



Published in final edited form as:

Nat Neurosci. 2014 June ; 17(6): 851–857. doi:10.1038/nn.3707.

Population code in mouse V1 facilitates read-out of natural scenes through increased sparseness

Emmanouil Froudarakis¹, Philipp Berens^{1,2,3,*}, Alexander S. Ecker^{1,2,3,4}, R. James Cotton¹, Fabian H. Sinz^{2,5}, Dimitri Yatsenko¹, Peter Saggau^{1,6}, Matthias Bethge^{2,3,4}, and Andreas S. Tolias^{1,2,7,*}

¹Department of Neuroscience, Baylor College of Medicine, Houston, TX, USA

²Bernstein Center for Computational Neuroscience, Tübingen, Germany

³Werner-Reichardt-Center for Integrative Neuroscience and Institute for Theoretical Physics, University of Tübingen, Germany

⁴Max Planck Institute for Biological Cybernetics, Tübingen, Germany

⁵Institute for Neurobiology, Department for Neuroethology, University Tübingen, Germany

⁶Department of Bioengineering, Rice University, Houston, TX, USA

⁷Department of Computational and Applied Mathematics, Rice University, Houston, TX, USA

Abstract

The neural code is believed to have adapted to the statistical properties of the natural environment. However, the principles that govern the organization of ensemble activity in the visual cortex during natural visual input are unknown. We recorded populations of up to 500 neurons in the mouse primary visual cortex and characterized the structure of their activity, comparing responses to natural movies with those to control stimuli. We found that higher-order correlations in natural scenes induce a sparser code, in which information is encoded by reliable activation of a smaller set of neurons and can be read-out more easily. This computationally advantageous encoding for natural scenes was state-dependent and apparent only in anesthetized and active awake animals, but not during quiet wakefulness. Our results argue for a functional benefit of sparsification that could be a general principle governing the structure of the population activity throughout cortical microcircuits.

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use:http://www.nature.com/authors/editorial_policies/license.html#terms

*Corresponding authors: Andreas S. Tolias (astolias@bcm.edu), Philipp Berens (philipp.berens@uni-tuebingen.de).

Authors contributions

EF, PB, MB and AST designed the study; EF and PB created the stimuli; RJC, EF and PS built the two-photon setups; EF, RJC and DY wrote preprocessing code; EF performed the experiments; EF, PB and ASE analyzed the data; EF, ASE, FHS and MB performed the modeling; EF and PB wrote the paper with input from ASE, FHS, MB and AST. PB and AST supervised all stages of the project.

Introduction

Neurons in the early visual system are believed to have adapted to the statistical properties of the organism's natural environment¹⁻³. In particular, the response properties of neurons in the primary visual cortex are hypothesized to be optimized to provide a sparse representation of natural scenes: Theoretical work suggests that neural networks optimized for sparseness yield receptive fields similar to those observed in primary visual cortex (V1)³⁻⁵. In the resulting population code, only a small subset of the neurons should be active to encode each image and neural responses should be sparser for natural scenes than for stimuli from which the critical higher-order correlations have been removed⁶. These higher order correlations are reflected in the phase spectrum of a scene (as opposed to the amplitude spectrum) and drive the emergence of localized oriented bandpass filters resembling V1 receptive fields in sparse coding models⁴. In addition, they carry the perceptually relevant content of an image^{3,7}.

Indeed, single neurons in V1 respond highly selectively to image sequences as they occur during natural vision, showing high lifetime sparseness⁸⁻¹² (but see¹³). In addition, identical visual features activate complex cells more strongly when embedded in a natural scene compared to a noise stimulus without spatial structure¹⁴. However, sparseness in single neurons does not guarantee sparseness in a population¹⁰. For example, consider a local population of neurons tuned to diverse stimulus features becoming more selectively tuned to one particular stimulus due to learning. These neurons will exhibit high lifetime sparseness. However, since all neurons will become similarly tuned, population sparseness will decrease. Until recently, it has not been possible to record from sufficiently large and dense neuronal populations in order to empirically study the representation of natural scenes on the population level and measure the effect of natural stimulus statistics on properties such as population sparseness.

Despite the intriguing theoretical ideas and experimental advances¹⁵, it is thus still unclear whether and how the response properties of V1 neurons have adapted to the statistical regularities of natural scenes and whether this optimized representation has functional or computational benefits. Here, we use a novel high-speed 3D in vivo two-photon microscope¹⁶ to record the activity of nearly all of the hundreds of neurons in small volumes of the visual cortex of anesthetized and awake mice. The animals viewed natural movies and phase-scrambled movies. The latter were generated from natural movies by removing the higher-order correlation structure, resulting in two types of movies with identical power spectra but different phase spectra.

We found that higher-order correlations in the visual stimulus indeed change the structure of population activity in both anesthetized and awake, active animals: firing patterns are reorganized such as to facilitate decoding of individual movie scenes. In particular, we provide empirical evidence that sparse encoding of natural stimuli in neural populations leads to this improvement in read-out accuracy. This effect could be reproduced by a standard linear-nonlinear population model of V1 including a normalization stage. Interestingly, during quiet wakefulness, we did not observe the same reorganization, suggesting that visual processing of natural scenes in mouse V1 depends on brain state and

normalization mechanisms may improve the representation of natural scenes in certain brain states, such as when the animal is actively engaged with its sensory environment.

Results

To study how neural populations in V1 encode natural scenes we recorded stimuli matching the typical statistics encountered by these neurons. To this end, we mounted a miniature camera on the head of a mouse and recorded movies from its point of view, capturing body and head movement dynamics (but not eye movements; ‘natural movies’; Fig. 1a and Supplementary Movie 1). As we were specifically interested in how the spatio-temporal higher-order structure of natural movies impacts V1 population activity, we created a second stimulus set, in which we kept the first- and second-order statistics of the natural movies intact but randomized all higher-order correlations (‘phase-scrambled movies’; Fig. 1a and Supplementary Movie 2). These two movie types had the same spatio-temporal power spectrum, but different phase spectra¹⁷.

We then recorded the activity of neural ensembles in mouse V1 with up to 510 neurons using random-access 3D scanning¹⁶ and conventional 2D two-photon calcium imaging¹⁸ (Fig. 1b and Methods) while presenting multiple repetitions of natural and phase-scrambled movies. The recorded populations were dense in that we recorded most neurons in a small plane or cube of cortical tissue (Fig. 1b, expected 480–640 neurons in a volume of $200 \times 200 \times 100 \mu\text{m}$ in layer 2/3 for 3D scanning¹⁹). We verified that we were able to infer the spike rates of individual neurons from the fluorescence traces using simultaneous two-photon imaging and cell-attached recordings of action potentials (Fig. 1c, average linear correlation between inferred and true rate: $r = 0.79$, $n = 16$ cells). We imaged a total of 492 sites in 59 anesthetized and 100 sites in 9 awake animals mostly in layer 2/3, allowing us to characterize the representation of natural scenes across different cortical states (for details on the dataset, see Supplementary Table 1).

Neural populations responded well to both natural and phase-scrambled movies (Fig. 1d and Supplementary Fig. 1). Population receptive fields typically covered $25 \pm 8^\circ$ ($n=129$ sites; mean \pm SD) of visual angle (Fig. 2a, shown for 16 sites from one exemplary animal). The population response was highly reliable even across sites if the locations of their population receptive fields matched (Fig. 2b, c; $n = 120$ pairs of sites, linear correlation coefficient, $r = -0.77$; $p < 0.001$). We assessed which features of the natural stimulus were most important for driving the neural response. To this end, we predicted the average population response as well as single cell responses from the average brightness, the contrast, the kurtosis and the overall motion strength within the receptive field (Fig. 2d–g; see Methods for details). This model performed reasonably well in predicting population activity for most sites ($r^2 = 0.15$; $n = 21$ sites with at least four movies shown). By analyzing the weights of the model, we found that populations were driven most strongly by dark image regions for both classes of stimuli (Fig. 2g). Motion strength led to increases in population activity as well (Fig. 2g).

