Article

State-dependent pupil dilation rapidly shifts visual feature selectivity

https://doi.org/10.1038/s41586-022-05270-3

Received: 5 December 2021

Accepted: 23 August 2022

Published online: 28 September 2022

Check for updates

Katrin Franke[1,2,3,4,9]✉, Konstantin F. Willeke[5,6,9], Kayla Ponder[3,4], Mario Galdamez[3,4], Na Zhou[3,4], Taliah Muhammad[3,4], Saumil Patel[3,4], Emmanouil Froudarakis[3,4,7], Jacob Reimer[3,4], Fabian H. Sinz[3,4,5,6,10] & Andreas S. Tolias[3,4,8,10]

To increase computational flexibility, the processing of sensory inputs changes with behavioural context. In the visual system, active behavioural states characterized by motor activity and pupil dilation[1,2] enhance sensory responses, but typically leave the preferred stimuli of neurons unchanged[2–9]. Here we find that behavioural state also modulates stimulus selectivity in the mouse visual cortex in the context of coloured natural scenes. Using population imaging in behaving mice, pharmacology and deep neural network modelling, we identified a rapid shift in colour selectivity towards ultraviolet stimuli during an active behavioural state. This was exclusively caused by state-dependent pupil dilation, which resulted in a dynamic switch from rod to cone photoreceptors, thereby extending their role beyond night and day vision. The change in tuning facilitated the decoding of ethological stimuli, such as aerial predators against the twilight sky[10]. For decades, studies in neuroscience and cognitive science have used pupil dilation as an indirect measure of brain state. Our data suggest that, in addition, state-dependent pupil dilation itself tunes visual representations to behavioural demands by differentially recruiting rods and cones on fast timescales.

Neuronal responses in animals are modulated by their behavioural and internal states to flexibly adjust information processing to different behavioural contexts. This phenomenon has been well described across animal species, from invertebrates[11,12] to primates[4,9]. In the mammalian visual cortex, neuronal activity is desynchronized and sensory responses are enhanced during an active behavioural state[1–3,5,7,8], which is characterized by pupil dilation[1] and locomotion activity[2]. Mechanistically, these effects have been linked to neuromodulators such as acetylcholine and noradrenaline (reviewed in refs. [13,14]). Other than changes in response gain, the tuning of visual neurons, such as orientation selectivity, typically does not change across quiet and active states[2,3,5,7,8]. So far, however, this has largely been studied in non-ecological settings using simple synthetic stimuli.

In this work, we study how behavioural state modulates cortical visual tuning in mice in the context of naturalistic scenes. Crucially, these scenes include the colour domain of the visual input due to its ethological relevance across species (reviewed in ref. [15]). Mice, like most mammals, are dichromatic and have two types of cone photoreceptor that express ultraviolet (UV)-sensitive and green-sensitive short-wavelength and medium-wavelength opsins (S-opsin and M-opsin, respectively)[16]. These UV-sensitive and green-sensitive cone photoreceptors predominantly sample the upper and the lower visual field, respectively, through uneven distributions across the retina[16,17].

To systematically study the relationship between neuronal tuning and behavioural state in the context of naturalistic scenes, we combined in vivo population calcium imaging of the primary visual cortex (V1) in awake, head-fixed mice with deep convolutional neural network (CNN) modelling. We extended a recently described model[18,19] to predict neuronal responses on the basis of both the visual input and the behaviour of the animal jointly. This enabled us to characterize the relationship between neuronal tuning and behaviour in extensive in silico experiments without the need to experimentally control the behaviour. Finally, we experimentally confirmed in vivo the in silico model predictions[18,20].

Using this approach, we demonstrate that colour tuning of mouse V1 neurons rapidly shifts towards higher UV sensitivity during an active behavioural state. By pharmacologically manipulating the pupil, we show that this is solely caused by pupil dilation. Dilation during active behavioural states sufficiently increases the amount of light entering the eye to cause a dynamic switch between rod-dominated and cone-dominated vision, even for constant ambient light levels. Finally, we show that the increased UV sensitivity during active periods may tune the mouse visual system to improved detection of predators against the UV background of the sky. Our results identify a new functional role of state-dependent pupil dilation: to rapidly tune visual feature representations to changing behavioural requirements in a bottom-up manner.

[1]Institute for Ophthalmic Research, Tübingen University, Tübingen, Germany. [2]Center for Integrative Neuroscience, Tübingen University, Tübingen, Germany. [3]Department of Neuroscience, Baylor College of Medicine, Houston, TX, USA. [4]Center for Neuroscience and Artificial Intelligence, Baylor College of Medicine, Houston, TX, USA. [5]Institute for Bioinformatics and Medical Informatics, Tübingen University, Tübingen, Germany. [6]Department of Computer Science, Göttingen University, Göttingen, Germany. [7]Institute of Molecular Biology and Biotechnology, Foundation for Research and Technology Hellas, Heraklion, Greece. [8]Department of Electrical and Computer Engineering, Rice University, Houston, TX, USA. [9]These authors contributed equally: Katrin Franke and Konstantin F. Willeke. [10]These authors jointly supervised this work: Fabian H. Sinz and Andreas S. Tolias. ✉e-mail: katrin.franke@uni-tuebingen.de

CNNs identify optimal coloured stimuli

Here we studied the relationship between neuronal tuning in mouse V1 and the behaviour of the animal, specifically focusing on colour processing because of its behavioural relevance (reviewed in ref. [15]). We presented coloured naturalistic images (Extended Data Fig. 1) to awake, head-fixed mice positioned on a treadmill (Fig. 1a) while recording the calcium activity of L2/3 neurons in V1 using two-photon imaging (Fig. 1c,d). We simultaneously recorded locomotion activity, pupil size and instantaneous changes in pupil size, which have all been associated with distinct behavioural states[1,2]. Visual stimuli were presented using a projector with UV and green light-emitting diodes (LEDs)[21] (Fig. 1b), which enabled the differential activation of UV-sensitive and green-sensitive mouse photoreceptors. We recorded neuronal responses along the posterior–anterior axis of V1 (Fig. 1c), sampling from various vertical positions across the visual field. This choice was motivated by the gradient of spectral sensitivity of mouse cone photoreceptors across the retina[16,17].

We used a deep CNN to learn an in silico model of the recorded neuron population as a function of the visual input and the behaviour of the animal[18] (Fig. 1e). The CNN had the following input channels: (1) UV and green channels of the visual stimulus; (2) three channels set to the recorded behavioural parameters (that is, pupil size, change in pupil size and locomotion); and (3) two channels that were shared across all inputs encoding the x and y pixel positions of the stimulus image. The third criterion was previously shown to improve CNN model performance in cases for which feature representations depend on image position[22], similar to the gradient in mouse colour sensitivity across visual space. Our neural predictive models also included a shifter network[18] that spatially shifted the receptive fields of model neurons according to the recorded pupil position traces. For each dataset, we trained an ensemble of four-layer CNN models end-to-end[19] to predict the neuronal responses to individual images and behavioural parameters. The prediction performance of the resulting ensemble model (Extended Data Fig. 2) was comparable to state-of-the-art predictive models of mouse V1 (ref. [19]).

Using our CNN ensemble model as a 'digital twin' of the visual cortex, we synthesized maximally exciting inputs (MEIs) for individual neurons (Fig. 1f and Extended Data Fig. 3a). To this end, we optimized the UV and green colour channels of a contrast-constrained image to produce the highest activation in the given model neuron using regularized gradient ascent[18,20]. For most of the neurons, MEI colour channels were positively correlated, which indicated that colour opponency is rare given our stimulus paradigm (Extended Data Figs. 3 and 4). Inception loop experiments[18] confirmed that the computed MEIs strongly drive the recorded neurons. For these experiments, we randomly selected MEIs of 150 neurons above a response reliability threshold for presentation on the next day (Fig. 1g). For most neurons, the MEIs were indeed the most exciting stimuli: responses of neurons to their own MEI were significantly larger than to other MEIs (Fig. 1h; for statistics, see figure legends and Supplementary Methods). Together, these findings demonstrate that our modelling approach accurately captures the tuning properties of mouse V1 neurons in the context of coloured naturalistic scenes.

V1 colour tuning changes with behaviour

To study how cortical colour tuning changes with behavioural state, we performed detailed in silico characterizations using the above-described trained CNN model. To that end, we focused on two well described and spontaneously occurring behavioural states[1,2]: (1) a quiet state with no locomotion and a small pupil (3rd percentile of locomotion and pupil size across all trials) and (2) an active state indicated by locomotion and a larger pupil (97th percentile). For each neuron and distinct behavioural state, we optimized a MEI and then generated a

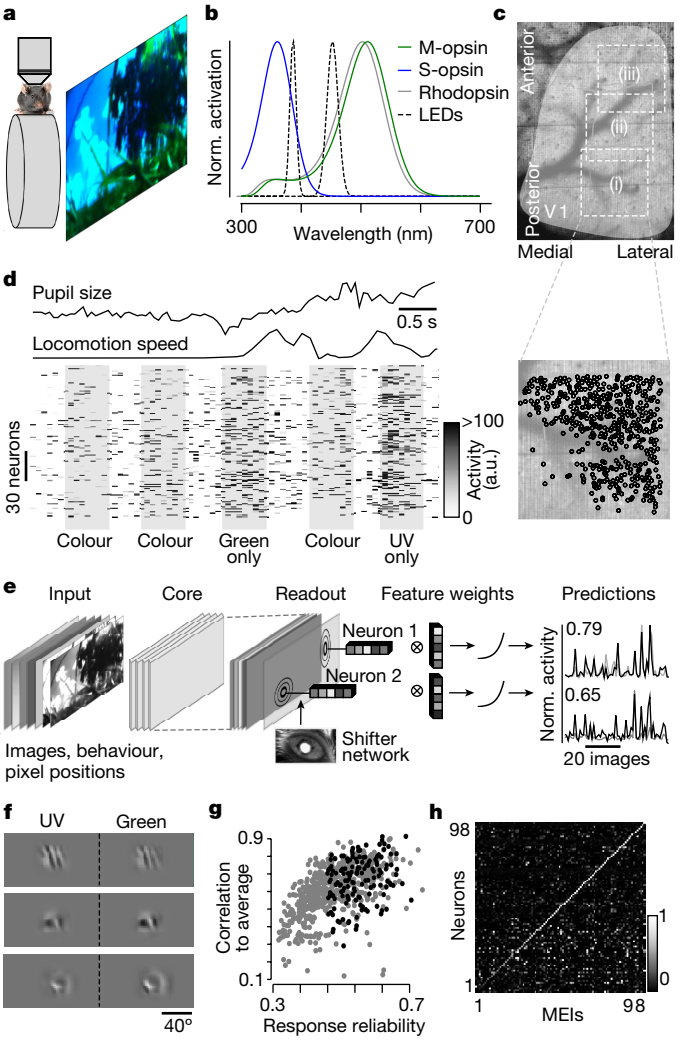


Fig. 1 | Deep neural networks capture mouse V1 tuning properties in the context of coloured naturalistic scenes. a, Schematic of the experimental setup. Awake, head-fixed mice on a treadmill were presented with UV-coloured and green-coloured naturalistic scenes (Extended Data Fig. 1). **b**, Normalized (Norm.) sensitivity spectra of mouse S-opsin and M-opsin expressed by cones and rhodopsin expressed by rods, with LED spectra for visual stimulation. **c**, Cortical surface of a transgenic mouse expressing GCaMP6s, with positions of three scan fields ((i)–(iii), 650 × 650 µm each). The bottom image shows cells (n = 478) selected for further analysis. **d**, Neuronal activity (shown in arbitrary units (a.u.); n = 150 cells) in response to coloured naturalistic scenes and simultaneously recorded behavioural data (pupil size and locomotion speed). **e**, Schematic of the model architecture. The model input consists of two image channels, three behaviour channels and two position channels that encode the x and y pixel position of the input images[22]. A four-layer convolutional core is followed by a Gaussian readout and a nonlinearity[19]. Readout positions were adjusted using a shifter network[18]. Traces on the right show average responses (grey) to test images of two example neurons and corresponding model predictions (black). **f**, MEI images of three example neurons (from n = 658). See also Extended Data Fig. 3. **g**, Response reliability to natural images plotted against model prediction performance of all cells of one scan. Neurons selected for experimental verification (inception loop) are indicated in black. **h**, Confusion matrix of the inception loop experiment[18] depicting the activity of each selected neuron to presented MEIs. Neurons are ordered on the basis of the response to their own MEI (>65% showed the strongest response to their own MEI). Responses of neurons to their own MEI (along the diagonal) were significantly larger than to other MEIs (P = 0 for a one-sided permutation test, n = 10,000 permutations).

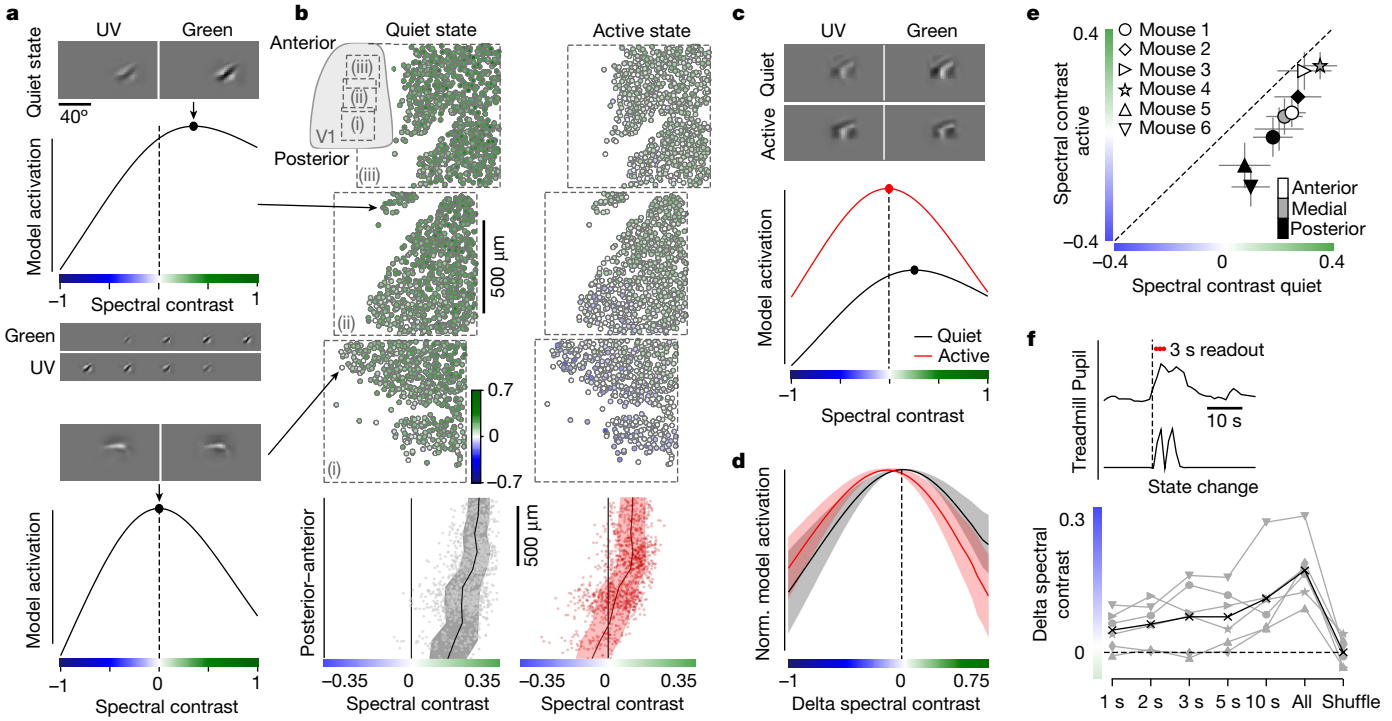


Fig. 2 | V1 colour-tuning changes with the behavioural state. a, MEIs optimized for a quiet state (3rd percentile of pupil and locomotion) and model activations for varying MEI spectral contrasts ($n = 50$) of two example neurons (from $n = 1{,}759$). Example stimuli are shown below. Arrows indicate the cortical position of neurons. **b**, Neurons ($n = 1{,}759$ neurons, $n = 3$ scans, $n = 1$ mouse) along the posterior–anterior V1, colour-coded on the basis of the spectral contrast of quiet and active state (97th percentile) MEIs. Inset shows the scan positions within V1. Bottom shows MEI spectral contrasts of neurons from the top, with binned average and s.d. shading. The spectral contrast significantly varied across the anterior–posterior V1 axis ($P = 10^{-16}$ for the smooth term on the cortical position of the generalized additive model (GAM); see Supplementary Methods for more details). **c**, MEIs of an example neuron optimized for a quiet and an active state, with colour-tuning curves shown below. **d**, Population mean with s.d. shading of peak-normalized colour-tuning curves from **b** and **c** aligned with respect to the peak of the tuning curves from the quiet state. The optimal spectral contrast shifted significantly towards higher UV sensitivity during active periods ($P = 10^{-16}$ for the behavioural state coefficient of the GAM). **e**, Mean MEI spectral contrast of quiet and active states across animals ($n = 478$ (mouse 1, posterior), 623 (mouse 1, medial), 658 (mouse 1, anterior), 843 (mouse 2), 711 (mouse 3), 822 (mouse 4), 769 (mouse 5), 706 (mouse 6) cells, $n = 8$ scans, $n = 6$ animals). Error bars indicate the s.d. across neurons. Wilcoxon signed-rank test (two-sided): $P = 10^{-78}$ (mouse 1, posterior), 10^{-103} (mouse 1, medial), 10^{-109} (mouse 1, anterior), 10^{-139} (mouse 2), 10^{-50} (mouse 3), 10^{-136} (mouse 4), 10^{-127} (mouse 5), 10^{-111} (mouse 6). **f**, Pupil size and treadmill velocity over time. Dashed line indicates the state change from quiet to active. Red dots indicate active trials used for analyses for a 3-s readout period. Bottom, change in mean MEI spectral contrast ($n = 6$ animals) between quiet and active states for different readout lengths after the state change, with mean across animals (black). All, all trials; Shuffle, shuffled behaviour relative to responses. One-sample t-test across animals (two-sided): $P = 0.038$ (1 s), $P = 0.029$ (2 s), $P = 0.053$ (3 s), $P = 0.03$ (5 s), $P = 0.021$ (10 s), $P = 0.001$ (All), $P = 0.92$ (Shuffled).

colour-tuning curve by predicting the activity of the neuron to varying colour contrasts of this MEI (Fig. 2a and Extended Data Fig. 5).

