particular scenes in a natural movie can recruit specific ensembles (Miller et al., 2014; Carrillo-Reid et al., 2015), but in our data we do not find any specific orientation tuning of recurring assemblies (Fig. 4C,D). This could be due to ensembles not being directly related to processing orientation, a different computational definition of assemblies, the use of a paradigm that involved behavioral stimulus detection, but not orientation discrimination, or a lack of power in our study. In our behavioral task, the orientation of stimuli was behaviorally irrelevant to the mice, which may explain why we did not find any involvement of clusters in orientation-based neural responses. This could also explain why Jones et al. (2007) found a correlation of cortical state sequences with different olfactory stimuli, whereas we found a correlation with stimulus detection, but not stimulus orientation. Perhaps a correlation between stimulus identity and cluster sequences in sensory cortex only appears when stimuli are behaviorally relevant.

Alternatively, a lack of power seems unlikely to explain the difference in results, as we do find a highly significant anatomical grouping of a cluster's core neurons (Fig. 4A,B) that was not reported previously (Miller et al., 2014; Carrillo-Reid et al., 2015). Notably, while Yuste et al. (Carrillo-Reid et al., 2015) found a significant correlation between certain stimuli and the presence of particular ensembles, these authors mention that their ensembles regularly contain neurons with broad orientation tuning. This suggests that their ensembles might be correlated with certain stimuli because of the underlying noise correlation structure of their members in response to those stimuli. It has been reported that noise correlations between pairs of neurons are weaker with longer intersomatic distances (Hansen et al., 2012; Montijn et al., 2014; Goltstein et al., 2015), which may help explain why cluster core members tend be anatomically grouped. The results from our study and the aforementioned authors may therefore be more complementary than contradictory.

Hypothesized function of temporal cluster sequences and rate codes

Our data show that recurring clusters are present in L2/L3 populations of mouse V1, but the presence of such assemblies per se does not seem to serve a particular function in the context of visual detection. Instead, we found that neurons show a preferred temporal position across all PEs and that the precision of this temporal positioning, as well as the sequence of clusters, appears behaviorally relevant for the detection of stimuli. Together, these observations fit well within the framework of the packet-communication hypothesis as described previously using electrophysiology (Luczak et al., 2007, 2013, 2015; Junek et al., 2010). The hypothesized function of these packets is to gate the flow of information; every event represents a "read-out" of that cortical population by its postsynaptic targets (Luczak et al., 2013). Our current results show that a similar mechanism may exist in visual cortex (Figs. 678-9) (Carrillo-Reid et al., 2015). However, one could ask how the relative unimportance of the identities of neurons participating in PEs may be reconciled with a transfer of specific stimulus information. One explanation may be that the function of PEs is to provide a general level of subthreshold depolarization in postsynaptic populations (i.e., to evoke an "up-state" of several tens to hundreds of milliseconds), during which neurons with specific stimulus tuning can more easily cause action potential generation in postsynaptic neurons. This would explain why the specific configuration of active neurons is not important, as well as why such PEs are relatively long (often >100 ms) (Luczak et al., 2007, 2013). Inconsistent temporal sequences

might then reflect suboptimal presynaptic driving of postsynaptic up-states and would explain why sequence consistency is correlated with stimulus detection. Moreover, our observation that core neurons of recurring clusters are anatomically grouped suggests that the generation of these postsynaptic up-states may be local in retinotopic space. Finally, our finding on cluster sequence order suggests that V1 assembly behavior related to stimulus detection cannot be simply explained by enhanced excitability. If distinct neuronal clusters in V1 differentially project to specific postsynaptic subpopulations in downstream areas, particular cluster sequences may regulate assembly behavior and network dynamics in V1's projection targets.

In conclusion, our data suggest that recurring assemblies are present in layer 2/3 of primary visual cortex in mice performing a stimulus detection task. V1 response patterns appear to be characterized by stereotyped events of activity in a packet-like manner, as previously described in other cortical regions of passively stimulated animals (Luczak et al., 2015). Rather than assembly presence *per se* being involved in stimulus detection, we found that the sequential activation of clusters, as well as the precision of temporal positioning of neuronal responses within population AEs, is correlated with visual detection.

References

Adibi M, McDonald JS, Clifford CW, Arabzadeh E (2014) Population decoding in rat barrel cortex: optimizing the linear readout of correlated population responses. PLoS Comput Biol 10:e1003415. CrossRef Medline

Averbeck BB, Latham PE, Pouget A (2006) Neural correlations, population coding and computation. Nat Rev Neurosci 7:358–366. CrossRef Medline

Balleine BW, Dickinson A (1998) Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. Neuropharmacology 37:407–419. CrossRef Medline

Beck JM, Ma WJ, Kiani R, Hanks T, Churchland AK, Roitman J, Shadlen MN, Latham PE, Pouget A (2008) Probabilistic population codes for Bayesian decision making. Neuron 60:1142–1152. CrossRef Medline