We systematically assessed how higher-order correlations in the stimulus affect V1 ensemble activity in the output layers 2/3 during anesthesia by comparing neural firing statistics between the two stimulus conditions in 500 ms windows (Fig. 3a, b;

Supplementary Fig. 2; $n = 315$ sites). Interestingly, the mean and the variance of the firing rate were hardly changed for natural compared to phase-scrambled movies (median change and 95% upper/lower CI, mean: 0.23%, $-0.37/0.74\%$; variance: 0%, $0/2.41\%$). In contrast, the signal correlation between pairs of neurons across time was higher during stimulation with natural movies (0.47 vs. 0.43 in natural vs. phase scrambled movies; median change 10.59%; $8.3/12.02\%$). Also, the neurons responded more reliably in time to the natural movies than to phase scrambled ones, as measured by the variance explained by the PSTH (0.24 vs. 0.18; 22.52%; $19.74/24.76\%$), and showed higher lifetime sparseness, responding more selectively to individual movie frames (0.35 vs. 0.28; 18.33%; $16.87/20.1\%$). In addition, population sparseness was higher for the natural movies than for the phase scrambled ones (0.26 vs. 0.21; 16.4%; $14.93/17.91\%$). This increase was present across almost all recorded sites (293/315 sites; Fig. 3c; two-sided binomial test, $p < 0.001$, $n = 315$). While the absolute level of sparseness depends on the bin width, population sparseness was increased for natural compared to phase-scrambled movies at all bin widths and population sizes (Supplementary Fig. 3). In addition, population sparseness correlated strongly with lifetime sparseness (Supplementary Fig. 4a; $r = 0.87$, $p < 0.001$), as did the differences in the two measures between natural and phase-scrambled movies (Supplementary Fig. 4b, linear correlation coefficient, $r = 0.61$, $p < 0.001$).

These differences in the population response caused by higher-order correlations in the visual stimulus were not restricted to the output layer 2/3 of V1: Reliability, lifetime sparseness and population sparseness also increased in the input layer 4 of V1 ($n = 46$ sites) and in the layer 2/3 of the secondary visual areas LM and AL ($n = 117$ sites; Fig. 3b). Interestingly, the average activity in areas LM and AL increased for natural stimuli. In addition, we verified that the observed effects were not related to the imaging method (2D or 3D imaging; all V1 data: $n = 332$ and $n = 30$ sites, respectively; Fig. 3d) and compared the results from calcium imaging to spikes, recording their spiking activity of single cells directly while presenting the movies ($n = 29$ cells). We found that all single-neuron measures agreed between spiking and calcium data (Fig. 3d; overlapping 95%-CIs). Finally, as our stimulus movies did not contain eye movements and eye movements contribute to shaping the response reliability of V1 neurons²⁰, we performed additional experiments using two virtual reality movies including rodent eye movements²¹. We found that most effects were compatible with those obtained from our original dataset ($n = 13$ sites; Supplementary Fig. 5; see Methods).

Next, we studied how the changes in the neural representation affected the encoding of natural and phase-scrambled movies. For this analysis, we pooled all anesthetized data from V1 ($n = 362$ sites). Specifically, we tested how well movie scenes could be discriminated based on the population activity in a 500-ms window. Using a simple linear classifier, we found that the discriminability was higher for scenes from natural movies than for scenes from phase-scrambled movies (0.25 vs. 0.19 bits; change: 24.22%; 95%-CI: 22.11/26.35%; Fig. 4a). To identify which of the differences in the population response best predicted this enhanced discriminability, we performed a multivariate linear regression analysis ($r^2 = 0.73$; see Methods). Population sparseness was the best predictor of classification performance, followed by reliability and lifetime sparseness (Fig. 4b). We verified that this result was largely insensitive to the time bin size, the decoding algorithm, the regression technique and

the sparseness measure (Supplementary Table 2). Reliability was also an influential factor in most of these control analyses, in particular for short time bins and without trial averaging.

We studied the link between population sparseness and discrimination performance in more detail by looking at the structure of the neural response space. using locally linear embedding²². This dimensionality reduction technique preserves local neighborhood relationships of the high-dimensional space while embedding the data points into a low-dimensional space (see Methods). We observed that the representations elicited by natural movies spread out more than those of phase-scrambled movies (Fig. 4c). We quantified this effect by measuring how far the population response vectors spread out in the neural space (measured by their spherical variance, see Methods). We found that neural responses to the natural stimuli were distributed more (0.13 vs. 0.10; 18.25%; 16.46/20.02%), and that this increased spread was highly correlated with increased population sparseness (Fig. 4d, $r = 0.98$, $p = 0.001$). Indeed, population sparseness is proportional to the average sine squared of the angle between the representations of different scenes and the main diagonal (Fig. 4e; see Methods). We studied whether this increased spread in the neural space is a straightforward consequence of differences between the two conditions in the stimulus space. While the distance between natural movie frames is somewhat larger than the distance between frames from phase scrambled movies (~ 10% measured in pixel space), a large distance between two movie frames does not lead to a large distance in the neural response space: both measures were only weakly correlated (Spearman's $r = 0.05$ and $r = 0.03$ for natural and phase scrambled stimuli, respectively; evaluated for pixels within the population receptive field). Thus, discriminability increases with population sparseness because of a nonlinear transformation through which the activity vectors in the sparser representation “fan out” in the neural space.

We next asked what ingredients a simple model of V1 needs to reproduce these results. We started with a model population consisting of linear-nonlinear Poisson (LNP) neurons, where the linear stage consists of filters learned by a sparse coding model²³ and the non-linearity was a static half-wave rectification function (Model 1, Fig. 5a). In this simple model, sparseness was higher for natural scenes compared to phase-scrambled ones, but, at the same time, the mean response was significantly lower (Fig. 5b and Supplementary Fig. 6a, d). As a consequence, phase-scrambled images could be classified better under this model, in contrast to what we observed in the neural data (Fig. 5B and Supplementary Fig. 6a,d). Thus, a simple LNP population with static non-linearity seems not compatible with our anesthetized data. We therefore added an adaptive non-linearity²⁴ to match the average responses between the two conditions (Fig. 5a, Model 2) or a normalization stage, where a “surround” pool of neurons divisively affects the activity of each neuron (Fig. 5a, Model 3). These models qualitatively reproduced the experimental results (Fig. 5b and Supplementary Fig. 6b–d). In summary, a non-linearity rapidly adapting to the statistics of the visual stimulus^{20,24} or divisive normalization, which is commonly employed in cortical computations²⁵, seems to be required for a sparse population representation with high discriminability of natural scenes.

To understand how neural responses to natural movies depend on the brain state, we repeated the same experiment in awake mice ($n = 9$ animals). We monitored eye

movements, running and whisking activity of the animals and separated the recordings into 'active' periods, where the animal was whisking and/or running ($n = 23$ sites; Supplementary Fig. 7; see Methods for inclusion criteria), and 'quiet wakefulness', when the animal was sitting still and not whisking ($n = 100$ sites; Supplementary Fig. 7). During active wakefulness higher-order correlations in the stimulus influenced the response statistics in a manner very similar to the anesthetized state (Fig. 6a, b; overlapping 95%-confidence intervals for all measures between both states). In particular, discriminability for natural scenes was improved over phase-scrambled scenes in many sites (Fig. 6b, c; 19/27 sites, two-sided binomial test, $P < 0.05$) and differences in population sparseness were an important predictor of differences in discriminability (Fig. 6d and Supplementary Table 2). During quiet wakefulness, however, the mean activity and thus the discriminability were lower for natural movies than for phase-scrambled ones (Fig. 6a–c). Also, the differences in mean activity were the best predictor for differences in discriminability (Fig. 6d and Supplementary Table 2).

Discussion

We showed that higher-order correlations in natural movies change the structure of the firing patterns in V1 during anesthesia and during active wakefulness without changing the overall activation level, in particular increasing the population sparseness of the neuronal representation. This increased sparseness leads to an improvement of the discriminability of different movie scenes, arguing for a computational benefit of population sparseness. During quiet wakefulness, the mean activity was lower for natural movies than for phase-scrambled ones and discriminability was not improved. Together with our modeling work, these results indicate that processing of natural scenes in mouse V1 depends on cortical state, with normalization mechanisms improving the representation of natural scenes when the animal is actively engaged with its sensory environment.

Relationship to previous studies

Our results agree with previous single unit studies of natural scenes showing that responses of single neurons in V1 of cats and monkeys show high lifetime sparseness and reliability^{8–12}. In our experiments, the movies covered the receptive fields and its surround (Fig. 2a). The surround circuitry likely contributes to the computations underlying the high sparseness and discriminability levels observed during stimulation with natural movies^{8,26}. We found that dark image regions lead to highest population activity, in agreement with the dark dominance reported in layers 2/3 of macaque V1²⁷. Our results are also compatible with an imaging study recording neural populations of about 80 neurons in mouse V1 showing two natural and two highly artificial movies¹⁵. It found that ~25% of the neurons are reliably activated by each stimulus and that stimulus identity could be read out with high accuracy. However, its experimental design did not allow the authors to relate their findings to specific statistical properties of the natural stimuli. Finally, our results agree with a recent study which found no systematic difference in firing rates between phase-scrambled and natural movies of neurons in V1 but did find a difference in V2 of anesthetized monkeys¹⁷, as present in our data (Fig. 3b). This suggests that there are specific changes in the

organization of the population activity in V1, such as increased population sparseness, that could lead to neurons in V2 having higher firing rates for natural scenes.