For both behavioural states, the optimal spectral contrast of neurons systematically varied along the anterior–posterior axis of V1 (Fig. 2b). The UV sensitivity significantly increased from anterior to posterior V1, which is in line with the distribution of cone opsins across the retina[16,17] and with previous studies of V1 (ref. [23]) and the dorsal lateral geniculate nucleus[24]. Nevertheless, for quiet behavioural periods, nearly all neurons preferred a green-biased stimulus (Fig. 2b, left), even the ones positioned in the posterior V1, which receives input from the ventral retina, where cones are largely sensitive to UV light[17]. This distribution of V1 colour preferences indicates that visual responses during quiet states are largely driven by rod photoreceptors that are sensitive to green light[25].

By contrast, during active periods, the colour tuning of neurons systematically shifted towards higher UV sensitivity (Fig. 2b–d). This was accompanied by an overall increase in neuronal activation predicted by the model (Fig. 2c and Extended Data Fig. 6a,d), which is in agreemnt with previous results[2,5]. The shift in colour selectivity was observed across animals for both the posterior and anterior V1 (Fig. 2e). As a result, neurons in the posterior V1 exhibited UV-biased MEIs, whereas neurons in the anterior V1 largely maintained their preference for green-biased stimuli. This is consistent with a cortical distribution of colour tuning expected from a shift from rod-dominated to cone-dominated visual responses[25]. Notably, the spatial structure of the MEIs was largely unchanged across behavioural states (Fig. 2c and Extended Data Fig. 5).

The shift in colour selectivity with behavioural state was fast, operating on the timescale of seconds (Fig. 2f). To test the temporal dynamics of the shift in tuning, we identified state changes from quiet to active periods by detecting rapid increases in pupil size after a prolonged quiet period. Then we sampled active trials within different time bins after the state change, trained CNN models on this subselection of active trials and all quiet trials and optimized MEIs as described above. The shift in colour selectivity with behavioural state was evident for a 10-s readout window for all animals tested. Notably, for the majority of animals ($n = 4$ out of 6), the shift was already present when training a model based on active trials that sampled just 1 s after the state change.

We wanted to confirm the above prediction from our in silico analysis that mouse V1 colour tuning rapidly shifts towards higher UV sensitivity during active periods. To that end, we used a well-established sparse noise paradigm for mapping the receptive fields of visual neurons

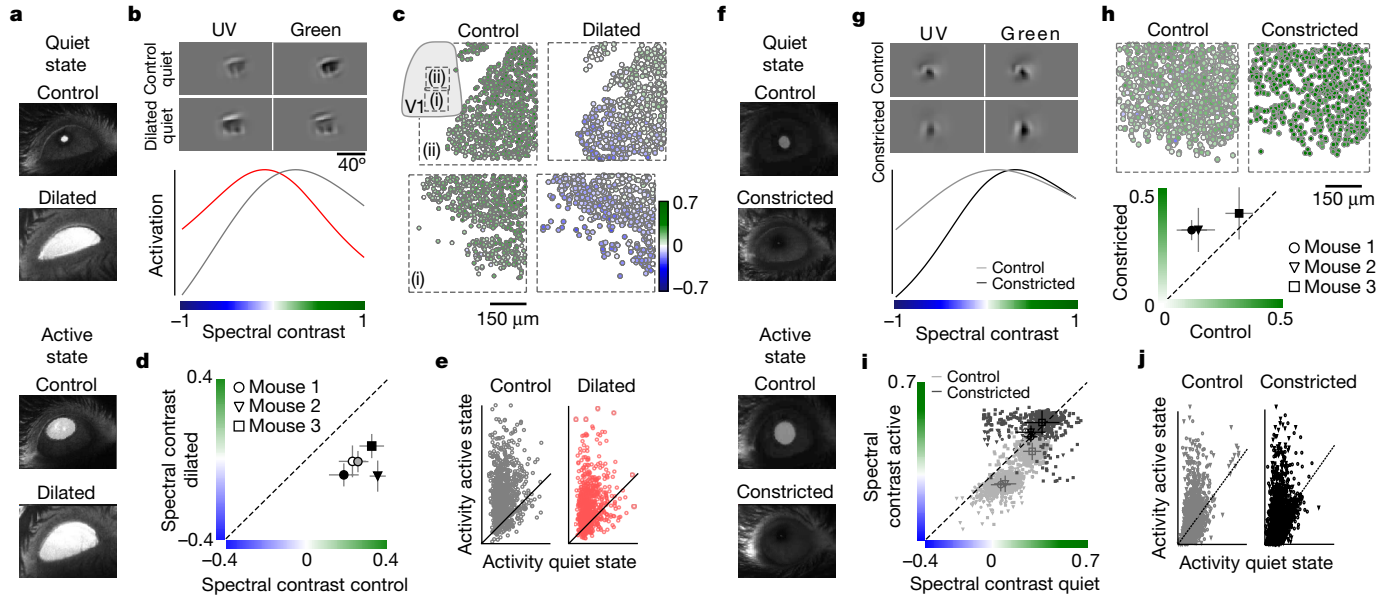


Fig. 3 | Pupil dilation causes the state-dependent shift in V1 colour selectivity. a, Example images from the eye camera during a quiet and an active state and for control and dilated conditions (atropine). **b**, MEIs of an example neuron (from $n = 478$) optimized for a quiet state for the control (black) and dilated (red) conditions (top) and peak-normalized colour-tuning curves (bottom). Neurons were matched anatomically across recordings. **c**, Neurons ($n = 1{,}101$) recorded in two experiments for the control (from Fig. 2) and the dilated condition, colour coded on the basis of the spectral contrast of the quiet state MEI. The spectral contrast significantly varied across the anterior–posterior V1 axis for the dilated condition ($n = 1{,}859$, $P = 10^{-16}$ for the smooth term on the cortical position of the GAM; see Supplementary Methods for more details). **d**, Mean spectral contrasts of quiet state MEIs in the control compared with the dilated condition ($n = 478$ (mouse 1, posterior, control), 623 (mouse 1, medial, control), 658 (mouse 1, anterior, control), 711 (mouse 2, control), 1,109 (mouse 3, drug), 464 (mouse 1, posterior, drug), 689 (mouse 1, medial, drug), 706 (mouse 1, anterior, drug), 723 (mouse 2, drug), 1,090 (mouse 3, drug) cells, $n = 10$ scans, $n = 3$ animals). Error bars indicate the s.d. across neurons. Two-sample t-test (two-sided): $P = 0$ for all scans. **e**, Mean activity of neurons from **c** during the quiet and active behavioural periods in the control and dilated conditions. **f**,**g**, Same as **a** (**f**) and **b** (**g**), but for pupil constriction with carbachol. **h**, Neurons recorded in posterior V1 ($n = 751$ (control) and 518 (constricted)), colour coded on the basis of the spectral contrast of a quiet state MEI. Bottom shows the mean spectral contrast of quiet state MEIs in control compared with the constricted condition ($n = 822$ (mouse 1, control), 769 (mouse 2, control), 1,109 (mouse 3, control), 751 (mouse 1, drug), 1,037 (mouse 2, drug), 1,028 (mouse 3, drug) cells, $n = 6$ scans, $n = 3$ mice). Error bars indicate the s.d. across neurons. Two-sample t-test (two-sided): $P = 0$ (mouse 1), 0 (mouse 2), 10^{-38} (mouse 3). **i**, Spectral contrast of quiet state MEIs compared with the spectral contrast of active state MEIs ($n = 778$ neurons, $n = 6$ scans, $n = 3$ mice), for the control (grey) and the constricted conditions (black). Only neurons with a test correlation value of >0.3 are shown. Wilcoxon signed-rank test (two-sided): $P = 10^{-134}$ (mouse 1, control), 10^{-127} (mouse 2, control), 10^{-170} (mouse 3, control), $P = 0.98$ (mouse 1, constricted), 0.0003 (mouse 2, constricted), 10^{-6} (mouse 3, constricted). **j**, Same as **e**, but for neurons from **h** in the control and the constricted conditions.

(Extended Data Fig. 7a). Trials were separated into quiet (<50th percentile) and active periods (>75th percentile) using the simultaneously recorded pupil size trace. For each neuron and behavioural state, we estimated a spike-triggered average (STA) that represented the preferred stimulus of the neuron in the context of the sparse noise input (Extended Data Fig. 7b). Consistent with the in silico analysis, we observed that most V1 neurons preferred a green-biased stimulus during the quiet behavioural state (Extended Data Fig. 7c). Moreover, neurons in the posterior and medial V1 showed increased UV sensitivity during active periods (Extended Data Fig. 7c,d). The UV shift was also present in the anterior V1, but only for more extreme pupil size thresholds (20th and 85th percentiles; Extended Data Fig. 7e). Finally, we confirmed that V1 colour preference shifted within a few seconds after onset of an active behavioural state (Extended Data Fig. 7e). Together, these results confirm the prediction of the CNN model that mouse V1 colour tuning rapidly changes with behavioural state, particularly for neurons that sample the upper visual field.

Pupil dilation shifts neuronal tuning

Next, we investigated the mechanism underlying the observed behaviour-related changes in colour tuning of mouse V1 neurons. On the one hand, the behavioural state of the animal affects neuronal activity through neuromodulation that acts on multiple stages of the visual system[6,8,26–28]. On the other hand, state-dependent pupil dilation results in higher light intensities at the level of the retina that might also affect visual processing[29,30].

To experimentally test the relative contribution of these two mechanisms, we dissociated state-dependent neuromodulatory effects from changes in pupil size by pharmacologically dilating and constricting the pupil with atropine and carbachol eye drops, respectively (Fig. 3a,f). We recorded visual responses to naturalistic scenes during control and pharmacology conditions and trained separate CNN models (Extended Data Fig. 2c).

Pupil dilation with atropine eye drops was sufficient to shift the colour tuning of neurons towards higher UV sensitivity, whereas locomotion activity was not necessary. During a quiet state with no locomotion, MEI colour tuning systematically shifted towards higher UV sensitivity for the dilated pupil compared with the control condition (Fig. 3b–d). We confirmed the role of pupil size in modulating colour tuning of mouse V1 neurons by also recording visual responses to the sparse noise stimulus after dilating the pupil with atropine (Extended Data Fig. 8).

To test whether pupil dilation is not only sufficient but also necessary for the behavioural shift in colour tuning, we dissociated pupil dilation from neuromodulation during active periods by temporarily constricting the pupil with carbachol eye drops (Extended Data Fig. 2f). The gain increase of neuronal responses with locomotion persisted under these pharmacological manipulations of the pupil[6,26,28]

(Fig. 3e,j), which indicated that this well-known effect of neuromodulation was unaffected. For quiet periods, pupil constriction resulted in a systematic shift towards higher green sensitivity compared with the control condition (Fig. 3g,h). Notably, we did not observe a significant shift towards higher UV sensitivity during active periods for the constricted condition, whereas the shift was evident in the control condition (Fig. 3i). This suggests that neuromodulation or other internal state-dependent mechanisms during active behavioural periods are not sufficient to drive the shift in colour tuning with behaviour, whereas state-dependent pupil dilation is necessary for the effect.

Tuning shift is caused by photoreceptors

Previous studies have shown that in mice, pupil size regulates retinal illuminance levels by more than one-order of magnitude[31]. This affects the relative activation levels of the green-sensitive rods and UV-sensitive and green-sensitive cones, thereby changing cortical colour preferences in anaesthetized mice[25]. To test whether our data could be explained by a shift from rod to cone photoreceptors during active behavioural periods because of a larger pupil (Fig. 4a), we estimated activation levels of mouse photoreceptors as a function of pupil size[10]. For our experiments, we observed up to a tenfold increase in pupil area and an equal increase in the estimated photoisomerization rate for an active compared with a quiet behavioural state (Fig. 4a, bottom). Therefore, the change in retinal light level due to pupil dilation during an active state is probably sufficient to dynamically shift the mouse visual system from a rod-dominated to a cone-dominated operating regimen.

If this was true, we would expect that the shift in colour selectivity can be reproduced for constant pupil sizes by changing ambient light levels. We experimentally confirmed this prediction by reducing the light intensity of the visual stimulus by 1.5-orders of magnitude while keeping the pupil size constant across recordings through pharmacological dilation with atropine (Fig. 4b). The low-light-intensity condition was expected to predominantly activate rod photoreceptors, which are green sensitive. Indeed, V1 neurons exhibited more green-biased MEIs for the low compared with the high light condition. Together with our pupil dilation and constriction experiments, this result strongly suggests that pupil dilation during active states results in a dynamic shift from rod-driven to cone-driven visual responses and a corresponding shift in spectral sensitivity.

Tuning shift affects population decoding

Next, we tested whether the shift in colour tuning during an active state might increase visual performance at the level of large populations of neurons in response to naturalistic stimuli. First, we applied a contrast-constrained image reconstruction paradigm[32] using the above-described trained CNN model (Extended Data Fig. 9a). Stimulus reconstruction from neuronal activity has previously been used to infer the most relevant visual features encoded by the neuron population[33], such as the colour sensitivity of neurons. Most reconstructed images for a quiet behavioural state exhibited higher contrast in the green channel, whereas the contrast was shifted towards the UV channel during active states (Extended Data Fig. 9b,c). This indicated that the increase in UV sensitivity during active periods observed at the single-cell level might contribute to specific visual tasks such as stimulus discrimination performed by populations of neurons in mouse V1.

We experimentally confirmed this prediction by showing that the decoding of UV objects selectively improved during active periods. To that end, we modified a recent object-decoding paradigm[34]. Mice passively viewed movie clips with two different objects presented in either the UV or green image channel (Fig. 5b) while recording the population calcium activity in the posterior V1 as described above. We estimated

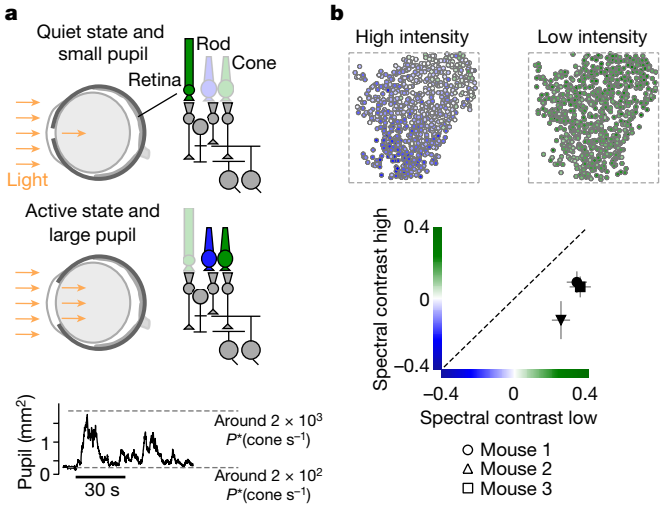


Fig. 4 | Pupil dilation during an active behavioural state differentially recruits rod and cone photoreceptors. a, Schematic of the mouse eye for a quiet behavioural state with a small pupil (top) and an active state with a large pupil (middle). Right, simplified circuit diagrams of the vertebrate retina. Activation of rod and cone photoreceptors are indicated by the degree of transparency. Arrows indicate the amount of light entering the eye through the pupil. Photoreceptors are coloured on the basis of their peak wavelength sensitivity. Bottom, pupil area recorded during functional imaging, with estimated photoisomerization rates (P^* (cone s^{-1})) for a small pupil and a large pupil. **b**, Top, neurons recorded in the posterior V1, colour coded on the basis of the spectral contrast of their quiet state MEI under the dilated condition for a high monitor intensity (n = 1,125 cells) and a 1.5-order of magnitude lower monitor intensity (n = 1,059 cells). Bottom, the mean spectral contrast of quiet state MEIs in low compared with a high monitor intensity condition (n = 1,125 (mouse 1, low), 651 (mouse 2, low), 1,090 (mouse 3, low), 1,059 (mouse 1, high), 627 (mouse 2, high), 1,068 (mouse 3, high) cells, n = 6 scans, n = 3 mice). Error bars indicate the s.d. across neurons. Two-sample t-test (two-sided): P = 0 for all scans.

the discriminability of object identity of UV and green objects from the recorded neuronal responses using a nonlinear support vector machine (SVM) decoder (Fig. 5a). Consistent with previous reports[1,35,36], decoding discriminability was higher during active compared with quiet behavioural periods (Fig. 5c). However, the increase in decoding discriminability of UV objects was larger than for green objects, which is consistent with an increase in UV sensitivity during active behavioural periods. This result was statistically significant compared with the result of a permutation test that shuffled quiet and active trials. The selective increase in decoding discriminability of UV objects was also present for a subset of recordings with modified stimuli, such as with reduced object contrast or different object polarity (Extended Data Fig. 10).