Benucci A, Saleem AB, Carandini M (2013) Adaptation maintains population homeostasis in primary visual cortex. Nat Neurosci 16:724–729. CrossRef Medline

Brainard DH (1997) The Psychophysics toolbox. Spat Vis 10:433–436. CrossRef Medline

Buran BN, von Trapp G, Sanes DH (2014) Behaviorally gated reduction of spontaneous discharge can improve detection thresholds in auditory cortex. J Neurosci 34:4076–4081. CrossRef Medline

Cafaro J, Rieke F (2010) Noise correlations improve response fidelity and stimulus encoding. Nature 468:964–967. CrossRef Medline

Carnevale F, de Lafuente V, Romo R, Parga N (2013) An optimal decision population code that accounts for correlated variability unambiguously predicts a subject's choice. Neuron 80:1532–1543. CrossRef Medline

Carrillo-Reid L, Miller JE, Hamm JP, Jackson J, Yuste R (2015) Endogenous sequential cortical activity evoked by visual stimuli. J Neurosci 35:8813–8828. CrossRef Medline

Chen JL, Carta S, Soldado-Magraner J, Schneider BL, Helmchen F (2013) Behaviour-dependent recruitment of long-range projection neurons in somatosensory cortex. Nature 499:336–340. CrossRef Medline

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. CrossRef Medline

Churchland MM, Yu BM, Cunningham JP, Sugrue LP, Cohen MR, Corrado GS, Newsome WT, Clark AM, Hosseini P, Scott BB, Bradley DC, Smith MA, Kohn A, Movshon JA, Armstrong KM, Moore T, Chang SW, Snyder LH, Lisberger SG, Priebe NJ, et al. (2010) Stimulus onset quenches neural variability: a widespread cortical phenomenon. Nat Neurosci 13:369–378. CrossRef Medline

Cohen MR, Maunsell JH (2009) Attention improves performance primarily by reducing interneuronal correlations. Nat Neurosci 12:1594–1600. CrossRef Medline

Cohen MR, Maunsell JH (2011) Using neuronal populations to study the mechanisms underlying spatial and feature attention. Neuron 70:1192–1204. CrossRef Medline

Girardeau G, Benchenane K, Wiener SI, Buzsáki G, Zugaro MB (2009) Selective suppression of hippocampal ripples impairs spatial memory. Nat Neurosci 12:1222–1223. CrossRef Medline

Glickfeld LL, Histed MH, Maunsell JH (2013) Mouse primary visual cortex is used to detect both orientation and contrast changes. J Neurosci 33:19416–19422. CrossRef Medline

Goltstein PM, Montijn JS, Pennartz CM (2015) Effects of isoflurane anesthesia on ensemble patterns of Ca^{2+} activity in mouse V1: reduced direction selectivity independent of increased correlations in cellular activity. PLoS One 10:e0118277. CrossRef Medline

Greenberg DS, Houweling AR, Kerr JN (2008) Population imaging of ongoing neuronal activity in the visual cortex of awake rats. Nat Neurosci 11:749–751. CrossRef Medline

Guizar-Sicairos M, Thurman ST, Fienup JR (2008) Efficient subpixel image registration algorithms. Opt Lett 33:156–158. CrossRef Medline

Gutnisky DA, Dragoi V (2008) Adaptive coding of visual information in neural populations. Nature 452:220–224. CrossRef Medline

Hansen BJ, Chelaru MI, Dragoi V (2012) Correlated variability in laminar cortical circuits. Neuron 76:590–602. CrossRef Medline

Harris KD (2005) Neural signatures of cell assembly organization. Nat Rev Neurosci 6:399–407. CrossRef Medline

Hebb DO (1949) The organization of behavior: a neuropsychological theory. Mahwah, NJ: Psychology.

Hubel DH, Wiesel TN (1959) Receptive fields of single neurones in the cat's striate cortex. J Physiol 148:574–591. CrossRef Medline

Ito M, Gilbert CD (1999) Attention modulates contextual influences in the primary visual cortex of alert monkeys. Neuron 22:593–604. CrossRef Medline

Jones LM, Fontanini A, Sadacca BF, Miller P, Katz DB (2007) Natural stimuli evoke dynamic sequences of states in sensory cortical ensembles. Proc Natl Acad Sci U S A 104:18772–18777. CrossRef Medline

Junek S, Kludt E, Wolf F, Schild D (2010) Olfactory coding with patterns of response latencies. Neuron 67:872–884. CrossRef Medline

Kerr JN, Greenberg D, Helmchen F (2005) Imaging input and output of neocortical networks in vivo. Proc Natl Acad Sci U S A 102:14063–14068. CrossRef Medline

König P, Engel AK, Roelfsema PR, Singer W (1995) How precise is neuronal synchronization? Neural Comput 7:469–485. CrossRef Medline