State dependence of neural processing in mouse V1

We found that the population representation of natural movies was highly state-dependent in awake animals. This seems to be a general feature of visual processing in the primary visual cortex of mice. For example, firing rates are higher during running than during quiet wakefulness^{28,29} when stimulated with oriented gratings. In addition, activity during running resembles activity under anesthesia²⁸, in agreement with our data (Fig. 6). In contrast, neurons appear “visually detached” during quiet wakefulness²⁸ and visual performance is lower³⁰. Interestingly, the origins of this state dependence seem to lie in the cortex, as responses in LGN do not differ strongly between active and quiet wakefulness²⁸. In the context of our models, it seems like a normalization stage as commonly found in cortical circuits acts to balance the mean activity between the movie types during anesthesia and active wakefulness. During quiet wakefulness, in contrast, the structure of the population response agreed better with a purely static linear-nonlinear model of V1.

Implications for models of V1

Most theoretical studies relating natural image statistics to response properties of V1 have modeled the neurons as linear rate-based units without spike generation mechanisms^{4,5}. Recent theoretical studies highlight the role of normalization for removing contrast-correlations^{31,32}, which are one of the main sources of higher-order correlations in natural images. This view is supported by our data, where a normalization step is needed to qualitatively reproduce the finding that the average activity is closely matched between natural and phase scrambled movies, yet population sparseness and discriminability are increased for natural movies. In hierarchical sparse coding networks, such nonlinearities could be used to make higher-order correlations explicit by transferring them into second-order correlations. This process would make the information carried by these higher-order correlations easier to read-out by subsequent processing stages. Note, however, that our models are highly simplified versions of cortical processing and the precise effect size for the models depends to some extent on the parameter settings. Therefore, we refrained from more detailed quantitative comparisons. To obtain a more precise link between sparse coding models and neural data, it may be useful to study models with a higher degree of biological accuracy including spiking neurons and local learning³³.

In addition, our data does not provide evidence for sparse coding per se: our measurements do not allow us to determine the absolute level of sparseness due to its dependence on the size of the time bins and two-photon imaging parameters. In addition, we and others¹⁵ observed fairly high levels of signal correlations in our data. Whether these signal correlations reflect the true response similarity of the neurons is not clear since measuring signal correlations with two-photon imaging is quite sensitive to correlated imaging noise e.g. from cardiac or respiratory cycles¹⁶. Some authors have tried to remove this correlated noise by projecting out the first principle component of the population response and obtained lower estimates of signal correlation³⁴. To what degree this procedure removes correlated signal as well is an open question. Another study found lower correlations using

highly decorrelated receptive field mapping stimuli³⁵. If our measurements reflect the signal correlations generated by the neural circuit under natural stimulation conditions, they speak against conventional interpretation of sparse coding models.

While our study therefore does not rule out the possibility that the population code in V1 is relatively dense, it establishes, however, that a relative increase in population sparseness greatly increases the discriminability of natural scenes by ‘fanning out’ their representation in neural space. This relationship is not a straightforward consequence of mathematical identities: beyond a certain point, increasing sparseness will impair classification performance^{36,37}. Theoretical studies have argued for different functional advantages of sparse coding³⁸, for instance that they enable efficient memory storage, are energy-efficient or make the structure of the sensory input signal explicit. Interestingly, sparseness seems an important objective for learning image representation useful for image discrimination, but enforcing sparseness during the encoding of images does not seem advantageous³⁷. How the increase in discriminability and its relationship to sparseness we observed is mathematically related to the diverse theoretical proposals will therefore be an important direction for future research.

Methods

Animal preparation

We used both male and female mice (62 and 6 mice respectively, age: p22 to p185) of either the *C57CL/6J* wild type strain (54 mice) or heterozygous for either the *SST-Cre*, *PV-Cre* or *Viaat-Cre* genes and heterozygous for the *Ai9* reporter gene (5, 5 and 4 mice respectively). Mice were initially anesthetized with Isoflurane (3%) and for the anesthetized experiments anesthesia was maintained by either Isoflurane (2%, 4 mice) or a mixture of Fentanyl (0.05 mg/kg), Midazolam (5 mg/kg), and Medetomidin (0.5 mg/kg) with anesthesia boosts consisting half of the initial dose every three hours (48 mice). The temperature of the mouse was maintained between 36.5°C and 37.5°C using a homeothermic blanket system (Harvard Instruments). In some experiments we applied eye oil ointment (polydimethylsiloxane) to prevent dehydration of the cornea. In the surgical procedure part of the skin over the skull was resected and the location of the craniotomy determined (approximately 2.5–2.9 mm lateral to the mid-line sagittal suture and anterior to the Lambda suture). A head-post bar was secured to the skull using bone cement and then a 0.5–1.5 mm craniotomy was performed over the region of interest. The mouse was transferred to either a traditional galvanometric imaging system with a 10x objective (Nikon CFI Plan Apo 10×), or a custom epi-fluorescence stereoscope, for guided injections of the membrane-permeable calcium indicator Oregon green 488 BAPTA-1 AM (OGB-1, Invitrogen) and in some experiments the astrocyte specific fluorescent dye sulforhodamine 101 (SR-101, Invitrogen). Dye solution was prepared as previously described¹⁸ and bolus loading injections were performed by using a continuous-pulse low pressure protocol with a glass micropipette and a computer controlled pressure system (Picospritzer II, Parker) to inject ~ 300 μm below the surface of the cortex. The window was then sealed using a glass coverslip secured with dental cement. After allowing 1 h for dye uptake, each injection resulted in a stained area of < 400 μm in diameter. For details on the awake experiments, see below.

All procedures performed on mice were carried out in accordance with the ethical guidelines of the National Institutes of Health and approved by the institutional review board (IACUC Baylor College of Medicine).

Intrinsic optical imaging and visual area identification

We used stereotactic information to locate the primary visual cortex of the mouse (V1). In some experiments we additionally used intrinsic imaging as previously described³⁹. We measured the change in cortical light reflectance with illumination wavelength of 610 nm. 512×512 pixel images were captured at the rate of 4 Hz with a CMOS camera (MV1-D1312-160-CL, PhotonFocus). For retinotopic mapping we stimulated with upward and rightward drifting white bars on black background (Frequency: 0.16 cycles/deg, Width: 2 deg). Images were analyzed by a custom-written code in MATLAB. In brief, we extracted the phases of the fundamental Fourier component of each pixel. We then constructed the 2D phase maps for the two directions and the retinotopic maps were constructed as previously described³⁹. We used published retinotopic maps⁴⁰ to identify the location of the three areas V1, LM and AL.

Two-photon imaging and electrophysiology

We recorded calcium traces using a custom-built two-photon microscope equipped with a Chameleon Ti-Sapphire laser (Coherent) tuned at 800 nm and a 20x, 1.0 NA Olympus objective. Scanning was controlled by either a set of galvanometric mirrors (2D imaging) or a custom-built acousto-optic deflector system (AODs; 3D imaging)¹⁶. The average power output of the objective was kept at less than 50 mW for 2D imaging and 120 mW for 3D imaging for the L2/3 experiments and less than 100mW for the L4 experiments. Calcium activity was typically sampled at ~ 12 Hz with the galvanometric mirrors and at ~ 260 Hz (78 – 450 Hz) with the AODs. We recorded data from depths of 100–540 μm below the cortical surface. The border between L2/3 and L4 was assumed to be at 350 μm ⁴¹.

To perform simultaneous loose-patch and two-photon calcium imaging recordings we replaced sulforhodamine 101 with Alexa Fluor 594 (Invitrogen) and used glass pipettes with 5–7 M Ω resistance for targeted two-photon-guided loose-cell patching as previously described⁴².

Awake experiments

Animal preparation was done as for anesthetized experiments. After the surgery the animal was removed from the anesthesia and was placed head-fixed on a spherical treadmill that allowed only single axis movements (forwards-backwards). We started the experiments once the animal was awake for at least 30 min and comfortably running on the ball. During the whole duration of the experiment the animal was monitored for any signs of discomfort and few drops of sugary water were delivered every ~ 30 min. All the awake recordings were done with the 2D imaging system. Whisking and eye movements were simultaneously monitored with one camera (CR-GM00-H6400, Teledyne Dalsa) and a high-dpi optical mouse (Deathadder, Razer) placed close to the ball detected the mouse movements. All the behavioral traces were recorded using a custom-written software in LabView (National Instruments).