We then considered the behavioural relevance of this increase in UV sensitivity during an active state for mice. It has recently been shown that during dusk and dawn, aerial predators in the natural environment of mice are more visible in the UV than the green wavelength range[10] (Fig. 5d). Therefore, an increase in UV sensitivity of mouse visual neurons for an alert behavioural state might facilitate the detection of predators visible as dark silhouettes in the sky. To investigate this hypothesis on the level of populations of neurons, we presented parametric stimuli inspired by these natural scenes, which contained either only noise or an additional dark object in the green or UV image channel, to passively viewing mice (Fig. 5e). This experiment revealed that decoding detection of the behaviourally relevant stimulus—corresponding to the dark object being presented in the UV channel—was substantially increased for an active behavioural state. Decoding detection of the green objects did not increase to a similar extent (Fig. 5f). This result

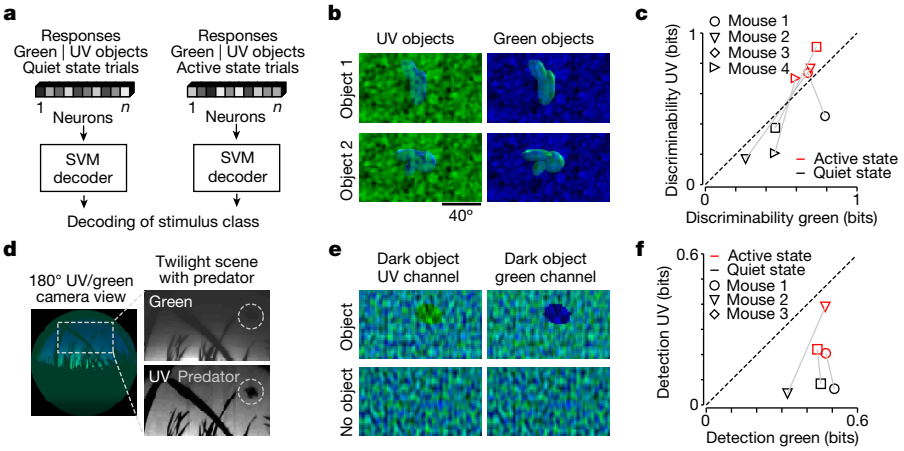


Fig. 5 | Shift in colour preference during an active state facilitates decoding of behaviourally relevant stimuli. a, Schematic illustrating the decoding paradigm. Neuronal responses for either quiet or active trials to green or UV objects were used to train a nonlinear SVM decoder to predict stimulus classes. **b**, Example stimulus frames of green and UV objects on top of noise. Stimulus conditions were presented as 5-s movie clips in random order. **c**, Scatter plot of the decoding discriminability of green compared with UV objects for quiet and active trials ($n = 4$ animals) for a SVM decoder trained on all neurons of each scan ($n = 1{,}090$ (mouse 1), 971 (mouse 2), 841 (mouse 3), 918 (mouse 4) cells). Grey lines connect the quiet and the active state performance of the same animal, with slopes larger than one indicating a larger increase in decoding performance for UV versus green objects. *P* values obtained from one-sided permutation test: $P < 0.002$ (mouse 1), $P < 0.044$ (mouse 2), $P < 0.024$ (mouse 3), $P < 0.01$ (mouse 4). **d**, Natural scene recorded at sunrise with a custom camera adjusted to the spectral sensitivity of mice[10], with a drone mimicking an aerial predator. Right images show single colour channels of a cropped image from the left, with the mock predator highlighted by a white dashed open circle. **e**, Parametric stimuli inspired by natural scene in **d** showing a dark object in either the UV or the green image channel (top) or noise only (bottom), with the object present or absent as the decoding objective. Stimuli were shown for 0.5 s with 0.3–0.5-s periods of grey screen in between. **f**, Similar to **c**, but for decoding detection of green versus UV dark objects from **e** ($n = 773$ (mouse 1), 1,049 (mouse 2), 1,094 (mouse 3) cells). *P* values were obtained from one-sided permutation test (see Methods for detail): $P < 0.008$ (mouse 1), $P < 0.009$ (mouse 2), $P < 0.008$ (mouse 3).

suggests that on the population level, the shift towards higher UV sensitivity might be behaviourally relevant as it selectively improves the decoding detection of dark objects in the UV channel, analogous to a predatory bird flying in a UV-bright sky.

Discussion

Our work identified a new mechanism by which state-dependent pupil dilation dynamically tunes the feature selectivity of the mouse visual system to behaviourally relevant stimuli.

The fact that sensory responses are modulated by the motor activity and the internal state of the animal was first demonstrated in elegant studies of invertebrates many decades ago[11,37]. Since then, modulation of sensory responses as a function of behavioural and internal states, such as attention, has been described in many animals[2,4,38,39]. Across animal species, state-dependent modulation predominantly affects neuronal responsiveness[2,9,27,28], which results in better behavioural performance[7,35,36,40]. In a few cases, however, the tuning properties of sensory circuits are also affected by this modulation. In the visual system, this has been reported, for instance, for temporal tuning in *Drosophila*[12], rabbits[39] and mice[41], as well as for direction selectivity in primates[4]. In these cases, the visual system might bias processing towards visual features relevant for current behavioural goals, such as higher temporal frequencies during periods of walking, running or flying.

Here, we demonstrated a shift in neuronal tuning with behavioural state in mice, focusing on the colour domain, which has rarely been studied in the context of behavioural modulation. Our results suggested that the shift towards higher UV sensitivity during active behavioural periods may help support ethological tasks, such as the detection of predators in the sky. In particular, UV vision has been implicated in predator and prey detection in several animal species as an adaptation to living in different natural environments (reviewed in ref. [42]). This is related to the stronger scattering of short wavelength light in general as well as ozone absorption[43] in the sky, which probably facilitate the detection of objects as dark silhouettes against a UV-bright background in the sky[10], underwater and against the snow[42]. However, it will be important to directly test the behavioural relevance of the described shift in colour tuning during an active state for mouse predator detection. For example, combining an overhead detection task of a looming stimulus presented in UV or green light conditions[44] with pharmacological pupil manipulations or careful tracking of pupil dynamics[45] will reveal whether pupil dilation results in better behavioural detection of UV stimuli, as suggested by our results.

Mechanistically, state-dependent modulation of visual responses has been linked to neuromodulators such as acetylcholine and noradrenaline (reviewed in refs. [13,14]), which are released with active behavioural states and alert internal states. Our results demonstrated that in addition to internal brain state mechanisms, dynamic changes in pupil size are both sufficient and necessary to affect cortical tuning (see also Supplementary Discussion). We propose that this mechanism changes colour sensitivity through differential rod versus cone activation, which is reminiscent of the Purkinje shift described in humans[46], although acting on faster timescales. A recent neurophysiological study[25] that used anaesthetized mice demonstrated that pharmacological pupil dilation at constant ambient light levels is sufficient to induce a shift from rod-driven to cone-driven visual responses in V1. Our data indicated that a switch between the rod and cone system can also happen dynamically at the timescale of seconds in behaving mice as a consequence of changes in pupil size across distinct behavioural states. As rod and cone photoreceptors differ with respect to spatial distribution, temporal resolution and degree of nonlinearity (discussed in ref. [47]), dynamically adjusting their relative activation might influence the sensory representation of the visual scene far beyond the colour domain of the visual input.

Changes in pupil size driven by behavioural and internal states of the animal are common features shared across most vertebrate species studied so far (reviewed in ref. [48]), including amphibians, birds and mammals (see also Supplementary Discussion). Notably, pupil

dilation is probably under voluntary control for some animals such as birds and reptiles (discussed in ref. [49]), and potentially even for some humans[50]. We propose that state-dependent pupil size changes might act as a general mechanism across species to rapidly switch between the rod-driven and cone-driven operating regimen, thereby tuning the visual system to different features, as suggested here for predator detection in mice during dusk and dawn. Our findings provide a functional explanation to the long-standing debate of why pupil size is modulated with internal and behavioural states.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-022-05270-3.

1. Reimer, J. et al. Pupil fluctuations track fast switching of cortical states during quiet wakefulness. *Neuron* **84**, 355–362 (2014).
2. Niell, C. M. & Stryker, M. P. Modulation of visual responses by behavioral state in mouse visual cortex. *Neuron* **65**, 472–479 (2010).
3. Vinck, M., Batista-Brito, R., Knoblich, U. & Cardin, J. A. Arousal and locomotion make distinct contributions to cortical activity patterns and visual encoding. *Neuron* **86**, 740–754 (2015).
4. Treue, S. & Maunsell, J. H. Attentional modulation of visual motion processing in cortical areas MT and MST. *Nature* **382**, 539–541 (1996).
5. Erisken, S. et al. Effects of locomotion extend throughout the mouse early visual system. *Curr. Biol.* **24**, 2899–2907 (2014).
6. Reimer, J. et al. Pupil fluctuations track rapid changes in adrenergic and cholinergic activity in cortex. *Nat. Commun.* **7**, 13289 (2016).
7. Bennett, C., Arroyo, S. & Hestrin, S. Subthreshold mechanisms underlying state-dependent modulation of visual responses. *Neuron* **80**, 350–357 (2013).
8. Liang, L. et al. Retinal inputs to the thalamus are selectively gated by arousal. *Curr. Biol.* **30**, 3923–3934.e9 (2020).
9. McAdams, C. J. & Maunsell, J. H. Effects of attention on orientation-tuning functions of single neurons in macaque cortical area V4. *J. Neurosci.* **19**, 431–441 (1999).
10. Qiu, Y. et al. Natural environment statistics in the upper and lower visual field are reflected in mouse retinal specializations. *Curr. Biol.* **31**, 3233–3247.e6 (2021).
11. Rowell, C. H. Variable responsiveness of a visual interneurone in the free-moving locust, and its relation to behaviour and arousal. *J. Exp. Biol.* **55**, 727–747 (1971).
12. Chiappe, M. E., Seelig, J. D., Reiser, M. B. & Jayaraman, V. Walking modulates speed sensitivity in *Drosophila* motion vision. *Curr. Biol.* **20**, 1470–1475 (2010).
13. Busse, L. The influence of locomotion on sensory processing and its underlying neuronal circuits. *eNeuroforum* **24**, A41–A51 (2018).
14. Schneider, D. M. Reflections of action in sensory cortex. *Curr. Opin. Neurobiol.* **64**, 53–59 (2020).
15. Gerl, E. J. & Morris, M. R. The causes and consequences of color vision. *Evol. Educ. Outreach* **1**, 476–486 (2008).
16. Szél, A. et al. Unique topographic separation of two spectral classes of cones in the mouse retina. *J. Comp. Neurol.* **325**, 327–342 (1992).
17. Baden, T. et al. A tale of two retinal domains: near-optimal sampling of achromatic contrasts in natural scenes through asymmetric photoreceptor distribution. *Neuron* **80**, 1206–1217 (2013).
18. Walker, E. Y. et al. Inception loops discover what excites neurons most using deep predictive models. *Nat. Neurosci.* **22**, 2060–2065 (2019).
19. Lurz, K.-K. et al. Generalization in data-driven models of primary visual cortex. In *Proc. International Conference on Learning Representations* (2021).
20. Bashivan, P., Kar, K. & DiCarlo, J. J. Neural population control via deep image synthesis. *Science* **364**, eaav9436 (2019).
21. Franke, K. et al. An arbitrary-spectrum spatial visual stimulator for vision research. *eLife* **8**, e48779 (2019).
22. Liu, R. et al. An intriguing failing of convolutional neural networks and the CoordConv solution. In *Advances in Neural Information Processing Systems* (2018).
23. Rhim, I., Coello-Reyes, G., Ko, H.-K. & Nauhaus, I. Maps of cone opsin input to mouse V1 and higher visual areas. *J. Neurophysiol.* **117**, 1674–1682 (2017).
24. Denman, D. J., Siegle, J. H., Koch, C., Reid, R. C. & Blanche, T. J. Spatial organization of chromatic pathways in the mouse dorsal lateral geniculate nucleus. *J. Neurosci.* **37**, 1102–1116 (2017).
25. Rhim, I., Coello-Reyes, G. & Nauhaus, I. Variations in photoreceptor throughput to mouse visual cortex and the unique effects on tuning. *Sci. Rep.* **11**, 11937 (2021).
26. Fu, Y. et al. A cortical circuit for gain control by behavioral state. *Cell* **156**, 1139–1152 (2014).
27. Schröder, S. et al. Arousal modulates retinal output. *Neuron* **107**, 487–495.e9 (2020).
28. Eggermann, E., Kremer, Y., Crochet, S. & Petersen, C. C. H. Cholinergic signals in mouse barrel cortex during active whisker sensing. *Cell Rep.* **9**, 1654–1660 (2014).
29. Tikidji-Hamburyan, A. et al. Retinal output changes qualitatively with every change in ambient illuminance. *Nat. Neurosci.* **18**, 66–74 (2015).
30. Grimes, W. N., Schwartz, G. W. & Rieke, F. The synaptic and circuit mechanisms underlying a change in spatial encoding in the retina. *Neuron* **82**, 460–473 (2014).
31. Pennesi, M. E., Lyubarsky, A. L. & Jr. Pugh, E. N. Extreme responsiveness of the pupil of the dark-adapted mouse to steady retinal illumination. *Invest. Ophthalmol. Vis. Sci.* **39**, 2148–2156 (1998).
32. Safarani, S. et al. Towards robust vision by multi-task learning on monkey visual cortex. In *Advances in Neural Information Processing Systems* (2021).
33. Bialek, W., Rieke, F., de Ruyter van Steveninck, R. R. & Warland, D. Reading a neural code. *Science* **252**, 1854–1857 (1991).
34. Froudarakis, E. et al. Object manifold geometry across the mouse cortical visual hierarchy. Preprint at *bioRxiv* https://doi.org/10.1101/2020.08.20.258798 (2020).
35. Dadarlat, M. C. & Stryker, M. P. Locomotion enhances neural encoding of visual stimuli in mouse V1. *J. Neurosci.* **37**, 3764–3775 (2017).
36. Spitzer, H., Desimone, R. & Moran, J. Increased attention enhances both behavioral and neuronal performance. *Science* **240**, 338–340 (1988).
37. Wiersma, C. A. & Oberjat, T. The selective responsiveness of various crayfish oculomotor fibers to sensory stimuli. *Comp. Biochem. Physiol.* **26**, 1–16 (1968).
38. Maimon, G., Straw, A. D. & Dickinson, M. H. Active flight increases the gain of visual motion processing in *Drosophila*. *Nat. Neurosci.* **13**, 393–399 (2010).
39. Bezdudnaya, T. et al. Thalamic burst mode and inattention in the awake LGNd. *Neuron* **49**, 421–432 (2006).
40. de Gee, J. W. et al. Mice regulate their attentional intensity and arousal to exploit increases in task utility. Preprint at *bioRxiv* https://doi.org/10.1101/2022.03.04.482962 (2022).
41. Andermann, M. L., Kerlin, A. M., Roumis, D. K., Glickfeld, L. L. & Reid, R. C. Functional specialization of mouse higher visual cortical areas. *Neuron* **72**, 1025–1039 (2011).
42. Cronin, T. W. & Bok, M. J. Photoreception and vision in the ultraviolet. *J. Exp. Biol.* **219**, 2790–2801 (2016).
43. Hulburt, E. O. Explanation of the brightness and color of the sky, particularly the twilight sky. *J. Opt. Soc. Am.* **43**, 113–118 (1953).
44. Storchi, R. et al. Measuring vision using innate behaviours in mice with intact and impaired retina function. *Sci. Rep.* **9**, 10396 (2019).
45. Meyer, A. F., Poort, J., O'Keefe, J., Sahani, M. & Linden, J. F. A head-mounted camera system integrates detailed behavioral monitoring with multichannel electrophysiology in freely moving mice. *Neuron* **100**, 46–60.e7 (2018).
46. Wald, G. Human vision and the spectrum. *Science* **101**, 653–658 (1945).
47. Lamb, T. D. Why rods and cones? *Eye* **30**, 179–185 (2016).
48. Larsen, R. S. & Waters, J. Neuromodulatory correlates of pupil dilation. *Front. Neural Circuits* **12**, 21 (2018).
49. Douglas, R. H. The pupillary light responses of animals; a review of their distribution, dynamics, mechanisms and functions. *Prog. Retin. Eye Res.* **66**, 17–48 (2018).
50. Eberhardt, L. V., Grön, G., Ulrich, M., Huckauf, A. & Strauch, C. Direct voluntary control of pupil constriction and dilation: exploratory evidence from pupillometry, optometry, skin conductance, perception, and functional MRI. *Int. J. Psychophysiol.* **168**, 33–42 (2021).