Lansink CS, Goltstein PM, Lankelma JV, McNaughton BL, Pennartz CM (2009) Hippocampus leads ventral striatum in replay of place-reward information. PLoS Biol 7:e100173. CrossRef Medline

Lee AK, Wilson MA (2002) Memory of sequential experience in the hippocampus during slow wave sleep. Neuron 36:1183–1194. CrossRef Medline

Luczak A, Barthó P, Marguet SL, Buzsáki G, Harris KD (2007) Sequential structure of neocortical spontaneous activity in vivo. Proc Natl Acad Sci U S A 104:347–352. CrossRef Medline

Luczak A, Bartho P, Harris KD (2013) Gating of sensory input by spontaneous cortical activity. J Neurosci 33:1684–1695. CrossRef Medline

Luczak A, McNaughton BL, Harris KD (2015) Packet-based communication in the cortex. Nat Rev Neurosci 16:745–755. CrossRef Medline

Luna R, Hernández A, Brody CD, Romo R (2005) Neural codes for perceptual discrimination in primary somatosensory cortex. Nat Neurosci 8:1210–1219. CrossRef Medline

Miller JE, Ayzenshtat I, Carrillo-Reid L, Yuste R (2014) Visual stimuli recruit intrinsically generated cortical ensembles. Proc Natl Acad Sci U S A 111:E4053–E4061. CrossRef Medline

Mitchell JF, Sundberg KA, Reynolds JH (2009) Spatial attention decorrelates intrinsic activity fluctuations in Macaque area V4. Neuron 63:879–888. CrossRef Medline

Montijn JS, Vinck M, Pennartz CM (2014) Population coding in mouse visual cortex: response reliability and dissociability of stimulus tuning and noise correlation. Front Comput Neurosci 8:58. CrossRef Medline

Montijn JS, Goltstein PM, Pennartz CM (2015) Mouse V1 population correlates of visual detection rely on heterogeneity within neuronal response patterns. Elife 4:e10163. CrossRef Medline

Nádasdy Z, Hirase H, Czurkó A, Csicsvari J, Buzsáki G (1999) Replay and time compression of recurring spike sequences in the hippocampus. J Neurosci 19:9497–9507. Medline

Nimmerjahn A, Kirchhoff F, Kerr JN, Helmchen F (2004) Sulforhodamine

101 as a specific marker of astroglia in the neocortex in vivo. Nat Methods 1:31–37. CrossRef Medline

Pelli DG (1997) The VideoToolbox software for visual psychophysics: transforming numbers into movies. Spat Vis 10:437–442. CrossRef Medline

Pennartz CM (2015) The brain's representational power: on consciousness and the integration of modalities. Cambridge, MA: Massachusetts Institute of Technology.

Petersen RS, Panzeri S, Diamond ME (2002) Population coding in somatosensory cortex. Curr Opin Neurobiol 12:441–447. CrossRef Medline

Romo R, Hernández A, Zainos A, Salinas E (2003) Correlated neuronal discharges that increase coding efficiency during perceptual discrimination. Neuron 38:649–657. CrossRef Medline

Rousseeuw PJ (1987) Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. J Comput Appl Math 20:53–65. CrossRef

Sadacca BF, Mukherjee N, Vladusich T, Li JX, Katz DB, Miller P (2016) The behavioral relevance of cortical neural ensemble responses emerges suddenly. J Neurosci 36:655–669. CrossRef Medline

Safaai H, von Heimendahl M, Sorando JM, Diamond ME, Maravall M (2013) Coordinated population activity underlying texture discrimination in rat barrel cortex. J Neurosci 33:5843–5855. CrossRef Medline

Schölvinck ML, Saleem AB, Benucci A, Harris KD, Carandini M (2015) Cortical state determines global variability and correlations in visual cortex. J Neurosci 35:170–178. CrossRef Medline

Stosiek C, Garaschuk O, Holthoff K, Konnerth A (2003) In vivo two-photon calcium imaging of neuronal networks. Proc Natl Acad Sci U S A 100: 7319–7324. CrossRef Medline

van der Togt C, Kalitzin S, Spekreijse H, Lamme VA, Supèr H (2006) Synchrony dynamics in monkey V1 predict success in visual detection. Cereb Cortex 16:136–148. CrossRef Medline

Ward JH (1963) Hierarchical grouping to optimize an objective function. J Am Statist Assoc 58:236–244. CrossRef

Wilson MA, McNaughton BL (1994) Reactivation of hippocampal ensemble memories during sleep. Science 265:676–679. CrossRef Medline

Xu S, Jiang W, Poo MM, Dan Y (2012) Activity recall in a visual cortical ensemble. Nat Neurosci 15:449–455. CrossRef Medline

Yang H, Kwon SE, Severson KS, O'Connor DH (2016) Origins of choice-related activity in mouse somatosensory cortex. Nat Neurosci 19:127–134. CrossRef Medline

Zandvakili A, Kohn A (2015) Coordinated neuronal activity enhances corticocortical communication. Neuron 87:827–839. CrossRef Medline