The pupil was detected using circular Hough transform (`imfindcircles`, MATLAB) and its location was extracted for each recorded frame. We analyzed only periods where the eye did not move significantly (< 0.15 mm of linear pupil displacement; Supplementary Fig. 7) to keep the trial-to-trial responses similar. We classified the trials into quiet and active states using whisker and animal movements (Supplementary Fig. 7). To detect the whisker movement we first extracted the optic flow (OpticalFlow Object, Computer Vision System Toolbox, MATLAB) and then computed the average magnitude of velocity for all the pixels within each frame. An active period was defined as a period where the animal was either whisking (> 1.25 $\mu\text{m/s}$), running (> 24 mm/s) or both. A quiet state period was defined as a period where the animal was not whisking (< 1 $\mu\text{m/s}$) or running (< 18 mm/s). If the running and whisking activity was between the upper and lower boundaries, the period was discarded. Only sessions that had 10 or more trials with either whisking or running periods were included in the analysis.

Stimulus preparation and visual stimulation

Natural movies were acquired by a modified web camera (LifeCam VX-6000, Microsoft) mounted on the head of a mouse. The mouse was left to run freely in an enriched and spacious (~ 0.5 m²) environment. We used parts of the movie where the mouse was active, which contained periods of moving, looking around and resting. These movies thus included head and body movements but not eye movements of the mice. We converted the movies to grayscale and from those created new movies without higher-order correlations (phase scrambled movies) by randomizing the original natural phase spectrum using a 3D FFT, while preserving its complex conjugate symmetry. The randomization across the 3D space ensured that both spatial and temporal 1st and 2nd order statistics were equal between the two classes of stimuli, whereas both the spatial and temporal higher order correlations were randomized. Preserving the temporal 1st and 2nd order statistics led to the temporal smoothness necessary for the perception of stimulus continuity.

The natural movies used in this study did not include any compensatory eye movements, and thus might include motion that the animal would not experience in natural viewing conditions. Since we could not monitor the movements of the eye during video acquisition, we used a rendered virtual reality movie from Wallace et al.²¹ as a control analysis. This movie takes into account the eye movements of a rat. Additionally, to alleviate the effect of the artificial properties of the rendered movie, we created three “naturalistic” movies by extracting the 3D motion vectors of the virtual reality movie (`imregister`, MATLAB) and applied it to three different natural images.

Stimulus presentation routines were written in MATLAB using the Psychophysics Toolbox⁴³ and the experiment was controlled by a custom state system written in LabView. Multiple segments of the movies (10 s, 30 fps) were presented on either a 20" LCD monitor (2007WFPb, Dell, 60Hz refresh rate) or a 30" LCD monitor (U3011t, Dell, 60Hz refresh rate) with their default gamma curves, positioned ~ 30 cm away from the animal's eye, covering between 71 and 94 degrees in the azimuth and between 48 and 67 degrees in the elevation, respectively. We presented at least two different movie segments in each session for the anesthetized experiments and at least one movie for the awake experiments. Each

unique movie segment was presented monocularly between 6 and 30 times for the anesthetized experiments and between 10 and 28 times for the awake experiments.

Preprocessing

For data from the galvanometric scanning system, we performed motion compensation across the horizontal plane. Cell detection was done by a fully automated custom-written code in MATLAB that detects circular blobs within a specific parameter set of minimum diameter, cell contrast, edge sharpness and maximum number of cells per site. For data from the AOD system we used online motion detection¹⁶ and sessions with excessive motion (session with frequent periods where movement exceeded 5 μ m) were aborted. The data processing chain for this analysis was built using the DataJoint library for MATLAB (<http://datajoint.github.com/datajoint-matlab/>). Individual cell traces were computed as the average of the cell ROI. For data from the AOD system individual cell traces were computed from a single point in the center of the cell. Calcium signals were normalized to $\Delta F/F$, down-sampled to 10–20 Hz and then filtered by using a non-negative deconvolution algorithm⁴⁴. We assessed the quality of the reconstruction as previously described¹⁶. For this analysis we only included cells that were recording with both 2D and 3D imaging systems ($n = 16$ neurons). Reconstruction of the spike rate is generally very good (Fig 1c; 500 ms time bin; $R=0.79$, averaged across both systems).

Receptive field mapping and population response prediction

In some experiments we mapped the location and size of the receptive fields of the neurons using black and white squares that each covered $\sim 6^\circ$ of the visual field. The squares were presented across the entire range of the monitor in random order for 250 ms each. To map the receptive fields of individual neurons we averaged the first 500 ms of the activity of a cell across all repetitions of the stimulus for each location. We fit the resulting 2D map using an elliptic 2D Gaussian. For each neuron we computed the SNR as the ratio of the variance of this image within three SD of the receptive field center to the variance of the image outside of the three SD of the receptive field center. Well-defined receptive fields with $\text{SNR} > 1$ had a median size of $11 \pm 7^\circ$ ($n=5193$ neurons) and their locations covered a standard deviation of 10° across the visual field of each recording session. We computed population receptive fields by averaging the receptive fields of the neurons with $\text{SNR} > 1$. We again fit a 2D Gaussian and took the two-SD ellipse as the population receptive field to ensure we capture the almost all of the area that the neurons respond to.

To identify which features of the natural stimulus were most important for driving the neural responses, we computed four image statistics for each frame of the movie within the population receptive field and across the whole image: (1) Average pixel intensity, (2) Standard deviation of the pixel intensities (contrast), (3) Kurtosis of the pixel intensities, (4) Velocity of the global motion vector (OpticalFlow Object, Computer Vision System Toolbox, MATLAB). We used a linear regression model and leave-one-out cross-validation to estimate the weights of these features on both the individual cell responses and the population responses. For this analysis we only included sessions with four or more unique natural movies ($n = 21$ sites).

Measures of population activity

(1) Mean was measured across the entire presentation of a movie and then averaged across all neurons from a given site. (2) Variance was measured across the trials and then averaged across all neurons from a given site. (3) Signal correlations were measured as the Pearson correlation between the average response traces. (4) The reliability of each cell was defined as the variance of the average response across trials divided by the total variance across all responses. (5) Sparseness was computed using the definition of Vinje and Gallant⁸

$$s = \left(1 - \frac{1}{N} \frac{(\sum_i r_i)^2}{\sum_i r_i^2} \right) / \left(1 - \frac{1}{N} \right)$$

For lifetime sparseness, r_i is the response to the i^{th} frame of a movie averaged across trials and N is the number of movie frames. Lifetime sparseness was highly correlated with the peak rate of the neurons across sites ($r=0.82$). For population sparseness, r_i is the response of the i^{th} neuron in a site to a certain movie frame averaged across trials, and N is the number of neurons. This measure of population sparseness is sensitive to the binning size but not the population size (Supplementary Fig 1). Population sparseness was averaged across movie frames and lifetime sparseness was averaged across neurons. Spherical variance was measured as

$$SV = 1 - \left\| \frac{1}{N} \sum_i \frac{\vec{r}_i}{\|\vec{r}_i\|} \right\|,$$

where \vec{r}_i is the population response vector in response to movie frame i , averaged across trials. For all measures, the statistics were done over sites after averaging within each site and we report the median and 95% bootstrap confidence intervals with 5000 resamples unless otherwise stated.

Decoding and discriminability

We used the optimal linear classifier assuming isotropic covariance matrices and leave-one-out cross-validation to estimate the decoding error between the neural representations of pairs of 500 ms scenes^{45,46}. Each scene was represented as an N -dimensional vector of neural activity for each trial. We converted the decoding error to discriminability, the mutual information (measured in bits) between the true class label c and its estimate \hat{c} , by computing

$$MI(c, \hat{c}) = \sum_{i \in 1,2} \sum_{j \in 1,2} p_{ij} \log_2 \frac{p_{ij}}{p_i \cdot p_j},$$

where p_{ij} is the probability of observing true class i and predicted class j and p_i and p_j denote the respective marginal probabilities.

Regression analysis

We used a linear regression model to predict the difference in classification performance between natural and phase-scrambled scenes across sites from the differences in mean, variance, signal correlations, reliability, lifetime sparseness and population sparseness. Before fitting, the regressors were standardized. We show the total contribution towards explaining the average difference in decoding performance, i.e. the product between the estimated coefficient and the average of the factor difference $b_i x_i$, normalized to a total effect size of 1. Error bars shown are the standard errors of the coefficients in the linear regression model scaled by the mean difference x .

Locally linear embedding

In order to graphically illustrate the dispersion of the neural responses to the two classes of stimuli, we used local linear embedding to embed the high-dimensional neural representations of 500 ms scenes into a low-dimensional subspace²². This technique takes into account local neighborhood relationships in the original space and tries to preserve them in the low-dimensional space. In brief, the first step is to identify the k nearest neighbors in the full activity space, the second is to reconstruct the neighbor relationships with some linear weights by taking into account the local covariance structure and the third step consists of computing the embedding coordinates in the reduced space by using the weights from the previous step. We first averaged the responses across neurons and then convolved the data with a one second Gaussian window in time to reduce noise. To ensure the reduction of space is equal for the responses to the two classes of stimuli, we performed the local linear embedding simultaneously for the response vectors to both natural and phase-scrambled movies. As parameters for the algorithm, we used 12 neighbors and obtained 3 output dimensions. For visualization, we sampled the embedded trace every 50 ms using cubic spline interpolation.