Methods

Neurophysiological experiments

All procedures were approved by the Institutional Animal Care and Use Committee of Baylor College of Medicine. Owing to the explanatory nature of our study, we did not use randomization or blinding. No statistical methods were used to predetermine sample sizes.

Mice of either sex (*Mus musculus*, $n = 13$; 6 weeks to 5 months of age) expressing GCaMP6s in excitatory neurons through the Slc17a7-Cre and the Ai162 transgenic lines (stock numbers 023527 and 031562, respectively, The Jackson Laboratory) were anaesthetized, and a 4-mm craniotomy was made over the visual cortex of the right hemisphere as previously described[1,51]. For functional recordings, awake mice were head-mounted above a cylindrical treadmill, and calcium imaging was performed using a Ti-Sapphire laser tuned to 920 nm and a two-photon microscope equipped with resonant scanners (Thorlabs) and a ×25 objective (MRD77220, Nikon). The laser power after the objective was kept below 60 mW. The rostral–caudal treadmill movement was measured using a rotary optical encoder with a resolution of 8,000 pulses per revolution. We used light diffusing from the laser through the pupil to capture eye movements and pupil size. Images of the pupil were reflected through a hot mirror and captured with a GigE CMOS camera (Genie Nano C1920M; Teledyne Dalsa) at 20 fps at 1,920 × 1,200 pixel resolution. The contour of the pupil for each frame was extracted using DeepLabCut[52], and the centre and major radius of a fitted ellipse were used as the position and dilation, respectively, of the pupil.

For image acquisition, we used ScanImage. To identify V1 boundaries, we used pixelwise responses to drifting bar stimuli of a 2,400 × 2,400 μm scan at 200 μm depth from the cortical surface[53], recorded using a large-field-of-view mesoscope[54] not used for other functional recordings. In V1, imaging was performed using 512 × 512 pixel scans (650 × 650 μm) recorded at approximately 15 Hz and positioned within L2/3 at around 200 μm from the surface of the cortex. Imaging data were motion-corrected, automatically segmented and deconvolved using the CNMF algorithm[55]; cells were further selected by a classifier trained to detect somata based on the segmented masks. In addition, we excluded cells with low stimulus correlation. For this, we computed the first principal component (PC) of the response matrix of the size number of neurons × number of trials. For each neuron, we then estimated the linear correlation of its responses to the first PC, as the first PC captured unrelated background activity. We excluded neurons with a correlation lower or higher than −0.25 or 0.25, respectively. This resulted in 450–1,100 selected soma masks per scan depending on the response quality and the blood vessel pattern. A structural stack encompassing the scan plane and imaged at 1.6 × 1.6 × 1 μm *xyz* resolution with 20 repeats per plane was used to register functional scans of the same neurons into a shared *xyz* frame of reference. Cells registered to the same 3D stack were then anatomically matched for distances of <10 μm. For inception loop experiments, we confirmed the anatomical matching with a functional matching procedure, using the responses of cells to the same set of test images (see also ref. [18]) and only included anatomically matched neurons with a response correlation of >0.5 for further analysis. To bring different recordings of the same animal across the posterior–anterior axis of V1 into the same frame of reference, we manually aligned the mean image of each functional recording to the mean image of the 2,400 × 2,400 μm scan acquired at the mesoscope (see above) using the blood vessel pattern. Then, each cell within the functional scan was assigned a new *xy* coordinate (in μm) in the common frame of reference. To illustrate coarse differences across visual space, scan fields were manually assigned into three broad location categories within V1 (posterior, medial and anterior) depending on their position relative to V1 boundaries.

Visual stimulation

Visual stimuli were presented to the left eye of the mouse on a 42 × 26 cm light-transmitting Teflon screen (McMaster-Carr) positioned 12 cm from the animal, covering approximately 120 × 90° visual angle. Light was back-projected onto the screen by a DLP-based projector (EKB Technologies)[21] with UV (395 nm) and green (460 nm) LEDs that differentially activated mouse S-opsin and M-opsin. LEDs were synchronized with the scan retrace of the microscope. Note that the UV LED not only drives UV-sensitive S-opsin but also slightly activates green-sensitive M-opsin and rhodopsin because of their sensitivity tail for shorter wavelengths (β-band). This cross-activation could be addressed by using a silent substitution protocol, whereby one type of photoreceptor is selectively stimulated by presenting a steady excitation to all other photoreceptor types using a counteracting stimulus. However, this comes at the cost of overall contrast. We considered that our imperfect spectral separation of photoreceptor types was suitable to investigate most questions concerning chromatic processing in the visual system (discussed in ref. [21]), especially as photoreceptor-type-isolating stimulation in natural scenes is rare.

Light intensity (measured as the estimated photoisomerization rate, P^*(cone s^{-1})) was calibrated using a spectrometer (USB2000+, Ocean Optics) to result in equal activation rates for mouse M-opsin and S-opsin (for details see ref. [21]). In brief, the spectrometer output was divided by the integration time to obtain counts per s and then converted into electrical power (in nW) using the calibration data (in μJ per count) provided by Ocean Optics. The intensity (in μW) of the entire screen set to maximal intensity (255 pixel values) was approximately 1.28 and 1.39 for green and UV LEDs, respectively. To obtain the estimated photoisomerization rate per photoreceptor type, we first converted electrical power into energy flux (in eV s^{-1}) and then calculated the photon flux (in photons s^{-1}) using the photon energy (in eV). The photon flux density (in photons s^{-1} μm^{-2}) was then computed and converted into the photoisomerization rate using the effective activation of mouse cone photoreceptors by the LEDs and the light collection area of cone outer segments. In addition, we considered both the wavelength-specific transmission of the mouse optical apparatus[56] and the ratio between pupil size and retinal area[57]. See the calibration iPython notebook provided online (https://github.com/katrinfranke/open-visual-stimulator) for further details. For a pupil area of 0.2 mm^2 during quiet trials and maximal stimulus intensities (255 pixel values), this resulted in 400 P^*(cone s^{-1}) corresponding to the mesopic range. During active periods, the pupil area increased to 1.9 mm^2, resulting in 4,000 P^*(cone s^{-1}) corresponding to the low photopic regimen.

Before functional recordings, the screen was positioned such that the population receptive field across all neurons, estimated using an achromatic sparse noise paradigm, was within the centre of the screen. The screen position was fixed and kept constant across recordings of the same neurons. We used Psychtoolbox in MatLab for stimulus presentation and showed the following light stimuli.

Natural images. We presented naturalistic scenes from the available ImageNet online database[58]. We selected images on the basis of two criteria (Extended Data Fig. 1). First, to avoid an intensity bias in the stimulus, we selected images with no significant difference in the mean intensity of the blue and green image channels across all images. Second, we selected images with high pixelwise mean squared error (MSE > 85) across colour channels to increase chromatic contrast, resulting in a lower pixel-wise correlation across colour channels compared with a random selection. Then, we presented the blue and green image channels using the UV and green LEDs of the projector, respectively. For a single scan, we presented 4,500 unique coloured and 750 monochromatic images in UV and green, respectively. We added monochromatic images to the stimulus to include images without correlations across colour channels, thereby diversifying the input to the model. As the test set, we used 100 coloured and 2 × 25 monochromatic images that were repeated 10 times uniformly spread throughout the recording. Each image was presented for 500 ms, followed by a grey screen (UV and green LEDs at 127 pixel value) for 300–500 ms, sampled

uniformly from that range. The mean intensity of presented natural images across the green and UV colour channels varied between 5 and 204 (8-bit, gamma-corrected). For a small pupil during quiet states, this corresponded to approximately 8 and 320 photoisomerizations (P^*) per cone and second (P^*(cone s^{-1})). Each natural image was preceded by a grey blank period (all pixel values set to 127), which reduced the range of monitor intensities to approximately 57.2–213 P^*(cone s^{-1}) when integrating over 1 s, spanning less than one-order of magnitude. For the light intensities we were using, previous studies have found that the pupil size is relatively constant for changes in ambient light intensities below one-order of magnitude[31,59]. Indeed, we found that ambient monitor intensity does not contribute strongly to the recorded changes in pupil size (Extended Data Fig. 1).

Sparse noise. To map the receptive fields of V1 neurons, we used a sparse noise paradigm. UV and green bright (pixel value of 255) and dark (pixel value of 0) dots of approximately 10° visual angle were presented on a grey background (pixel value of 127) in a randomized order. Dots were presented for eight and five positions along the horizontal and vertical axis of the screen, respectively, excluding screen margins. Each presentation lasted 200 ms and each condition (for example, UV bright dot at position $x = 1$ and $y = 1$) was repeated 50 times. For a subset of recordings ($n = 2$ animals, $n = 3$ scan fields; compare with Extended Data Fig. 7e), each condition was repeated 150 times to increase the number of trials for more extreme behavioural states.

Full-field binary white noise. We used a binary full-field noise stimulus of UV and green LEDs to estimate temporal kernels of V1 neurons. The intensity of UV and green LEDs was determined independently by a balanced 15-min random sequence updated at 10 Hz. A similar stimulus was recently used in recordings of mouse[60] and zebrafish retina[61].

Coloured objects. To test for object discrimination, we used two synthesized objects rendered in Blender (https://www.blender.org) as previously described[34]. In brief, we smoothly varied object position, size, tilt and axial rotation. For bright objects, we also varied either the location or energy of four light sources. Stimuli were rendered as bright objects on a black screen and Gaussian noise in the other colour channel (condition 1), bright and dark objects on a grey screen and Gaussian noise in the other colour channel (conditions 2 and 3) or as bright objects on a black screen without Gaussian noise (condition 4). Per object and condition, we rendered movies of 875 s, which we then divided into 175 5-s clips. We presented the clips with different conditions and objects in a random order.

Images with dark objects. For the object detection task, we generated images with independent Perlin noise[62] in each colour channel using the perlin-noise package for Python (https://pypi.org/project/perlin-noise/). For all images except the noise images, we added a dark ellipse (pixel value of 0) of varying size, position and angle to one of the colour channels. We adjusted the contrast of all images with a dark object to match the contrast of noise images, such that the distribution of image contrasts did not differ between noise and object images. We presented 2,000 unique noise images and 2,000 unique images with a dark object in the UV and green image channels, respectively. Each image was presented for 500 ms, followed by a grey screen (UV and green LEDs at 127 pixel value) for 300–500 ms, sampled uniformly from that range.

For the presentation of naturalistic scenes and object movies and images, we applied a gamma function of 1.9 to the 8-bit pixel values of the monitor.

Preprocessing of neuronal responses and behavioural data

Neuronal responses were first deconvolved using constrained non-negative calcium deconvolution[55]. For all stimulus paradigms except the full-field binary white noise stimulus, we subsequently extracted the accumulated activity of each neuron between 50 ms after stimulus onset and offset using a Hamming window. For the presentation of objects, we segmented the 5-s clips into 9 bins of 500 ms, starting 250 ms after stimulus onset. Behavioural traces were extracted using the same temporal offset and integration window as deconvolved calcium traces. To train our models, we isotropically downsampled stimuli images to 64 × 36 pixels. Input images, the target neuronal activities, behavioural traces and pupil positions were normalized across the training set during training.

Pharmacological manipulations

To pharmacologically dilate and constrict the pupil, we applied 1–3% atropine and carbachol eye drops, respectively, to the left eye of the animal facing the screen for visual stimulation. Functional recordings started after the pupil was dilated or constricted. Pharmacological pupil dilation lasted >2 h, enabling the use of all the data for further analysis. By contrast, carbachol eye drops constricted the pupil for approximately 30 min and were re-applied once during the scan. For analysis, we only selected trials with constricted pupils and we matched data analysed in the control scans to the same trial numbers.

Sparse noise spatial receptive field mapping

We estimated spatial STAs of V1 neurons in response to the sparse noise stimulus by multiplying the stimulus matrix with the response matrix of each neuron[63] separately for each stimulus colour and polarity as well as behavioural state. For the behavioural state, we separated trials into small (<50th percentile) and large pupil trials (>75th percentile). We used different pupil size thresholds for the two behavioural states compared to the model owing to the shorter recording time. For recordings with pupil dilation, we used locomotion speed instead of pupil size to separate trials into two behavioural states. For each behavioural state, STAs computed on the basis of on and off dots were averaged to produce one STA per cell and stimulus colour. Green and UV STAs of the same behavioural state were peak-normalized to the same maximum. To assess STA quality, we generated response predictions by multiplying the flattened STA of each neuron with the flattened stimulus frames and compared the predictions to the recorded responses by estimating the linear correlation coefficient. For analysis, we only included cells for which the correlation was >0.2 for at least one of the stimulus conditions.

In contrast to the modelling results, the STA spectral contrast for a quiet state varied only slightly across the anterior–posterior axis of the V1. This was probably due to the different pupil size thresholds for quiet and active state used in the STA paradigm compared to the model. To verify this, we used the data in response to natural images (Fig. 2) to train a separate model without behaviour as input channels on trials with small pupil (<50th percentile) and subsequently optimized MEIs, which is a procedure more similar to the STA paradigm. When looking at the spectral contrast of the resulting MEIs, we observed a smaller variation of colour preference across the anterior–posterior axis of V1, thereby confirming our prediction (data not shown).

To confirm that the shift in colour preference with behaviour in response to the sparse noise was not dependent on the specific pupil size thresholds we used, we presented 150 instead of 50 repeats per stimulus condition in a subset of experiments. The larger number of trials for more extreme behavioural states allowed us to compute STAs for behavioural states more similar to the model (<20th versus >85th percentile). This resulted in a stronger shift in colour preference during active periods compared with the lower thresholds of pupil sizes (data not shown), which indicated that we had probably underestimated the effect for the shorter recordings shown in Extended Data Fig. 7a–c.

Full-field binary noise temporal receptive field mapping

We used the responses to the 10 Hz full-field binary noise stimulus of UV and green LEDs to compute temporal STAs of V1 neurons. Specifically,

we upsampled both stimulus and responses to 60 Hz and then multiplied the stimulus matrix with the response matrix of each neuron. Per cell, this resulted in a temporal STA in response to UV and green flickers, respectively. The kernel quality was measured by comparing the variance of each temporal STA with the variance of the baseline, defined as the first 100 ms of the STA. Only cells with at least five times more variance of the kernel compared with baseline were considered for further analysis.

Simulated data using Gabor neurons

We simulated neurons with Gabor receptive fields with varying Gabor parameters across the two colour channels. We normalized each Gabor receptive field to have a background of 0 and an amplitude range between −1 and 1. To generate responses of simulated neurons, we used the same set of training images presented during functional recordings. First, we subtracted the mean across all images from the training set, multiplied each Gabor receptive field with each training image and computed the sum of each multiplication across the two colour channels c. We then passed the resulting scalar response per neuron through a rectified linear unit (ReLU) to obtain the simulated response r, such that

$$r = \mathrm{ReLU}\left(\sum_{c,x,y} \mathrm{image}_{c,x,y}\mathrm{Gabor}_{c,x,y}\right),$$

where

$$\mathrm{Gabor}_{c,x,y} = \alpha_c \exp\left(-\frac{x'^2 + \gamma_c^2 y'^2}{2\sigma_c^2}\right)\cos\left(2\pi\frac{x'}{\lambda} + \psi_c\right)$$

with $x' = x\cos(\theta_c) + y\sin(\theta_c)$ and $y' = -x\sin(\theta_c) + y\cos(\theta_c)$. We varied orientation θ, size σ, spatial aspect ratio γ, phase ψ and colour preference α independently for each colour channel and neuron, while keeping spatial frequency λ constant across all neurons. Finally, we passed the simulated responses r through a Poisson process and normalized the responses by the respective standard deviation of the responses across all images. We used the responses of the simulated Gabor neurons together with the natural images to train the model (see below). Our model recovered both the colour opponency and the colour preference of simulated neurons. Only extreme colour preferences were slightly underestimated by our model, which is probably due to high correlations across the colour channels of natural scenes.