V1 model

For the V1 model, we used the filter responses to a set of 143 localized and oriented filters obtained with Independent Component Analysis (ICA) on 12×12 patches randomly sampled from natural images from the van Hateren database⁴⁷ using techniques previously described²³ (143 because the DC component filter was excluded). A copy of the filter responses was sign-flipped and combined with the original responses. This duplication was done to preserve the full information about the image, as the resulting responses of this population of 286 neurons were half-wave rectified (Model 1). For the adaptive model, the slope of the linear part was adjusted to match the means between phase-scrambled and natural image patches (Model 2). For the divisive normalization model (Model 3), each

rectified filter response x_i was transformed by $f(x_i) = \frac{x_i^2}{\sigma + \sum x_j^2}$. The constant σ was set to the 0.5% percentile of the raw filter responses to the natural image patches. All the resulting responses were scaled to obtain discrimination performances matching the in-vivo data. We used these responses as rates to draw spike counts from a Poisson distribution (20 trials). To obtain a similar processing procedure with the in vivo data, the simulated data was split into 200 sets of responses to 100 unique image patches and the averages for all the measures were included in the analysis (200 data points for each measure).

Derivation of relation between sparseness and angle

$$\begin{aligned}\theta &= \arccos\left(\frac{x \cdot d}{\|x\| \|d\|}\right) \\ \theta &= \arccos\left(\frac{\sum_i x_i}{\sqrt{\sum_i x_i^2} \sqrt{n}}\right) \\ \cos^2\theta &= \frac{(\sum_i x_i)^2}{n \sum_i x_i^2} \\ S &= 1 - \frac{(\sum_i x_i)^2}{n \sum_i x_i^2} \\ S &= 1 - \cos^2\theta \\ S &= \sin^2\theta\end{aligned}$$

θ : Angle from the diagonal

d: Diagonal

n: Number of neurons

x: Population Response

S: Population Sparseness

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank W. J. Ma for discussions, P. Storer for help with setting up the two-photon system, D. Murray, A. Laudano and A. M. Morgan for organizational help and V. Kalatsky for help with the intrinsic-imaging system.

This work was supported by grants NIH-Pioneer award DP1-OD008301 & NEI DP1EY023176 to A.S.T., NEI P30-EY002520, NIDA R01DA028525, the McKnight Scholar Award and Beckman Foundation Young Investigator Award to A.S.T; the German National Academic Foundation (P. B.), the Bernstein Center for Computational Neuroscience (FKZ 01GQ1002); the German Excellency Initiative through the Centre for Integrative Neuroscience Tübingen (EXC307).

References

1. Attneave F. Some informational aspects of visual perception. *Psychol Rev.* 1954; 61:183. [PubMed: 13167245]
2. Barlow HB. The coding of sensory messages. *Curr Probl Anim Behav.* 1961:331–360.
3. Simoncelli EP, Olshausen BA. Natural image statistics and neural representation. *Annu Rev Neurosci.* 2001; 24:1193–1216. [PubMed: 11520932]
4. Olshausen BA, Field DJ. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature.* 1996; 381:607–609. [PubMed: 8637596]
5. Bell AJ, Sejnowski TJ. The ‘independent components’ of natural scenes are edge filters. *Vision Res.* 1997; 37:3327–3338. [PubMed: 9425547]
6. Field DJ. What Is the Goal of Sensory Coding? *Neural Comput.* 1994; 6:559–601.
7. Gerhard HE, Wichmann FA, Bethge M. How Sensitive Is the Human Visual System to the Local Statistics of Natural Images? *PLoS Comput Biol.* 2013; 9:e1002873. [PubMed: 23358106]
8. Vinje WE, Gallant JL. Sparse Coding and Decorrelation in Primary Visual Cortex During Natural Vision. *Science.* 2000; 287:1273–1276. [PubMed: 10678835]
9. Yen SC, Baker J, Gray CM. Heterogeneity in the Responses of Adjacent Neurons to Natural Stimuli in Cat Striate Cortex. *J Neurophysiol.* 2007; 97:1326–1341. [PubMed: 17079343]

10. Willmore BDB, Mazer JA, Gallant JL. Sparse Coding in Striate and Extrastriate Visual Cortex. *J Neurophysiol.* 2011; 105:2907–2919. [PubMed: 21471391]
11. Weliky M, Fiser J, Hunt RH, Wagner DN. Coding of Natural Scenes in Primary Visual Cortex. *Neuron.* 2003; 37:703–718. [PubMed: 12597866]
12. Yao H, Shi L, Han F, Gao H, Dan Y. Rapid learning in cortical coding of visual scenes. *Nat Neurosci.* 2007; 10:772–778. [PubMed: 17468750]
13. Tolhurst DJ, Smyth D, Thompson ID. The Sparseness of Neuronal Responses in Ferret Primary Visual Cortex. *J Neurosci.* 2009; 29:2355–2370. [PubMed: 19244512]
14. Felsen G, Touryan J, Han F, Dan Y. Cortical Sensitivity to Visual Features in Natural Scenes. *PLoS Biol.* 2005; 3:e342. [PubMed: 16171408]
15. Kampa BM, Roth MM, Gobel W, Helmchen F. Representation of visual scenes by local neuronal populations in layer 2/3 of mouse visual cortex. *Front Neural Circuits.* 2011; 5
16. Cotton RJ, Froudarakis E, Storer P, Saggau P, Tolias AS. Three-dimensional mapping of microcircuit correlation structure. *Front Neural Circuits.* :7.
17. Freeman J, Ziemba CM, Heeger DJ, Simoncelli EP, Movshon JA. A functional and perceptual signature of the second visual area in primates. *Nat Neurosci.* 2013; 16:974–981. [PubMed: 23685719]
18. Garaschuk O, Milos RI, Konnerth A. Targeted bulk-loading of fluorescent indicators for two-photon brain imaging in vivo. *Nat Protoc.* 2006; 1:380–386. [PubMed: 17406260]
19. Schüz A, Palm G. Density of neurons and synapses in the cerebral cortex of the mouse. *J Comp Neurol.* 1989; 286:442–455. [PubMed: 2778101]
20. Levy M, Marre O, Monier C, Pananceau M, Frégnac Y. Animation of natural scene by virtual eye-movements evokes high precision and low noise in V1 neurons. *Front Neural Circuits.* 2013; 7:206. [PubMed: 24409121]
21. Wallace DJ, et al. Rats maintain an overhead binocular field at the expense of constant fusion. *Nature.* 2013; 498:65–69. [PubMed: 23708965]
22. Roweis ST, Saul LK. Nonlinear Dimensionality Reduction by Locally Linear Embedding. *Science.* 2000; 290:2323–2326. [PubMed: 11125150]
23. Eichhorn J, Sinz F, Bethge M. Natural Image Coding in V1: How Much Use Is Orientation Selectivity? *PLoS Comput Biol.* 2009; 5:e1000336. [PubMed: 19343216]
24. Hosoya T, Baccus SA, Meister M. Dynamic predictive coding by the retina. *Nature.* 2005; 436:71–77. [PubMed: 16001064]
25. Carandini M, Heeger DJ. Normalization as a canonical neural computation. *Nat Rev Neurosci.* 2012; 13:51–62. [PubMed: 22108672]
26. Haider B, et al. Synaptic and Network Mechanisms of Sparse and Reliable Visual Cortical Activity during Nonclassical Receptive Field Stimulation. *Neuron.* 2010; 65:107–121. [PubMed: 20152117]
27. Yeh CI, Xing D, Shapley RM. ‘Black’ Responses Dominate Macaque Primary Visual Cortex V1. *J Neurosci.* 2009; 29:11753–11760. [PubMed: 19776262]
28. Niell CM, Stryker MP. Modulation of Visual Responses by Behavioral State in Mouse Visual Cortex. *Neuron.* 2010; 65:472–479. [PubMed: 20188652]
29. Polack PO, Friedman J, Golshani P. Cellular mechanisms of brain state-dependent gain modulation in visual cortex. *Nat Neurosci.* 2013; 16:1331–1339. [PubMed: 23872595]
30. Bennett C, Arroyo S, Hestrin S. Subthreshold Mechanisms Underlying State-Dependent Modulation of Visual Responses. *Neuron.* 2013; 80:350–357. [PubMed: 24139040]
31. Sinz F, Bethge M. Temporal Adaptation Enhances Efficient Contrast Gain Control on Natural Images. *PLoS Comput Biol.* 2013; 9:e1002889. [PubMed: 23382664]
32. Schwartz O, Simoncelli EP. Natural signal statistics and sensory gain control. *Nat Neurosci.* 2001; 4:819–825. [PubMed: 11477428]
33. Zylberberg J, Murphy JT, DeWeese MR. A Sparse Coding Model with Synaptically Local Plasticity and Spiking Neurons Can Account for the Diverse Shapes of V1 Simple Cell Receptive Fields. *PLoS Comput Biol.* 2011; 7:e1002250. [PubMed: 22046123]