In silico tuning characterization

It has been our main interest to investigate the change in tuning properties with the behavioural state of animals. Ideally, this includes manipulating the behaviour of the animal and investigating the resulting effect on different visual tuning properties. Although this is experimentally challenging and time-consuming, it is straightforward with a deep-learning-based neuronal predictive model that emulates the biological circuit. This allowed us to selectively study how tuning to colour or spatial features changes with behaviour. To perform our in in silico tuning characterization, we created a CNN model, which was split into two parts: the core and the readout. The core computed latent features from the inputs, which were shared among all neurons. The readout was learned per neuron and mapped the output features of the core onto the neuronal responses through regularized regression.

Core of the CNN model. We based our model on the work from ref. [19], as it was demonstrated to set the state of the art for predicting the responses of a population of mouse V1 neurons. In brief, we modelled the core as a 4-layer CNN, with 64 feature channels per layer. Each layer consisted of a 2D convolutional layer followed by a batch-normalization layer and ELU nonlinearity[64,65]. Except for the first layer, all convolutional layers were depth-separable convolutions[66], which led to better performance while reducing the number of core parameters. Each depth-separable layer consisted of a 1×1 pointwise convolution followed by a 7×7 depth-wise convolution, again followed by a 1×1 pointwise convolution. Without stacking the outputs of the core, the output tensor of the last layer was passed on to the readout.

Readout of the CNN model. To obtain the scalar firing rate for each neuron, we computed a linear regression between the core output tensor of dimensions $\mathbf{x} \in \mathbb{R}^{w \times h \times c}$ (w, width; h, height; c, channels) and the linear weight tensor $\mathbf{w} \in \mathbb{R}^{c \times w \times h}$, followed by an exponential linear unit (ELU) offset by one (ELU+1) to keep the response positive. We made use of the recently proposed Gaussian readout[19], which considerably simplifies the regression problem. Our Gaussian readout learned the parameters of a 2D Gaussian distribution $\mathcal{N}(\mu_n, \Sigma_n)$ and sampled a location of height and width in the core output tensor in each training step for every image and neuron. Given a large enough initial Σ_n to ensure gradient flow, Σ_n, that is, the uncertainty about the readout location, decreased during training for more reliable estimates of the mean location μ_n, which represented the centre of the receptive field of a neuron. At inference time (that is, when evaluating our model), we set the readout to be deterministic and to use the fixed position μ_n. We therefore learned a position of a single point in core feature space for each neuron. In parallel to learning the position, we learned the weights of the weight tensor of the linear regression of size c per neuron. Furthermore, we made use of the retinotopic organization of V1 by coupling the recorded cortical 2D coordinates $\mathbf{p}_n \in \mathbb{R}^2$ of each neuron with the estimation of the receptive field position μ_n of the readout. We achieved this by learning a common function $\mu_n = f(\mathbf{p}_n)$, which is shared by all neurons. We set f to be a randomly initialized linear fully connected network of size 2-2 followed by tanh nonlinearity.

Shifter network. Because we used a free viewing paradigm when presenting the visual stimuli to the head-fixed mice, the receptive field positions of the neurons with respect to the presented images had considerable trial-to-trial variability due to eye movements. To inform our model of the trial-dependent shift of the receptive fields of neurons, we shifted μ_n, the receptive field centre of the model neuron, using the estimated pupil centre (see the section 'Neurophysiological experiments'). We accomplished this by passing the pupil centre through a small shifter network, a three-layer fully connected network with $n = 5$ hidden features, again followed by a tanh nonlinearity, that calculates the shift Δx and Δy per trial. The shift was then added to μ_n of each model neuron.

Input of behaviour and image position encoding. In addition to the green and UV channels of the visual stimulus, we appended five extra channels to each input to the model. We added three channels of the recorded behavioural parameters in each given trial (pupil size, instantaneous change of pupil size and locomotion speed), such that each channel simply consisted of the scalar for the respective behavioural parameter, transformed into the stimulus dimensions. This enabled the model to predict neuronal responses as a function of both visual input and behaviour and therefore to learn the relationship between behavioural states and neuronal activity. This modification enabled us to investigate the effect of behaviour by selecting different inputs in the behavioural channels while optimizing the image channels. Furthermore, we added a positional encoding to the inputs, which consisted of two channels that encoded the horizontal and vertical pixel positions of the visual stimulus. These encodings can be thought of as simple greyscale gradients in either direction, with values from [−1, ..., 1]. Appending position encodings of this kind has been shown to improve the ability of CNNs to learn spatial relationships between pixel positions of the input image and high level feature representations[22]. We found that including

the position embedding increased the performance of our model (Extended Data Fig. 2b). We also observed a smoother gradient of colour tuning across the different scan fields (Fig. 2b and Extended Data Fig. 6b) when adding the position encoding. This indicated that the model learned the well-described colour sensitivity tuning of mouse cone photoreceptors across visual space.

Model training and evaluation

We first split the unique training images into the training and validation set, using a split of 90% to 10%, respectively. Then we trained our networks with the training set by minimizing the Poisson loss $\frac{1}{m}\sum_{i=1}^{m}(\hat{r}^{(i)} - r^{(i)}\log\hat{r}^{(i)})$, where m denotes the number of neurons, $\hat{r}$ the predicted neuronal response and r the observed response. After each full pass through the training set (that is, epoch), we calculated the correlation between the predicted and the measured responses across all neurons on the validation set: if the correlation failed to increase during a fixed number of epochs, we stopped the training and restored the model to its state after the best performing epoch. After each stopping, we either decreased the learning rate or stopped training altogether if the number of learning-rate decay steps was reached. We optimized the parameters of the network through stochastic gradient descent using the Adam optimizer[67]. Furthermore, we performed an exhaustive hyperparameter selection using a Bayesian search on a held-out dataset. All parameters and hyperparameters can be found in our GitHub repository (see the Code availability section). When evaluating our models on the test set (Extended Data Fig. 2a–c), we used two different types of correlation. First, referred to as test correlation, we computed the correlation between the prediction by the model and neuronal responses across single trials, including the trial-by-trial variability across repeats. Second, we computed the correlation of the predicted responses with the average responses across repeats and refer to it here as the correlation to average. We also computed the fraction of variance explained, using $\hat{r}^2_{\mathrm{ER}}$ proposed in ref. [68], which provides an unbiased estimate of the variance explained based on the expected neuronal response across image repetitions. However, our model computed different predictions for each repetition of a given test set image because we also fed the behavioural parameters of each trial into the model. We therefore simply averaged the model responses across repetitions and calculated the $\hat{r}^2_{\mathrm{ER}}$ accordingly. When evaluating the model performance for the pharmacology conditions (Extended Data Fig. 2c), we found that they led to a lower model performance compared with the control condition. This could be due to the fact that for the dilated condition, we did not incorporate pupil-related behavioural parameters into the model owing to difficulties in pupil tracking for this pharmacological condition. For the drug condition with carbachol, we selected a subset of trials in which the pupil was constricted (see the 'Pharmacological manipulations' section), which led to fewer trials to train the models with. Finally, for some of our datasets that had either a low number of trials or a low yield of neurons, we trained a single model on multiple datasets[19], such that the convolutional core of the model was trained with more examples. The training of the per-neuron readout was unaffected by this joint training of datasets. We assigned a model identifier to each trained model (which can be found in Supplementary Table 1) such that datasets that were trained together in one model could be easily identified.

Ensemble models

For all analyses and for the generation of MEIs, we used an ensemble of models rather than individual models. Instead of training just one model for each dataset, we trained ten individual models that were initialized with different random seeds. We then selected the five best models as measured by their performance on the validation set to be part of a model ensemble. The inputs to the ensemble model were passed to each member, and the resulting predictions were averaged to obtain the final model prediction.

Generation of MEIs

We used a variant of regularized gradient ascent on our trained deep neural network models to obtain a MEI image for each neuron, given by $\mathbf{x} \in \mathbb{R}^{h\times w\times c}$. Because of our particular model inputs (see the section 'Input of behavioural parameters and image position encoding'), each MEI, like the natural images used for training, had seven channels of which we optimized only the first two: the green and UV colour channels. To obtain MEIs, we initialized a starting image with Gaussian white noise. We set the behavioural channels of the starting image to the desired behavioural values (usually <3rd and >97th percentile for quiet and active states, respectively). In addition, we set the position channels to the default position encoding. Then, in each iteration of our gradient ascent method, we presented the image to the model and computed the gradients of the first two image channels (green and UV) with respect to the model activation of a single neuron. During gradient descent optimization, we smoothed the gradient by applying Gaussian blur with a σ of 1 pixel. To constrain the contrast of the image, we calculated the Euclidean (L2) norm of the resulting MEI

$$\|\mathrm{MEI}\|_2 := \sqrt{\sum_{i=1}^{c}\sum_{j=1}^{w}\sum_{k=1}^{h}\mathrm{MEI}_{ijk}^2}$$

across all pixels MEI_{ijk} of the two colour channels c and compared the L2 norm to a fixed norm budget b, which we set to 10. The norm budget can be effectively thought of as a contrast constraint. An L2 norm of 10, calculated across all pixel intensities of the image, proved to be optimal such that the resulting MEI had minimal and maximal values similar to those found in our training natural image distribution. If the image exceeded the norm budget during optimization, we divided the entire image by factor f_{norm} with $f_{\mathrm{norm}} = \|\mathrm{MEI}\|_2/b$. Additionally, we made sure that the MEI could not contain values outside the 8-bit pixel range by clipping the MEI outside these bounds, corresponding to 0 or 255 pixel intensity. As an optimizer, we used stochastic gradient descent with a learning rate of 3. We ran each optimization for 1,000 iterations, without an option for early stopping. Our analyses showed that the resulting MEIs were spatially highly correlated across behavioural states (Extended Data Fig. 5a–c). To validate this finding, we performed a control experiment using two separate models exclusively trained on trials from active or quiet states. We again split the trials into quiet and active periods using pupil size (quiet, <50th percentile; active, >75th percentile). When inspecting the MEIs generated from these two models, we found that the MEIs were again highly correlated across colour channels, albeit less than for the model that was trained on the entire data. This can partially be explained by the limited amount of data for the model trained with trials from the active state that occurred less frequently in our data. Furthermore, we found that the spatial structure of MEIs of anatomically matched neurons across the control and pharmacology conditions was highly similar, which suggested that the two models trained separately both converged on the same tuning properties, despite differences in the prediction performance (Extended Data Fig. 2)

Spectral contrast

For estimating the chromatic preference of the recorded neurons, we used spectral contrast (SC). It is estimated as a Michelson contrast ranging from −1 to 1 for a neuron responding solely to UV and green image contrast, respectively. We decided to quantify the spectral sensitivity in relative terms for each behavioural state because visual responses to both green and UV stimuli are gain modulated in an active state. Therefore, interpretation of absolute response amplitudes to UV and green stimuli across behavioural states can be challenging. See Extended Data Fig. 6a,d for an illustration of how responses to stimuli of diverse spectral contrasts are gain modulated during an active state. We define SC as

$$SC = \frac{r_{green} - r_{UV}}{r_{green} + r_{UV}}$$

where r_{green} and r_{UV} correspond to the following criteria: (1) the norm of the green and UV MEI channels to estimate the chromatic preference of neurons in the context of naturalistic scenes; (2) the amplitude (mean of all pixels >90th percentile) of the UV and green spatial STAs to estimate the chromatic preference of neurons in the context of the sparse noise paradigm; (3) the norm of the green and UV channels of reconstructed images to quantify chromatic preference at a populational level; and (4) the norm of the green and UV channels of simulated Gabor receptive fields to obtain each simulated chromatic preference of neurons.

In silico colour-tuning curves

To generate in silico colour-tuning curves for recorded V1 neurons, we systematically varied the L2-norm of the green and UV MEI channels while keeping the overall norm across colour channels constant (with norm = 10). We used $n = 50$ spectral contrast levels, ranging from all contrast in the UV channel to all contrast in the green channel. We then presented the modified MEIs to the model and plotted the predicted responses across all $n = 50$ spectral contrast levels. Modified MEIs were either presented to the model for a quiet or active state (see also above).

Temporal dynamics of shift in colour tuning with behaviour

To investigate the timescale of the shift in colour selectivity with behaviour, we tested how fast we could observe the shift after a transition from a quiet to an active behavioural state. To achieve this, we identified state changes from quiet to active periods by detecting rapid increases in pupil size above a certain threshold (>95th percentile of differentiated pupil size trace) after a prolonged quiet state period (>5 s below the 50th percentile of pupil size). Results were consistent across varying thresholds (data not shown). We then sampled active trials with pupil sizes >75th percentile of pupil size for varying readout windows (1, 2, 3, 5 and 10 s) after that state change. Model training was performed on all quiet trials (<50th percentile of pupil size) and the selection of active trials. MEIs and STAs were then estimated as described above.

Reconstruction analysis

We visualized which image features the population of model neurons are sensitive to by using a new resource-constrained image reconstruction method based on the responses of a population of model neurons[32]. The reasoning behind the resource-constrained reconstruction is to recreate the responses of a population of neurons when presented with a target image by optimizing a new image and matching the responses of neurons given that new image as close as possible to the responses of the target image. By limiting the image contrast of the reconstructed image during the optimization, the reconstructions will only contain the image features that are most relevant to recreate the population responses, thereby visualizing the sensitivities and invariances of the population of neurons. As target images for our reconstruction, we chose natural images from our test set. For each reconstruction, we first calculated the responses $f(\mathbf{x}_0)$ of all model neurons when presented with target image $\mathbf{x}_0$. We then initialized an image ($\mathbf{x}$) with Gaussian white noise as the basis for reconstruction of the target image by minimizing the squared loss between the target responses and the responses from the reconstructed image $\ell(\mathbf{x}_0, \mathbf{x}) = \|f(\mathbf{x}) - f(\mathbf{x}_0)\|^2$ subject to a norm constraint. In this work, we set the contrast (that is, L2-norm, see section 'Generation of MEIs' for details) of the reconstructions to 40, which corresponds to about 60% of the average norm of our natural image stimuli. We chose this value to be high enough to still allow for qualitative resemblance between the reconstructed image and the target while keeping the constraint tight enough to avoid an uninformative trivial solution; that is, the identical reconstruction of the target. We improved the quality of the reconstructions by using an augmented version of our model, which reads out each neuronal response not from the actual receptive field position μ of the model neuron (see 'Readout' for details), but from all height × width positions in feature space, except the $n = 10$ pixels around each border to avoid padding artefacts. This yielded 18 × 46 = 828 copies per neuron, and with the $N = 478$ original model neurons of mouse 1 in Extended Data Fig. 9c, this resulted in overall $n = 395{,}784$ augmented neurons for our reconstruction analyses. A stochastic gradient descent with a learning rate of 1,000 produced the qualitatively best reconstructions, resulting in images with the least amount of noise. We always optimized for 5,000 steps per image, without the early stopping step of the optimization process.

Decoding analysis

We used a SVM classifier with a radial basis function kernel to estimate the decoding accuracy between the neuronal representations of two stimulus classes: either object 1 and object 2 (object discrimination) or dark object and no object (object detection). We used all neurons recorded within one scan and built four separate decoders for UV and green stimuli and small and large pupil trials, respectively. Then we trained each decoder with randomly selected training trials (usually 176 trials, but only 60–126 trials for $n = 3$ scans owing to the lower number of trials with locomotion activity), tested its accuracy with randomly selected test trials (15% of train trials) and computed the mean accuracy across $n = 10$ different training–test trial splits. Finally, we converted the decoding accuracy into discriminability, the mutual information (MI) between the true class and its estimate using

$$MI(c, \hat{c}) = \sum_i \sum_j P_{ij} \log_2 \frac{P_{ij}}{P_{i\cdot} P_{\cdot j}}$$

where P_{ij} is the probability of observing the true class i and predicted class j and $P_{i\cdot}$ and $P_{\cdot j}$ denote the respective marginal probabilities.

To quantify the significance for each animal, we compared the observed shift in decoding performance of UV versus green objects across behavioural states per animal with a distribution of shifts ($n = 500$) obtained when shuffling the labels of quiet and active trials using bootstrapping. Specifically, we sampled half of the training data and test data from quiet trials, and the other half from active trials at random. We then trained SVMs to compute the decoding accuracy based on this particular shuffling. We repeated this $n = 500$ times and obtained a P value by computing the upper quantile of the real shift given the distribution of shifts obtained when shuffling the behavioural states.

Response reliability

We calculated the signal-to-noise ratio (SNR)[68] as our measure for response reliability. It is defined as follows:

$$SNR = \frac{\frac{1}{m} \sum_{i=1}^{m} (\mu_i - \bar{\mu})^2}{\hat{\sigma}^2}$$

The SNR expresses the ratio of the variance in the expected responses against trial-by-trial variability across repeats. Here, μ_i corresponds to the expected response to the ith stimulus, with the average expected response given as

$$\bar{\mu} = \frac{1}{m} \sum_{i=1}^{m} \mu_i$$

The trial-by-trial variance $\hat{\sigma}^2$ was computed by averaging the variance across repeats over all stimuli. We assumed that $\hat{\sigma}^2$ is constant across all responses to different stimuli. This is achieved by a variance stabilizing transform of the responses r, for which we used the Anscombe transformation. We therefore obtained the transformed responses $\hat{r}$ as follows:

$$\hat{r} = 2\sqrt{r + \frac{3}{8}}$$

The SNR is a reliable estimate of data quality for neuronal responses across diverse recording modalities and brain regions[68].

Statistical analysis

We used generalized additive models (GAMs) to analyse the relationship of MEI spectral contrast, cortical position and behavioural state (see Supplementary Methods for details). GAMs extend the generalized linear model by allowing the linear predictors to depend on arbitrary smooth functions of the underlying variables[69]. In practice, we used the mgcv-package for R to implement GAMs and perform statistical testing. For all other statistical tests, we used Wilcoxon signed-rank test and two-sampled or one-sampled *t*-test.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The stimulus images and neuronal data used in this paper are stored at https://gin.g-node.org/cajal/Franke_Willeke_2022.

Code availability

Our coding framework uses general tools such as PyTorch, Numpy, scikit-image, matplotlib, seaborn, DataJoint[70], Jupyter and Docker. We also used the following custom libraries and code: neuralpredictors (https://github.com/sinzlab/neuralpredictors) for torch-based custom functions for model implementation; nnfabrik (https://github.com/sinzlab/nnfabrik) for automatic model training pipelines using DataJoint; nndichromacy for utilities, (https://github.com/sinzlab/nndichromacy); and mei (https://github.com/sinzlab/mei) for stimulus optimization.

51. Froudarakis, E. et al. Population code in mouse V1 facilitates readout of natural scenes through increased sparseness. *Nat. Neurosci.* **17**, 851–857 (2014).
52. Mathis, A. et al. DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. *Nat. Neurosci.* **21**, 1281–1289 (2018).
53. Garrett, M. E., Nauhaus, I., Marshel, J. H. & Callaway, E. M. Topography and areal organization of mouse visual cortex. *J. Neurosci.* **34**, 12587–12600 (2014).
54. Sofroniew, N. J., Flickinger, D., King, J. & Svoboda, K. A large field of view two-photon mesoscope with subcellular resolution for in vivo imaging. *eLife* **5**, e14472 (2016).
55. Pnevmatikakis, E. A. et al. Simultaneous denoising, deconvolution, and demixing of calcium imaging data. *Neuron* **89**, 285–299 (2016).
56. Henriksson, J. T., Bergmanson, J. P. G. & Walsh, J. E. Ultraviolet radiation transmittance of the mouse eye and its individual media components. *Exp. Eye Res.* **90**, 382–387 (2010).
57. Schmucker, C. & Schaeffel, F. A paraxial schematic eye model for the growing C57BL/6 mouse. *Vision Res.* **44**, 1857–1867 (2004).
58. Russakovsky, O. et al. ImageNet large scale visual recognition challenge. *Int. J. Comput. Vis.* **115**, 211–252 (2015).
59. Grozdanic, S. et al. Characterization of the pupil light reflex, electroretinogram and tonometric parameters in healthy mouse eyes. *Curr. Eye Res.* **26**, 371–378 (2003).
60. Szatko, K. P. et al. Neural circuits in the mouse retina support color vision in the upper visual field. *Nat. Commun.* **11**, 3481 (2020).
61. Yoshimatsu, T., Schröder, C., Nevala, N. E., Berens, P. & Baden, T. Fovea-like photoreceptor specializations underlie single UV cone driven prey–capture behavior in zebrafish. *Neuron* **107**, 320–337.e6 (2020).
62. Perlin, K. An image synthesizer. *SIGGRAPH Comput. Graph.* **19**, 287–296 (1985).
63. Schwartz, O., Pillow, J. W., Rust, N. C. & Simoncelli, E. P. Spike-triggered neural characterization. *J. Vis.* **6**, 484–507 (2006).
64. Ioffe, S. & Szegedy, C. Batch normalization: accelerating deep network training by reducing internal covariate shift. In *Proc. 32nd International Conference on Machine Learning* (2015).
65. Clevert, D.-A., Unterthiner, T. & Hochreiter, S. Fast and accurate deep network learning by exponential linear units (ELUs). In *Proc. International Conference on Learning Representations* (2016).
66. Chollet, F. Xception: Deep learning with depthwise separable convolutions. In *Proc. 30th IEEE Conference on Computer Vision and Pattern Recognition* (2017).
67. Kingma, D. P. & Ba, J. Adam: a method for stochastic optimization. In *Proc. International Conference on Learning Representations* (2015).
68. Pospisil, D. A. & Bair, W. The unbiased estimation of the fraction of variance explained by a model. *PLoS Comput. Biol.* **17**, e1009212 (2021).
69. Wood, S. N. *Generalized Additive Models: An Introduction with R* (Chapman and Hall/CRC, 2006).
70. Yatsenko, D. et al. DataJoint: managing big scientific data using MATLAB or Python. Preprint at *bioRxiv* https://doi.org/10.1101/031658 (2015).
71. Tan, Z., Sun, W., Chen, T.-W., Kim, D. & Ji, N. Neuronal representation of ultraviolet visual stimuli in mouse primary visual cortex. *Sci. Rep.* **5**, 12597 (2015).
72. Mouland, J. W. et al. Extensive cone-dependent spectral opponency within a discrete zone of the lateral geniculate nucleus supporting mouse color vision. *Curr. Biol.* **31**, 3391–3400.e4 (2021).

Acknowledgements We thank G. Horwitz, T. Euler, M. Mathis, T. Baden, L. Höfling and Y. Qiu for feedback on the manuscript and D. Kim, D. Sitonic, D. Tran, Z. Ding, K. Lurz, M. Bashiri, C. Blessing and E. Walker for technical support and helpful discussions. We also thank the International Max Planck Research School for Intelligent Systems (IMPRS-IS) for supporting Konstantin F. Willeke. This work was supported by the Carl-Zeiss-Stiftung (to F.H.S.), the DFG Cluster of Excellence 'Machine Learning—New Perspectives for Science' (to F.H.S.; EXC 2064/1, project number 390727645), an AWS Machine Learning research award (to F.H.S.), the Intelligence Advanced Research Projects Activity (IARPA) through the Department of Interior/Interior Business Center (DoI/IBC) contract number D16PC00003 (to A.S.T.), grant R01 EY026927 (to A.S.T.), grant U01 UF1NS126566 (to A.T.), a NEI/NIH Core Grant for Vision Research (P30EY002520) and an NSF NeuroNex grant 1707400 (to A.S.T.). The US Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright annotation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of IARPA, DoI/IBC, or the US Government.

Author contributions K.F.: conceptualization, methodology, validation, software, formal analysis, investigation, writing (original draft), visualization, supervision and project administration. K.F.W.: conceptualization, methodology, validation, software, formal analysis, investigation, writing (original draft), visualization and data curation. K.P.: investigation, validation and writing (reviewing and editing). M.G.: investigation and validation. N.Z.: investigation and methodology. T.M.: investigation. S.P.: methodology, software, validation and writing (reviewing and editing). E.F.: methodology, investigation and writing (reviewing and editing). J.R.: validation and writing (reviewing and editing). F.H.S.: conceptualization, writing (reviewing and editing), methodology, software, data curation, supervision and funding acquisition. A.S.T.: conceptualization, experimental and analysis design, supervision, funding acquisition and writing (reviewing and editing).

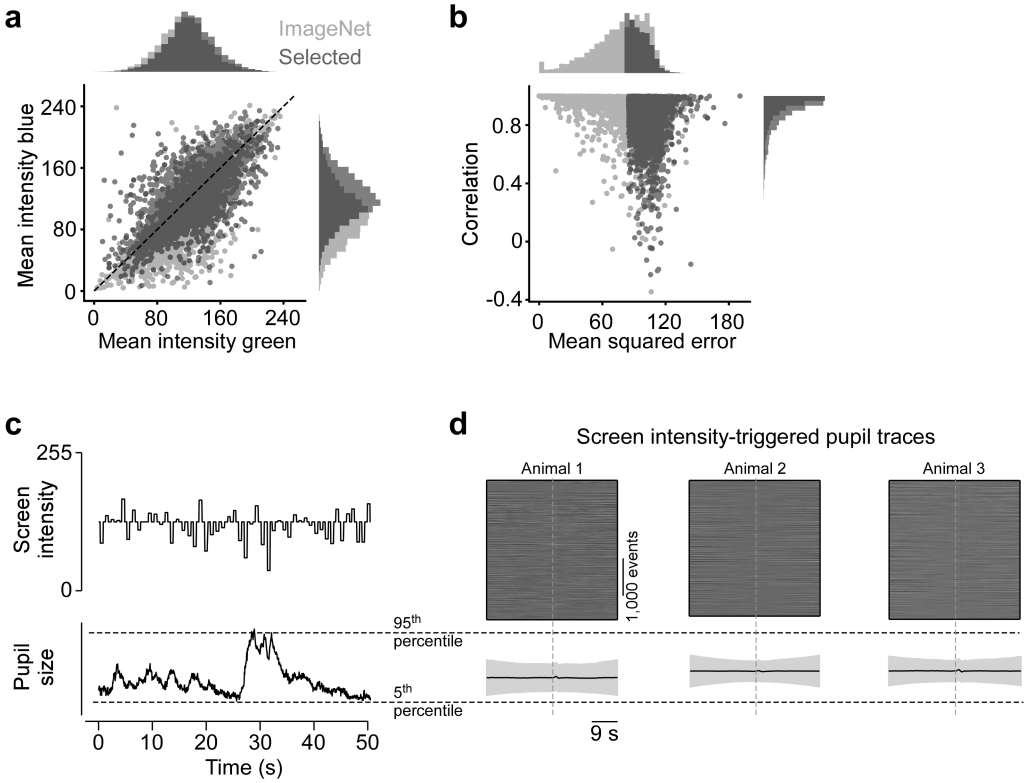
Competing interests The authors declare no competing interests.

Additional information

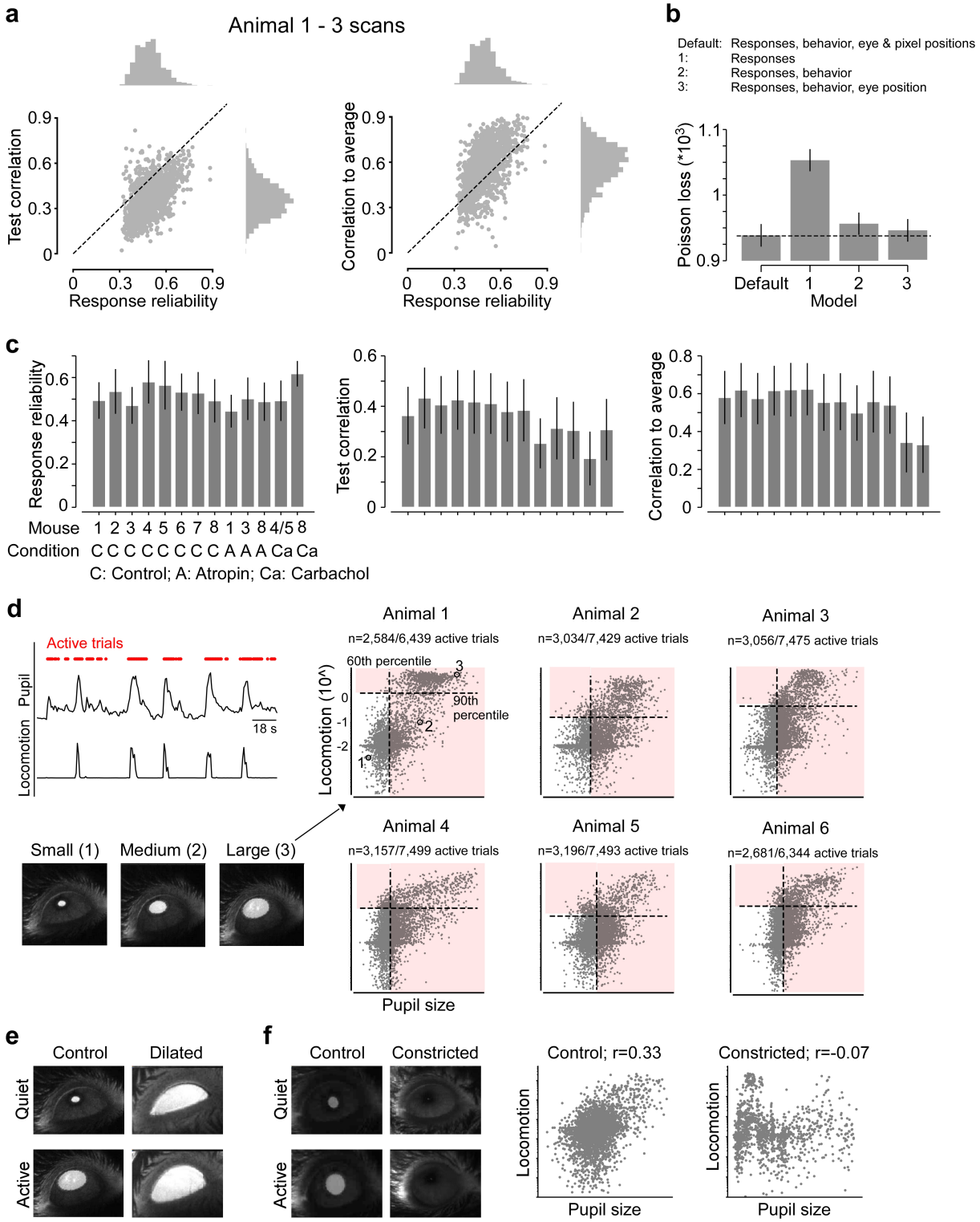
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41586-022-05270-3.

Correspondence and requests for materials should be addressed to Katrin Franke.

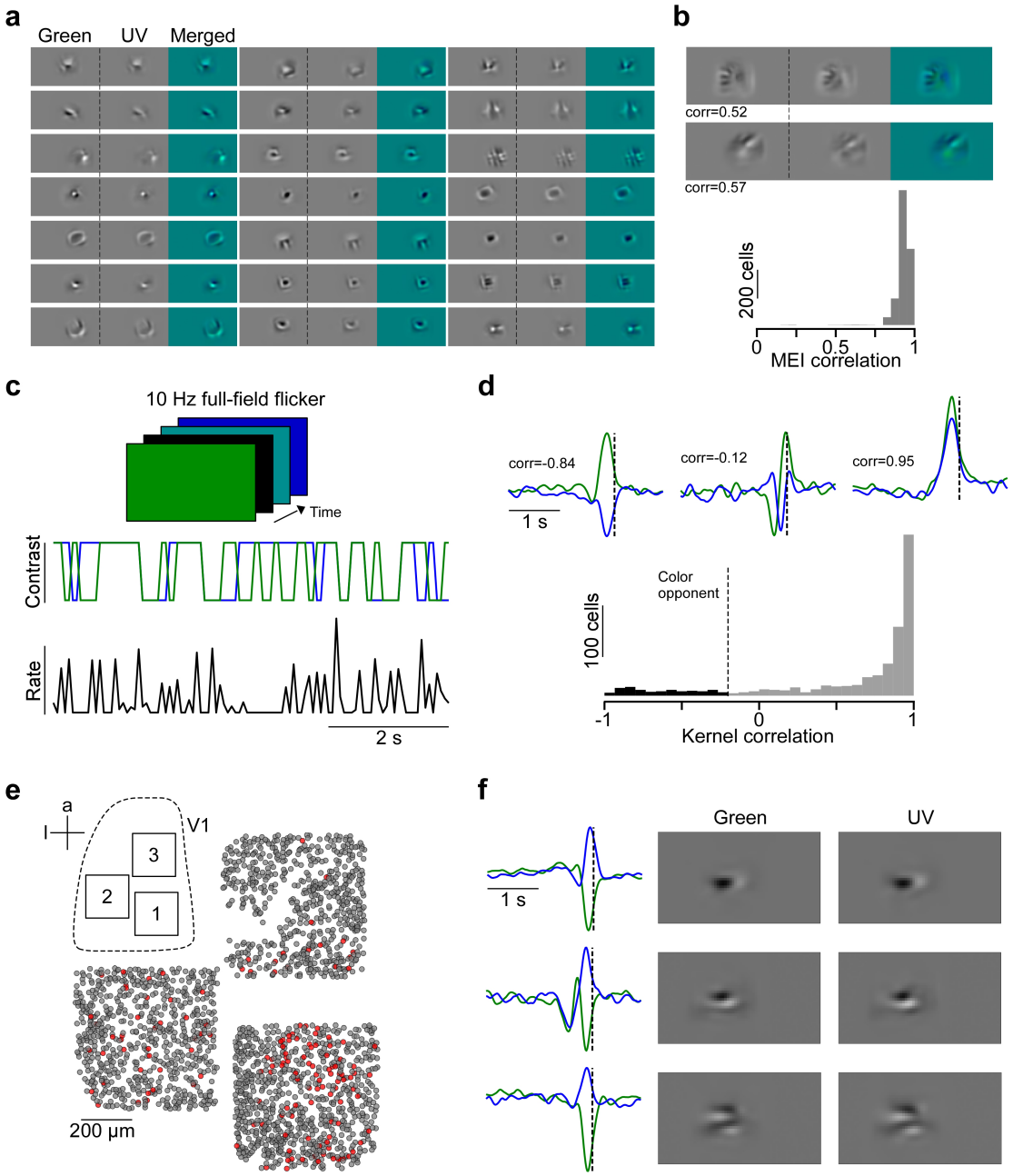
Peer review information *Nature* thanks Najib Majaj, Nathalie Rochefort and Aman Saleem for their contribution to the peer review of this work. Peer reviewer reports are available.