34. Bonin V, Histed MH, Yurgenson S, Reid RC. Local Diversity and Fine-Scale Organization of Receptive Fields in Mouse Visual Cortex. *J Neurosci.* 2011; 31:18506–18521. [PubMed: 22171051]
35. Smith SL, Hausser M. Parallel processing of visual space by neighboring neurons in mouse visual cortex. *Nat Neurosci.* 2010; 13:1144–1149. [PubMed: 20711183]
36. Barak O, Rigotti M, Fusi S. The Sparseness of Mixed Selectivity Neurons Controls the Generalization–Discrimination Trade-Off. *J Neurosci.* 2013; 33:3844–3856. [PubMed: 23447596]
37. Rigamonti, R.; Brown, MA.; Lepetit, V. Are sparse representations really relevant for image classification?. *Comput. Vis. Pattern Recognit. CVPR 2011 IEEE Conf. On; IEEE; 2011.* p. 1545-1552. at <http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=5995313>
38. Olshausen BA, Field DJ. Sparse coding of sensory inputs. *Curr Opin Neurobiol.* 2004; 14:481–487. [PubMed: 15321069]
39. Kalatsky VA, Stryker MP. New Paradigm for Optical Imaging: Temporally Encoded Maps of Intrinsic Signal. *Neuron.* 2003; 38:529–545. [PubMed: 12765606]
40. Wang Q, Burkhalter A. Area map of mouse visual cortex. *J Comp Neurol.* 2007; 502:339–357. [PubMed: 17366604]
41. Olsen SR, Bortone DS, Adesnik H, Scanziani M. Gain control by layer six in cortical circuits of vision. *Nature.* 2012; 483:47–52. [PubMed: 22367547]
42. Komai S, Denk W, Osten P, Brecht M, Margrie TW. Two-photon targeted patching (TPTP) in vivo. *Nat Protoc.* 2006; 1:647–652. [PubMed: 17406293]
43. Pelli DG. The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat Vis.* 1997; 10:437–442. [PubMed: 9176953]
44. Vogelstein JT, et al. Fast Nonnegative Deconvolution for Spike Train Inference From Population Calcium Imaging. *J Neurophysiol.* 2010; 104:3691–3704. [PubMed: 20554834]
45. Duda, RO.; Hart, PE.; Stork, DG. *Pattern Classification: Pattern Classification.* John Wiley & Sons; 2000.
46. Berens P, et al. A fast and simple population code for orientation in primate V1. *J Neurosci.* 2012
47. Hateren van JH, van der Schaaf A. Independent component filters of natural images compared with simple cells in primary visual cortex. *Proc R Soc Lond B Biol Sci.* 1998; 265:359–366.

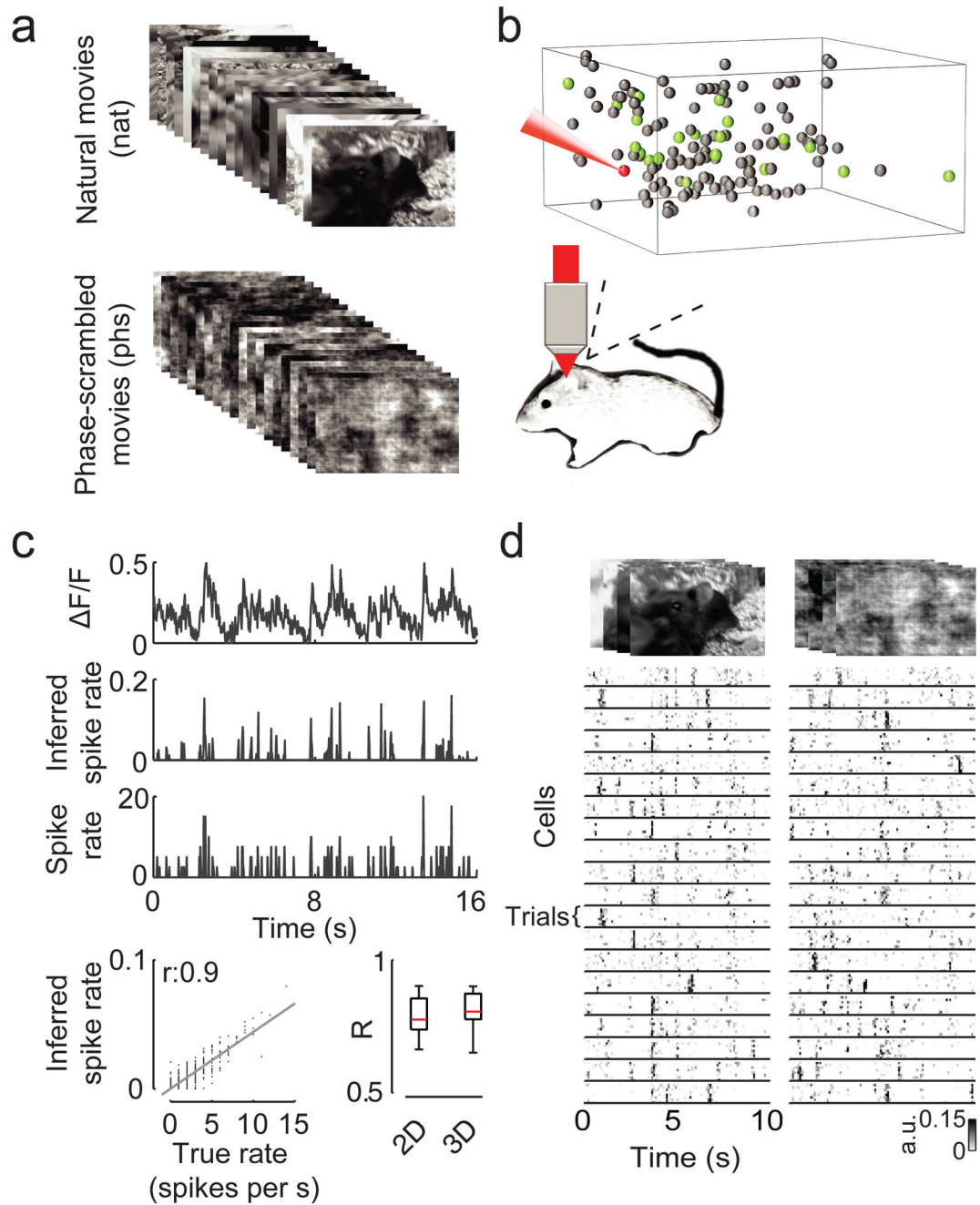


Figure 1. Experimental setup and example session

A. Illustration of the stimuli. Top, selected frames from a natural movie recorded from the viewpoint of a mouse using a MouseCam. Bottom, selected frames of the phase-scrambled movie, obtained by removing the higher-order correlations from the natural movie. First- and second-order correlations are matched between both stimulus types.

B. Example session of 3D AOD two-photon recordings with 139 neurons. Grey balls indicate the position of the cell bodies. The activity of green neurons is shown in D. The

spikes of the red neuron were recorded simultaneously using a patch pipette and are shown in C.

C. Reconstruction quality of the spike rate from Calcium signals. From top to bottom: dF/F , spike rate inferred using nonnegative deconvolution, measured spike rate using patching and a scatter plot of true vs. inferred rate with linear regression line (grey) for the red example neuron indicated in B. Bottom right: box plot of correlation of true vs. inferred rate for the same sample of 16 neurons imaged with 2D or 3D imaging (red line: median; black box: 25th and 75th percentile; whiskers: range).

D. Responses of the population of neurons shown in B (green neurons) for both stimulus types. Shown are responses of 20/139 cells, and 8 trials each.