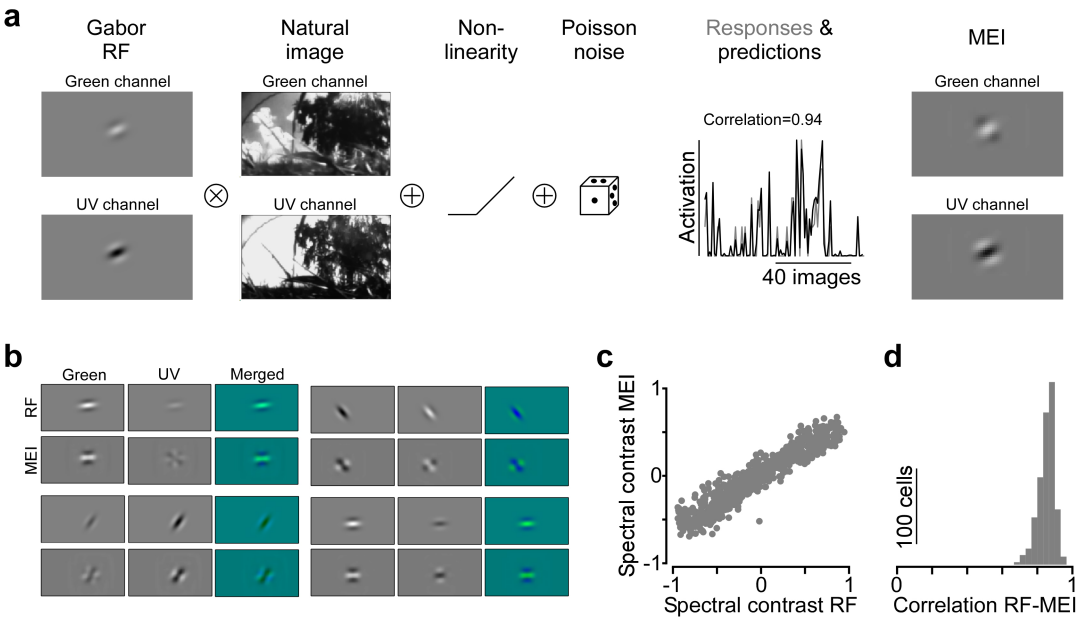
Extended Data Fig. 1 | Selection of coloured naturalistic scenes and pupil changes with monitor intensity. a, Mean intensity in 8-bit pixel space of green and blue channel of randomly sampled ImageNet images (light gray; n=6.000) and selected images (dark gray; n=6.000). Images were selected such that the distribution of mean intensities of blue and green image channels were not significantly different. Selected images can be downloaded from the online repository (see Data Availability in Methods section). **b**, Distribution of correlation and mean squared error (MSE) across green and blue image channels. To increase chromatic content, only images with MSE > 85 were selected for visual stimulation. **c**, Mean screen intensity (top) and pupil size changes (bottom) for n=50 trials. Dotted lines in the bottom indicate 5th and 95th percentile, respectively. **d**, Screen-intensity triggered pupil traces (top) for n=3 scans performed in different animals. Vertical dotted line indicates time point of screen intensity increase. Bottom shows mean change in pupil size (black; s.d. shading in gray) upon increase in screen intensity. Compared to pupil dilation induced by the behavioural state, the changes in monitor intensity over time only elicited minor changes in pupil size.



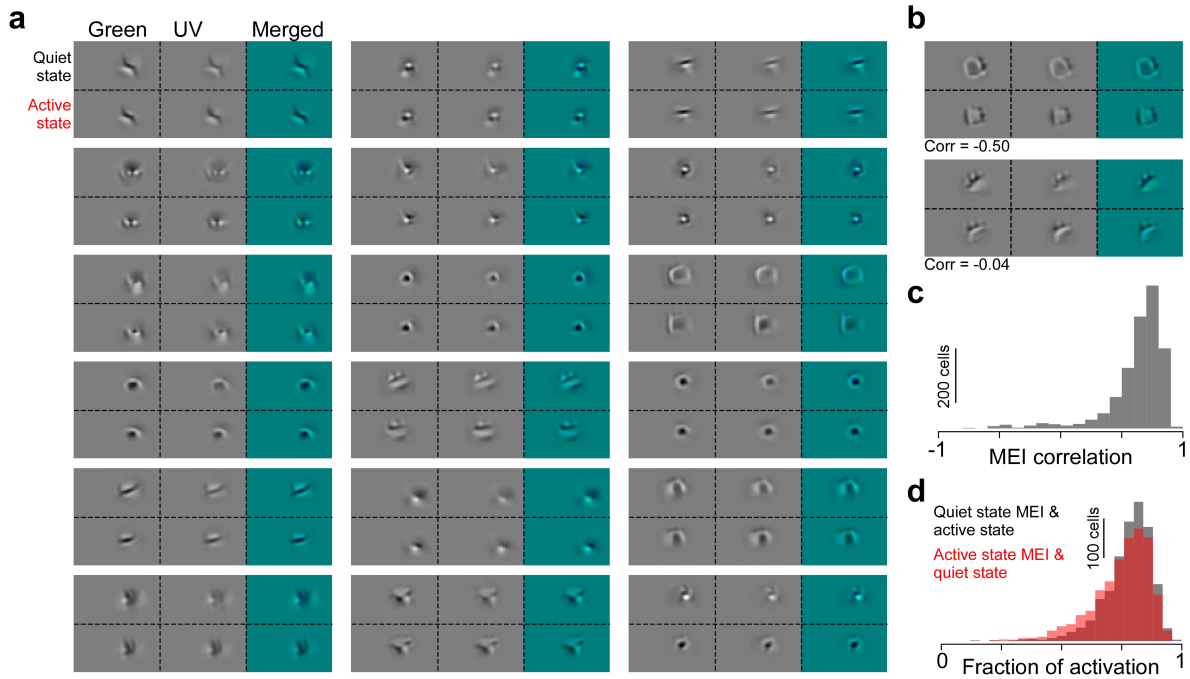
Extended Data Fig. 2 | Model performance and descriptive analysis of behaviour. a, Response reliability plotted versus test correlation (left) and correlation to average (right) for data shown in Fig. 2 (n=1.759 cells, n=3 scans, n=1 mouse). **b**, Mean Poisson loss (lower is better) for different models trained on the dataset from (a). The default model is used for all analysis, while models 1-3 are shown for comparison. Dotted line marks mean Poisson loss of default model. The default model had significantly lower Poisson loss values compared to all three alternative models (Wilcoxon signed rank test (two-sided), n=1,759: $p < 10^{-288}$ (model 1), 10^{-200} (model 2), 10^{-18} (model 3)). Error bars show 95% confidence interval. **c**, Mean response reliability, test correlation and correlation to average across neurons (error bars: s.d. across neurons; n=478 to n=1,160 neurons per recording) for n=10 models, with control and drug condition indicated below. **d**, Pupil size and locomotion speed trace of example animal, with active trials indicated by red dots. Trials were considered active if pupil size > 60th percentile and/or locomotion speed > 90th percentile. Plots on the right show mean pupil size across trials versus mean locomotion speed across trials. Dotted lines indicate 60th and 90th percentile of pupil size and locomotion speed, respectively. **e**, Example frames of eye camera for a quiet and active behavioural period for control and dilated condition. For the dilated condition, the eye was often squinted during quiet periods. **f**, Same as (e), but for control and constricted condition. Right plots show pupil size versus locomotion speed of trials used for model training for control and constricted condition.



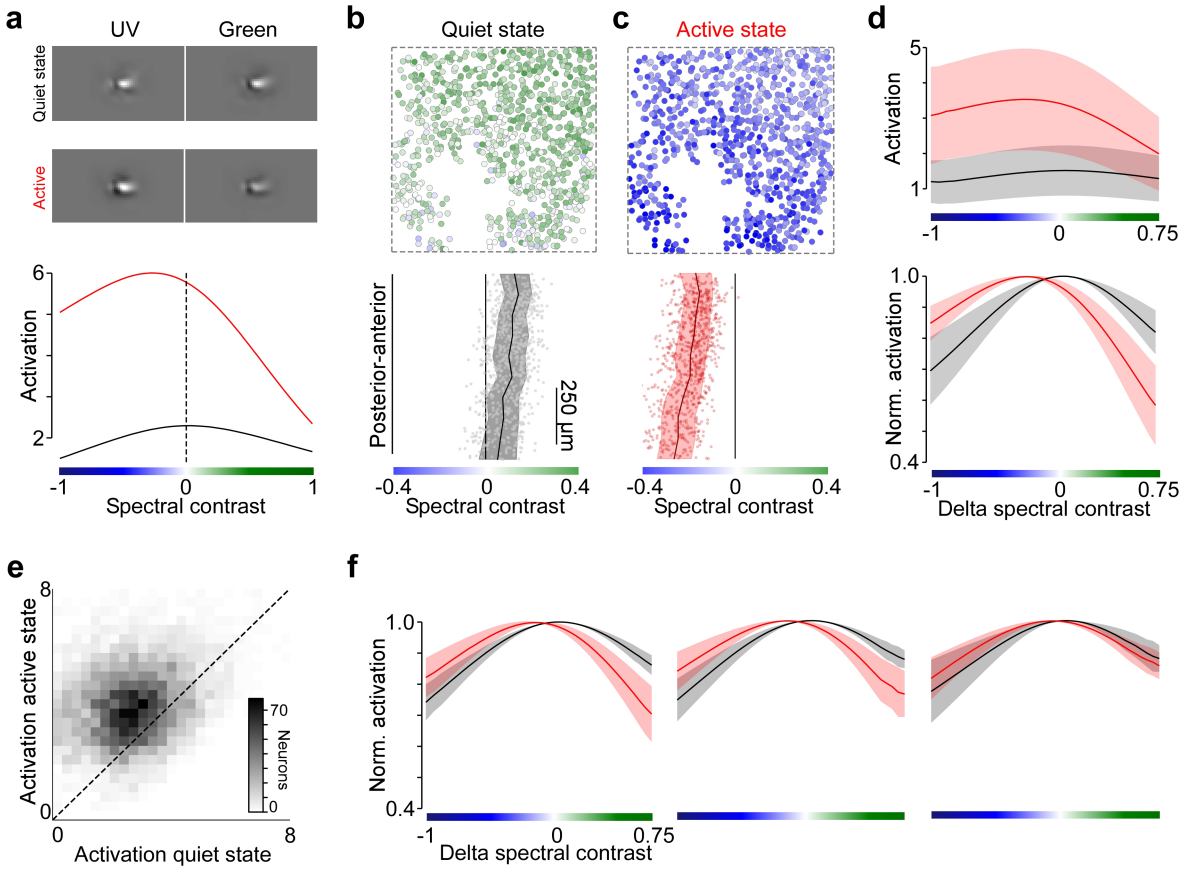
Extended Data Fig. 3 | Spatial and temporal colour opponency of mouse V1 neurons. a, MEIs of 21 exemplary neurons illustrate structural similarity across colour channels. **b**, Distribution of correlation across colour channels for dataset shown in Fig. 2. MEIs on top show example cells with relatively low correlation across colour channels. **c**, Schematic illustrating paradigm of 10 Hz full-field binary white noise stimulus and corresponding response of exemplary neuron. **d**, Temporal kernels estimated from responses to full-field noise stimulus from (c) of three exemplary neurons and distribution of kernel correlations (n=924 neurons, n=1 scan, n=1 mouse; scan 1 from (e)). Dotted line indicates correlation threshold of -0.25 – cells with a kernel correlation lower than this threshold were considered colour-opponent. A fraction of neurons (<5%) exhibited colour-opponent temporal receptive fields (see also [71]) in response to this full-field binary noise stimulus – in line with recent retinal work [60]. **e**, Neurons recorded in 3 consecutive scans at different positions within V1, colour-coded based on colour-opponency (red: opponent). **f**, Temporal kernels in response to full-field coloured noise stimulus of three exemplary neurons (left) and MEIs of the same neurons. Neurons were anatomically matched across recordings by alignment to the same 3D stack. This indicates that colour-opponency of mouse V1 neurons depends on stimulus condition, similar to neurons in mouse dLGN [72], which might be due to e.g. differences in activation of the neuron's surround or static versus dynamic stimuli.



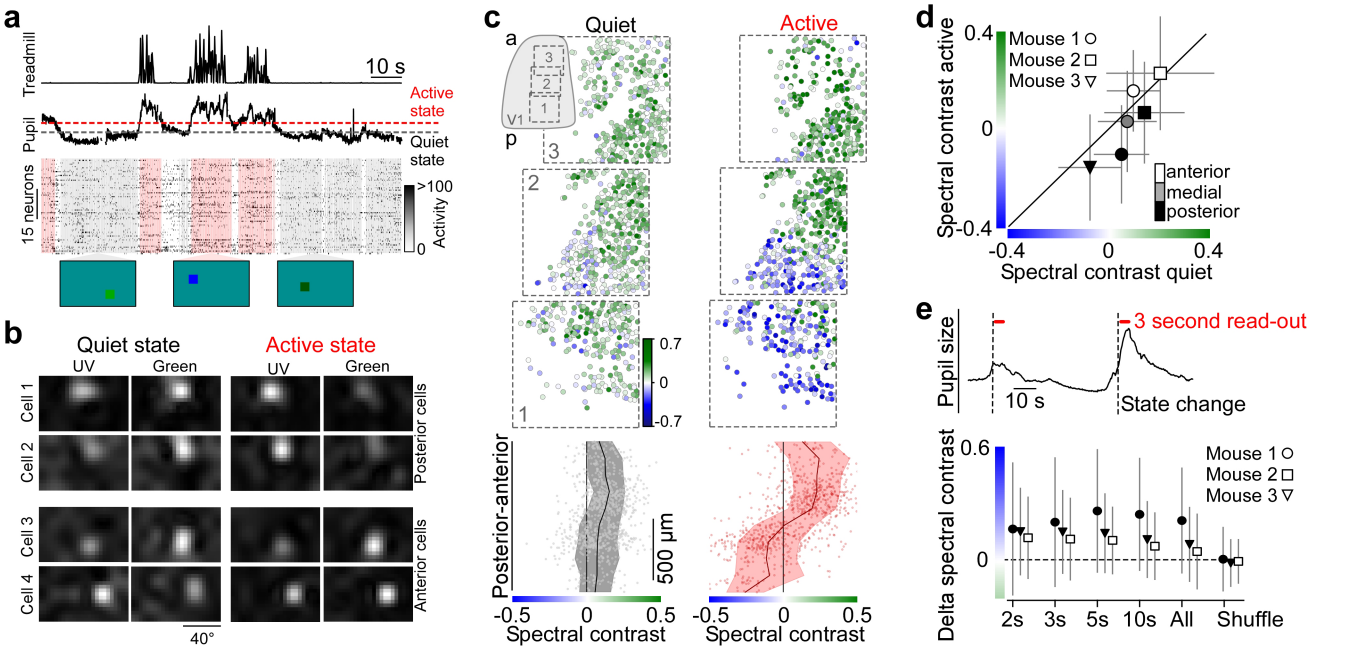
Extended Data Fig. 4 | Model recovers colour opponency and colour preference of simulated neurons. a, We simulated neurons with Gabor receptive fields (RFs) of varying size, orientation, spectral contrast and colour-opponency (correlation across colour channels). Then, responses of simulated neurons with Gabor RFs were generated by multiplication of the RFs with the natural images also used during experiments. Corresponding responses were passed through a non-linearity and a poisson process before model training. Model predictions and optimized MEIs closely matched the simulated responses and Gabor RFs, respectively. **b**, Gabor RFs and corresponding MEIs of four example neurons, some of them with colour-opponent RFs and MEIs. **c**, Spectral contrast of Gabor RFs plotted versus spectral contrast of computed MEIs. The model faithfully recovered the simulated neurons' colour preference. Only extreme colour preferences were slightly underestimated by our model, which is likely due to correlations across colour channels of natural scenes. This also suggests that it is unlikely that the low number of colour-opponent MEIs (Extended Data Fig. 3) is due to an artifact of modelling. **d**, Correlation of the MEI with the ground truth gabor RF.



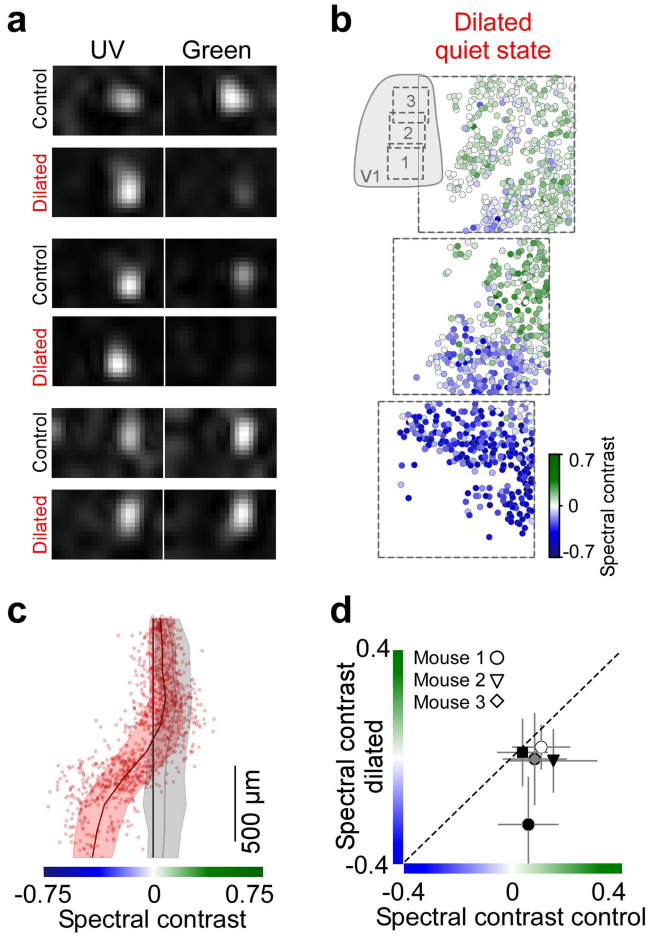
Extended Data Fig. 5 | MEI structure is consistent across quiet and active states. a, MEIs optimized for a quiet (top row of each sub-panel) and active (bottom row) behavioural state of 18 example neurons illustrate structural similarity of MEIs across states. **b**, MEIs of two exemplary neurons with low correlation across behavioural states. **c**, Distribution of MEI correlation across states (n=1,759 neurons, n=3 scans, n=1 mouse). **d**, MEI activation for incongruent behavioural state (n=1,759 neurons, n=3 scans, n=1 mouse). Gray: Model activation of MEI optimized for a quiet state presented to the model for active state relative to model activation of MEI optimized and presented for active state (activation=1). Red: Model activation of MEI optimized for active state presented to the model for quiet state relative to model activation of MEI optimized and presented for quiet state (activation=1). This suggests that MEIs optimized for different behavioural states lead to similar activations in the model and thus share similar tuning properties for the majority of neurons.



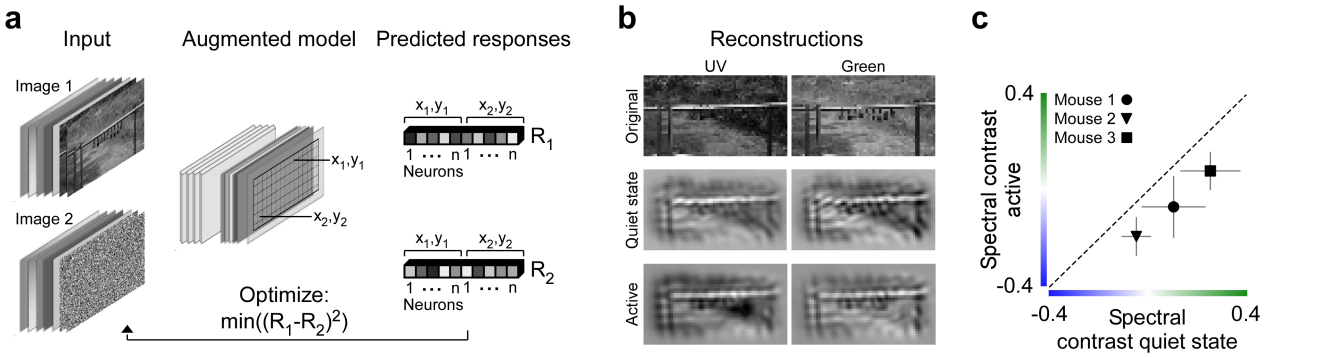
Extended Data Fig. 6 | Behavioural modulation of colour tuning of mouse V1 neurons - additional data. a, MEIs optimized for quiet and active state of exemplary neuron and corresponding colour tuning curves. **b**, Neurons recorded in posterior V1 colour coded based on spectral contrast of their quiet state MEI (top) and distribution of spectral contrast along posterior-anterior axis of V1 in an additional example animal. Black line corresponds to binned average (n=10 bins), with s.d. shading in gray. **c**, Like (b), but for active state. **d**, Mean of colour tuning curves of neurons from (b, c), aligned with respect to peak position of quiet state tuning curves. Shading: s.d. across neurons from this scan. Top shows higher model activation for active state tuning curves, in line with gain modulation of visual responses. Bottom shows peak-normalized tuning curves, illustrating (i) a shift towards lower spectral contrast values for the peak response, (ii) lower activation relative to peak for green-biased stimuli for an active state and (iii) stronger activation relative to peak for UV-biased stimuli for an active state. This suggests that during an active state, the increase in UV-sensitivity is accompanied by a decrease in green-sensitivity. **e**, Density plot of model activation in response to MEIs optimized for a quiet versus an active behavioural state, for n=6,770 neurons from n=7 mice. **f**, Mean of peak-normalized colour tuning curves of quiet (black) and active state (red), aligned with respect to peak position of quiet state tuning curves for n=3 scans from n=3 mice. Shading: s.d. across neurons.



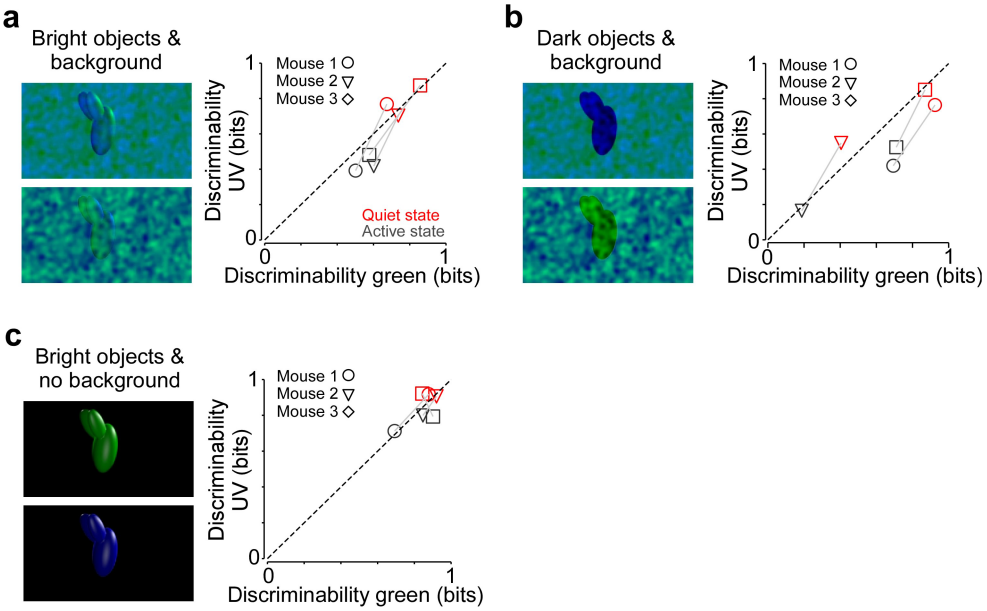
Extended Data Fig. 7 | Behavioural shift of colour preference of mouse V1 neurons in the context of a coloured sparse noise paradigm. a, Activity of n=50 exemplary V1 neurons in response to UV and green On and Off dots (10° visual angle) flashed for 0.2 seconds and simultaneously recorded locomotion speed and pupil size. Horizontal dashed lines indicate thresholds for quiet (black; < 50th percentile of pupil size) and active trials (red, > 75th percentile of pupil size). We adjusted the definition of quiet and active state compared to our *in-silico* analysis to ensure a sufficient number of trials in each state despite the shorter recording time (25 minutes for sparse noise versus 120 minutes for naturalistic images). Shading below in red and gray highlights trials above or below these thresholds. Bottom images show single stimulus frames. **b**, Spike-triggered average (STA) of 4 example neurons estimated from quiet and active trials, separated by posterior and anterior recording position. STAs estimated based on On and Off stimuli were combined to yield one STA per cell and pupil size. **c**, Neurons recorded in three consecutive experiments along the posterior-anterior axis of V1 (n=981 neurons, n=3 scans, n=1 mouse), colour coded based on spectral contrast of their STA estimated for quiet (left) and active trials (right). Bottom shows spectral contrast along the posterior-anterior axis of V1 of cells from (c, top), with binned average (black, n=10 bins) and s.d. shading (gray). Spectral contrast varied only slightly, but significantly along the anterior-posterior axis of V1 for quiet periods (n=981, p=10^{-7} for smooth term on cortical position of Generalized Additive Model (GAM); see Supplementary Methods). The small change in spectral contrast across the anterior-posterior axis of V1 is likely due to the fact that we pooled data from a wider range of pupil sizes. For an active state, optimal spectral contrast also changed with behavioural state (n=981, p=10^{-16} for behavioural state coefficient of GAM), with a significant interaction between cortical position and behavioural state modulation (p=10^{-7}; see Supplementary Methods). **d**, Mean STA spectral contrast of quiet versus active state for n=6 scans from n=3 mice. Error bars: s.d. across neurons recorded in one scan that passed quality threshold. Marker shape and filling indicate mouse ID and cortical position along the posterior-anterior axis, respectively. STA spectral contrast was significantly shifted (p=10^{-101}/3.68*10^{-51}/10^{-59}/10^{-303}, Wilcoxon signed rank test (two-sided)) towards UV for posterior and medial scan fields. The shift was not evident in anterior V1. This was likely due to the different definitions of quiet and active state in the model compared to the sparse noise recordings: For pupil size thresholds more similar to the ones used in the model (20th and 85th percentile), we observed a stronger UV-shift in STA colour preference with behaviour, also for anterior V1. **e**, Top: pupil size trace with state changes from quiet to active indicated by vertical dashed lines. Red dots show selected trials using a 3 second read-out window. Bottom: difference in STA spectral contrast of quiet versus active state for different read-out times after state change. All: all trials with quiet and active trials defined as < 20th and > 85th percentile of pupil size. Shuffle: all trials with shuffled behaviour parameters relative to neuronal responses. Dashed horizontal line indicates delta spectral contrast=0. Data shows mean and s.d. across neurons (n=996/702/964 cells, n=3 scans, n=3 animals).



Extended Data Fig. 8 | Pharmacological pupil dilation replicates shift in colour selectivity with sparse noise stimulus. a, STAs of three example neurons, estimated for quiet trials in control condition (black) and dilated condition (red). **b**, Neurons recorded in three consecutive experiments across the posterior-anterior axis of V1 (n=1,079 neurons, n=3 scans, n=1 mouse), colour coded based on STA estimated for quiet trials in the dilated condition. See Extended Data Fig. 7 for STAs estimated for the control condition of the same animal. **c**, Spectral contrast of STAs of neurons from (b) along the posterior-anterior axis of V1 (red dots), with binned average (n=10 bins; red line) and s.d. shading. Black line and gray shading corresponds to binned average and s.d. of neurons recorded at the same cortical positions in control condition (cf. Extended Data Fig. 7). Spectral contrast significantly varied across anterior-posterior axis of V1 for the dilated condition (n=1,079, $p=10^{-16}$ for smooth term on cortical position of GAM). Optimal spectral contrast changed with pupil dilation (n=1,079 (dilated) and n=943 (control), $p=10^{-16}$ for condition coefficient of GAM), with a significant interaction between cortical position and behavioural state modulation (see Supplementary Methods). **d**, Mean spectral contrast of quiet state STAs in control condition versus spectral contrast of quiet state STAs in dilated condition (n=10 scans, n=3 mice). Error bars: s.d. across neurons. Two-sample t-test (two-sided): $p=10^{-135}/10^{-20}/10^{-29}/10^{-194}/0.0006$.



Extended Data Fig. 9 | Reconstructions of coloured naturalistic scenes predict colour tuning shift for a population of neurons. a, Schematic illustrating reconstruction paradigm. As the receptive fields of neurons recorded within one of our scans only covered a fraction of the screen, we used an augmented version of our CNN model for image reconstruction where the receptive field of each model neuron was copied to each pixel position of the image except the image margins. For a given target input image (image 1), this results in a predicted response vector (R1) of length number of neurons times number of pixels. During image reconstruction, a novel image (image 2) is optimized such that its corresponding response vector (R2) matches the response vector of the target image as closely as possible. **b**, Green and UV image channels of exemplary test image (top) and reconstructions of this image for a quiet (middle) and active state (bottom). For reconstructions, neurons from scan 1 in Fig. 2 were used. **c**, Spectral contrasts of reconstructed test images (n=100) in quiet state versus active state for n=3 models trained on scans from n=3 animals. Wilcoxon signed rank test (two-sided): $p=10^{-18}/10^{-18}/10^{-18}$.



Extended Data Fig. 10 | Additional data and stimulus conditions for decoding paradigm. a, Exemplary frames of stimulus condition with lower object contrast than in Fig. 5c due to gray background in the object colour channel. Right: Scatter plot of decoding discriminability of green versus UV objects for quiet (gray) and active (red) trials for n=3 animals. Each marker represents the decoding performance of the SVM decoder trained on all neurons of the respective scan. The decoding performance for the two behavioural states are connected with gray lines, with slopes larger than one for all animals, corresponding to a larger increase in decoding performance for UV versus green objects. P-values obtained from a one-sided permutation test: < 0.012 (Mouse 1), < 0.032 (Mouse 2), < 0.112 (Mouse 3). **b**, Like (a), but for stimulus condition with objects as dark silhouettes and noise in the other colour channel. P-values obtained from a one-sided permutation test: < 0.02 (Mouse 1), < 0.1 (Mouse 2), < 0.038 (Mouse 3). **c**, Like (a), but for stimulus condition with high contrast objects and no noise in the other colour channel. P-values obtained from a one-sided permutation test (see Methods for detail): 0.44 (Mouse 1), 0.404 (Mouse 2), 0.024 (Mouse 3). The observed variability in (a) and (b) across animals might be related to different recording positions along the anterior-posterior axis of V1 and differences in the animal's behaviour, i.e. the time spent in a quiet versus active behavioural state. For the stimulus condition in (c), we might also observe a ceiling effect caused by the fact that these stimuli are relatively easy to discriminate, as indicated by high object discriminability even during quiet behavioural periods.

nature portfolio

Corresponding author(s): Katrin Franke

Last updated by author(s): Aug 4, 2022

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	For image acquisition, we used ScanImage 2017b. Stimuli were presented using PsychToolBox 3 and rendered (for 3D objects) using Blender 2.7.
Data analysis	We used DeepLabCut for automatic tracking of the pupil. Our coding framework used general tools like PyTorch (1.7.0), Numpy (1.19.4), scikit-image (0.17.2), matplotlib (3.3.2), seaborn (0.11.0), DataJoint (0.12.7), Jupyter (Core: 6.1.5), and Docker (19.03.13). We also used the following custom libraries and code: neuralpredictors (https://github.com/sinzlab/neuralpredictors) for torch-based custom functions for model implementation, nnfabrik (https://github.com/sinzlab/nnfabrik) for automatic model training pipelines using Datajoint, nndichromacy for utilities, (https: //github.com/sinzlab/nndichromacy), and mei (https://github.com/sinzlab/mei) for stimulus optimization.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The data used in this paper will be stored at an online GIN repository.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Due to the exploratory nature of this study, no statistical methods were used to predetermine sample size. All high-quality imaging fields collected from 3-7 animals per experiment were included in analysis, which resulted in large sample sizes of neurons per animal. At least three animals per experimental condition were used.
Data exclusions	Cells were excluded from analysis based on quality filtering with respect to their responses to visual stimuli and/or model prediction performance. This is described in detail in the Methods.
Replication	Datasets in the manuscript include at least 3 animals with a large number of samples per animal to assess reproducibility. In all main figures, we show data from all animals.
Randomization	Stimuli were presented in a pseudorandom order and the order of presentation of different stimuli was varied. No other randomization was used.
Blinding	Owing to the exploratory nature of our study, we did not use blinding.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☐	☒ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	Mice of either sex (Mus musculus, n=13) expressing GCaMP6s in excitatory neurons via Slc17a7-Cre and Ai162 transgenic lines (stock number 023527 and 031562, respectively; The Jackson Laboratory) were used for all experiments. Mice were between 6 weeks and 5 months old.
Wild animals	No wild animals were used in this study.
Field-collected samples	No field-collected samples were used in this study.
Ethics oversight	All procedures were approved by the Institutional Animal Care and Use Committee of Baylor College of Medicine.

Note that full information on the approval of the study protocol must also be provided in the manuscript.