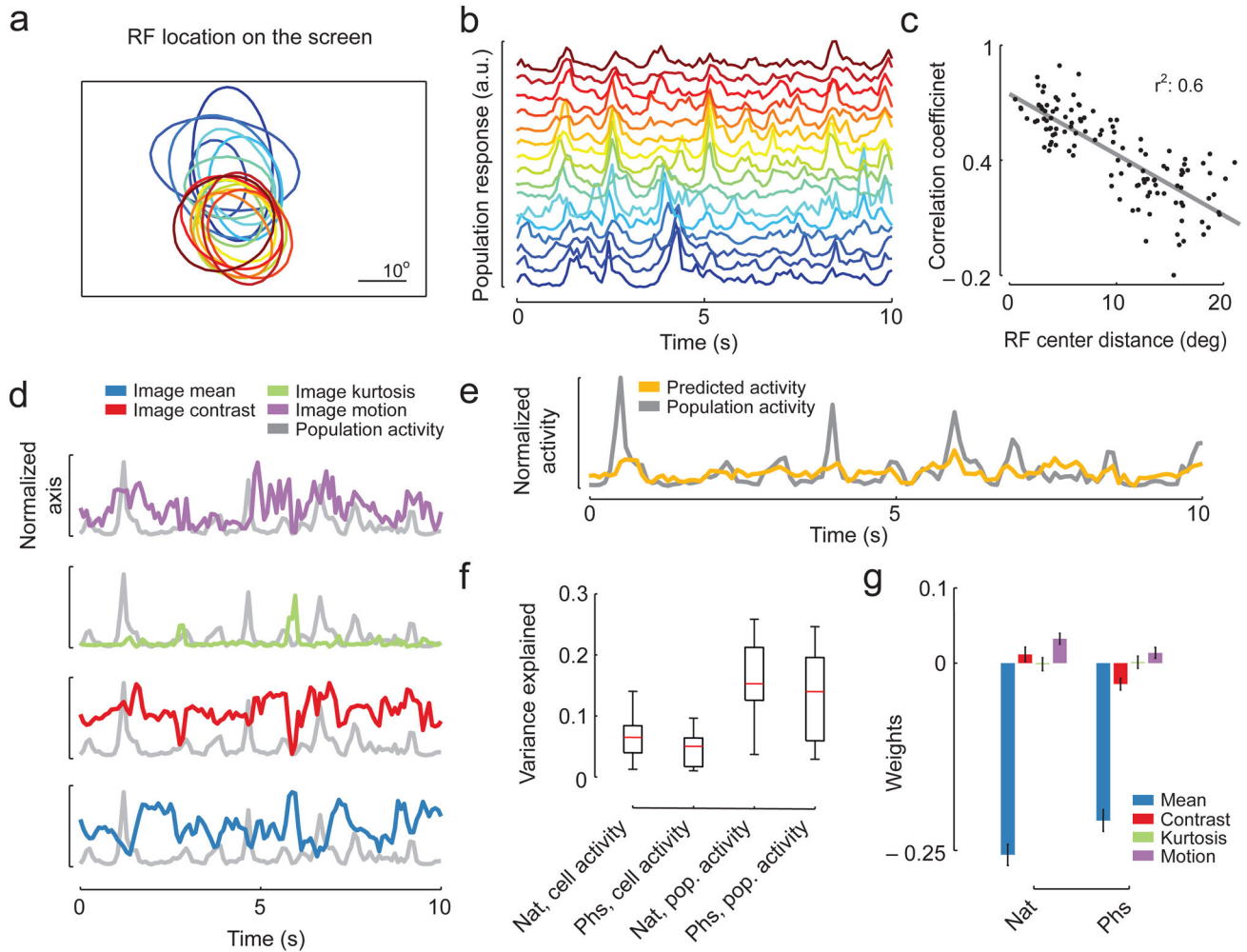


Figure 2. Image features contributing to population activation in V1

A. Population receptive fields from 16 sites recorded in a single animal superimposed on the screen. Different colors indicate different sites.

B. The population responses to one movie for the 16 sites shown in A.

C. Scatter plot of the linear correlation coefficients between the population responses of all pairs of sites shown in A. vs. their center to center distance in degrees ($N = 120$ pairs). The regression line is shown in gray.

D. Population responses and image statistics. The mean, contrast, kurtosis and motion of the movie for one example session evaluated within the population receptive field (blue, red, green and purple respectively) are plotted on top of the population response to that movie (gray).

E. A linear regression model from the image statistics from D can explain part of the population response. The population activity is partially followed by the predicted activity of the linear regression model (gray and yellow, respectively).

F. Boxplot of Variance Explained of a linear regression model ($N=21$ sites with at least 4 different movies; red line: median; black box: 25th and 75th percentile; whiskers: range), using the image statistics within the population receptive field of natural and phase

scrambled images for predicting the activity of either the individual cell responses (Cell Activity) or the population response (Pop. Activity).

G. Average weights of the regression model for the four image properties within the population receptive field of the two classes of stimuli (Nat, Phs) and the population response (N=21 sites, ± 1 SEM).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

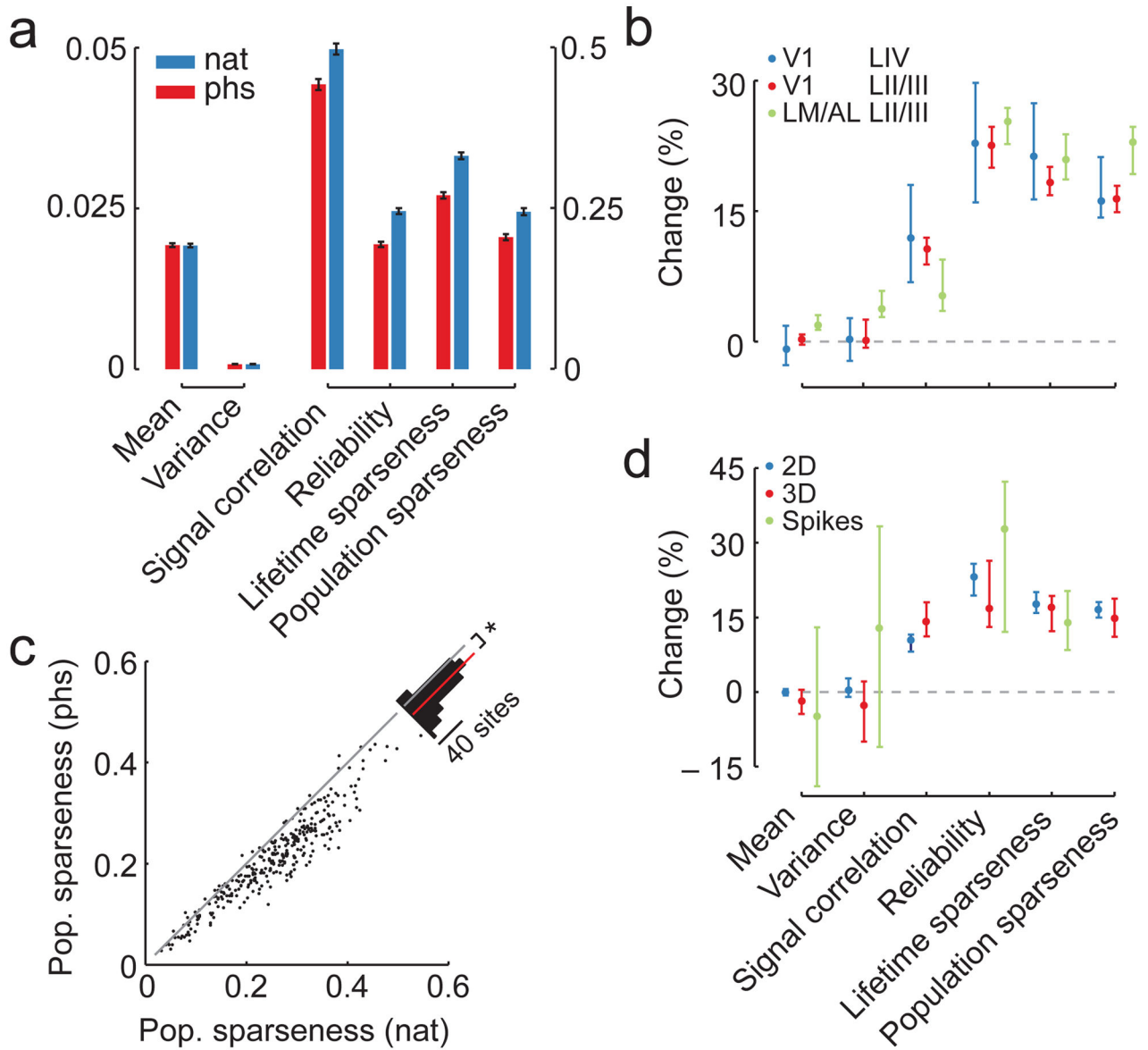


Figure 3. Higher order correlations in natural images change population structure

A. Mean measures of population activity (\pm 1SEM; $n = 315$ sites in layer 2/3 under anesthesia). Note the separate axis on the left and the right.

B. Median change in measures of population activity between stimulation with natural and phase scrambled movies separately for L4 ($n = 46$ sites) and L2/3 ($n = 315$ sites) of V1 and L2/3 of the secondary visual areas LM and AL ($n = 117$ sites; blue, red and green, respectively). Positive values indicate that the measure is higher under stimulation with natural movies. Error bars encompass the 95%-confidence intervals of the median.

C. Scatter plot of population sparseness for stimulation with natural and phase-scrambled movies. Inset shows histogram of absolute change with red bar indicating the mean difference ($*p < 0.001$, bootstrap).

D. Median change in measures of population activity between stimulation with natural and phase scrambled movies separately for 2D and 3D imaging (n = 332 sites for 2D imaging shown in blue; 30 sites for 3D imaging shown in red) and single neuron recordings (green; 29 cells). Pairwise signal correlations and population sparseness could not be evaluated for single neuron activity. Positive values indicate that the measure is higher under stimulation with natural movies. Error bars encompass the 95%-confidence intervals of the median.

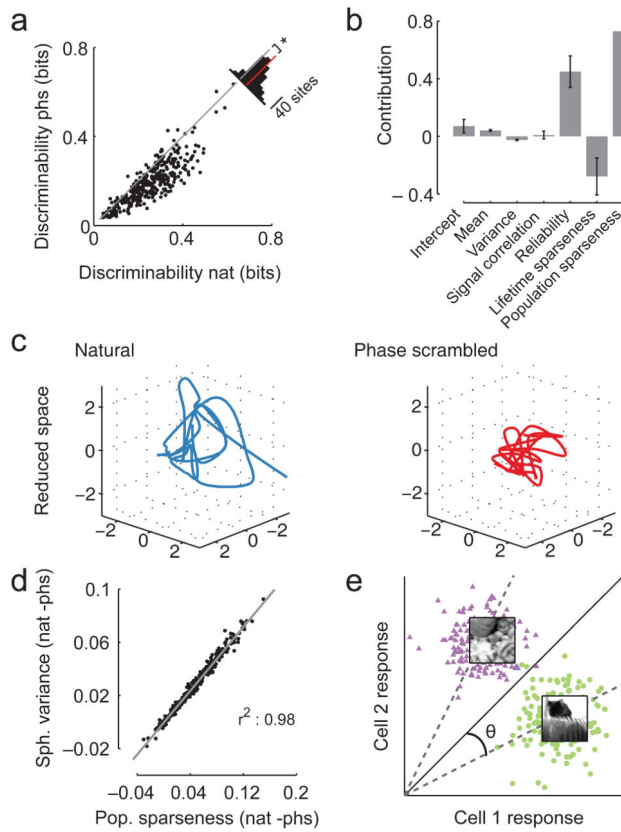


Figure 4. The population representation of natural movie scenes is easier to discriminate

A. Discriminability between pairs of scenes from natural movies and phase scrambled movies based on the population response ($n = 362$ sites). Inset shows histogram of absolute change with red bar indicating the mean difference ($*p < 0.001$, bootstrap).

B. Contributions of the change in different measures of the population structure towards predicting the average change in classification performance between natural and phase scrambled movies. Error bars show standard error for the regression coefficients scaled by the mean difference.

C. Low-dimensional representation of a 22-dimensional neural activity space from the responses of one site to natural and phase scrambled stimuli, illustrating how spread out the representation of different movie scenes in the reduced space.

D. Scatter plot of the difference in the spherical variance of each site between the natural and the phase scrambled stimuli, compared to the difference in the average population sparseness of each site between the two stimulus conditions.

E. Illustration of the geometrical intuition behind the relationship between population sparseness and classification performance: population sparseness measures the average \sin^2 of the angle between the main diagonal and the population response vector

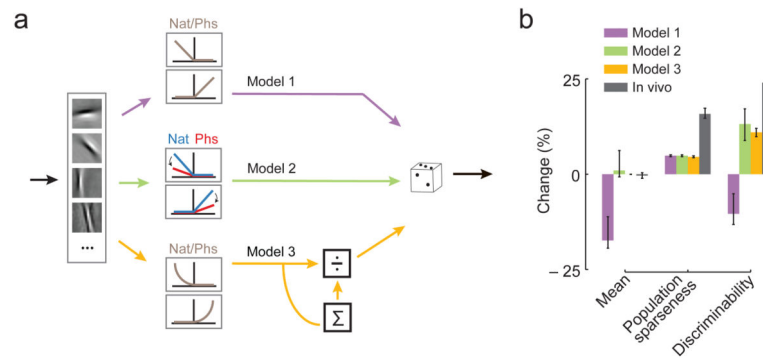


Figure 5. A sparse coding model with normalization is consistent with the results

A. Schematic of three LNP models used to generate the simulated responses. The initial linear filtering is the same for all (see methods). For Model 1, the filter output is rectified and used as a rate to generate Poisson spiking. An adaptive non-linearity is assumed that is adjusted to match the mean activity between the two classes for Model 2, and a normalization stage for Model 3.

B. Median change in average activity, population sparseness and discriminability in the sparse coding model (286 model neurons) for Model 1, Model 2 and Model 3 (purple, green and yellow, respectively) and the in vivo results (gray). Error bars encompass the 95%-confidence intervals of the median.

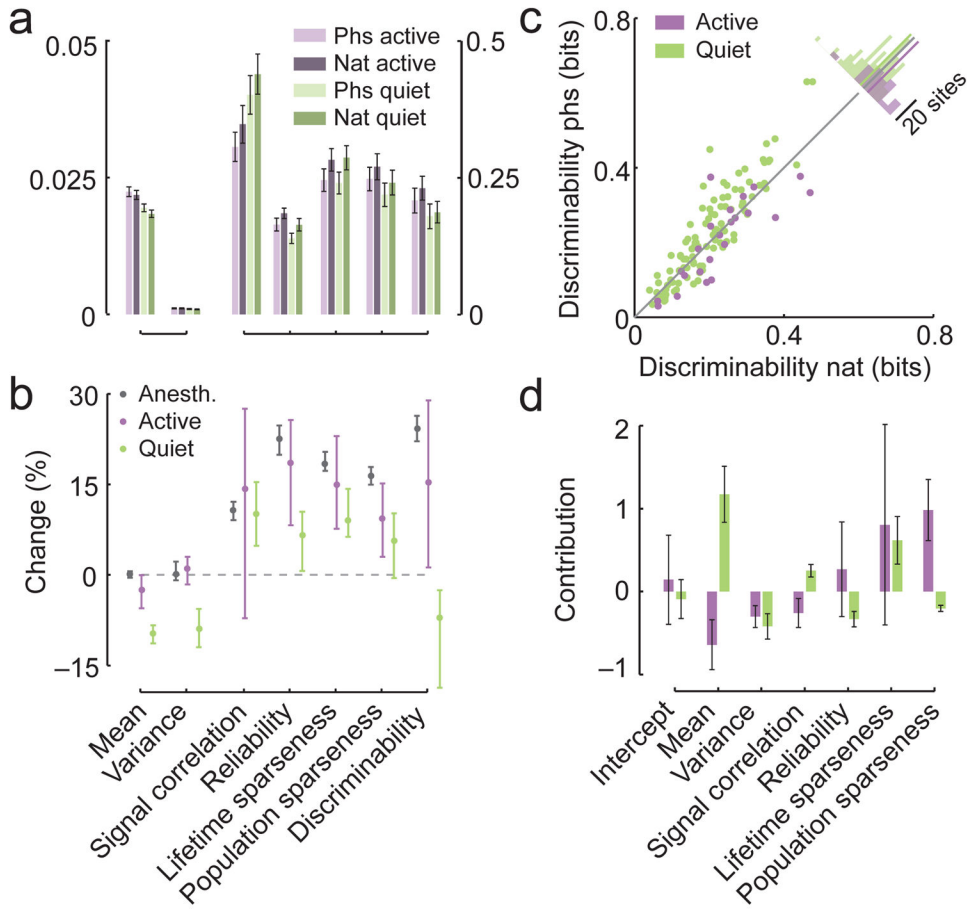


Figure 6. Active but not quiet wakefulness resemble anesthetized state

A. Mean measures of population activity for the responses to phase scrambled and natural stimuli (light and dark shading respectively) in the active ($n = 23$ sites) and quiet states ($n = 100$ sites) of awake animals (purple and green respectively; ± 1 SEM). Note the separate axis on the left and the right.

B. Median change in measures of population activity between stimulation with natural and phase scrambled movies separately for the anesthetized, active and quiet state of the animal (gray, purple and green, respectively). Positive values indicate that the measure is higher under stimulation with natural movies. Error bars encompass the 95%-confidence intervals of the median.

C. Discriminability between pairs of scenes from natural movies and phase scrambled movies based on the population response for the active and awake states (purple and green respectively). Inset shows histogram of absolute change with the colored bar indicating their mean difference.

D. Contributions of the change in different measures of the population structure towards predicting the average change in classification performance between natural and phase scrambled movies for the two awake states. Error bars show standard error for the regression coefficients scaled by the mean difference.