Coherent encoding of subjective spatial position in visual cortex and hippocampus

Aman B. Saleem[#1,2], **E. Mika Diamanti**[#1,3], **Julien Fournier**[1], **Kenneth D. Harris**[4,*], and **Matteo Carandini**[1,*]

[1]UCL Institute of Ophthalmology, University College London, London, EC1V9EL, UK

[2]Department of Experimental Psychology, University College London, London, WC1H 0AP, UK

[3]CoMPLEX, Department of Computer Science, University College London, London, WC1E 6BT, UK

[4]UCL Institute of Neurology, University College London, London, WC1N 3BG, UK

[#] These authors contributed equally to this work.

Abstract

A major role of vision is to guide navigation, and navigation is strongly driven by vision[1–4]. Indeed, the brain's visual and navigational systems are known to interact[5,6], and signals related to position in the environment have been suggested to appear as early as in visual cortex[6,7]. To establish the nature of these signals we recorded in primary visual cortex (V1) and in hippocampal area CA1 while mice traversed a corridor in virtual reality. The corridor contained identical visual landmarks in two positions, so that a purely visual neuron would respond similarly in those positions. Most V1 neurons, however, responded solely or more strongly to the landmarks in one position. This modulation of visual responses by spatial location was not explained by factors such as running speed. To assess whether the modulation is related to navigational signals and to the animal's subjective estimate of position, we trained the mice to lick for a water reward upon reaching a reward zone in the corridor. Neuronal populations in both CA1 and V1 encoded the animal's position along the corridor, and the errors in their representations were correlated. Moreover, both representations reflected the animal's subjective estimate of position, inferred from the animal's licks, better than its actual position. Indeed, when animals licked in a given location – whether correct or incorrect – neural populations in both V1 and CA1 placed the animal in the reward zone. We conclude that visual responses in V1 are controlled by navigational signals,

which are coherent with those encoded in hippocampus and reflect the animal's subjective position. The presence of such navigational signals as early as in a primary sensory area suggests that they permeate sensory processing in the cortex.

To characterise the influence of spatial position on the responses of area V1, we took mice expressing the calcium indicator GCaMP6 in excitatory cells and placed them in a corridor in virtual reality (VR; Figure 1**a**). The corridor had a pair of landmarks, a grating and a plaid, that repeated twice, thus creating two visually-matching segments 40 cm apart (Figure 1**a, b**; Extended Data Figure 1). We identified V1 based on the retinotopic map measured using wide-field imaging (Figure 1**c**). We then pointed a two-photon microscope on medial V1, focusing our analysis on neurons with receptive field centres >40° azimuth (Figure 1**c**), which were driven as the mouse passed the landmarks. As expected, given the repetition of visual scenes in the two segments of the corridor, some V1 neurons had a response profile with two equal peaks 40 cm apart (Figure 1**d**). Other V1 neurons, however, responded differently to the same visual stimuli in the two segments (Figure 1**d**). These results indicate that visual activity in V1 can be strongly modulated by the animal's position in an environment.

This modulation of visual responses by spatial position occurred in the majority of V1 neurons (Figure 1**e-g**). We imaged 8,610 V1 neurons across 18 sessions in 4 mice and selected 4,958 neurons with receptive field centres >40° azimuth and reliable firing along the corridor (see Methods). We divided the trials in half, and used the odd trials to find the position where each neuron fired maximally. The resulting representation reveals a striking preference of V1 neurons for spatial position (Figure 1**e**), with most neurons giving stronger responses in one position (preferred position) than in the visually-matching position 40 cm away (non-preferred position). To quantify this preference while avoiding circularity, we used the other half of the data, the even trials, and found that the preference for position was robust (Figure 1**f**). Among the neurons that responded when the animal traversed the visually-matching segments (n = 2,422), the responses at the non-preferred position were markedly smaller than at the preferred position (Figure 1**g**; Extended Data Figure 2). We defined a "spatial modulation ratio" for each cell as the ratio of responses at the two visually-matching positions (non-preferred/preferred, in the cross-validated trials). The median spatial modulation ratio was 0.61 ± 0.31 ($\pm$ m.a.d.), significantly <1 ($p < 10^{-104}$, Wilcoxon two-sided signed rank test). Neurons preferred the first or second sections in equal proportions (49% vs. 51%), making it unlikely that a global factor such as visual adaptation could explain their preference.

The modulation of V1 responses by spatial position could not be explained by visual factors. To confirm that the receptive fields of most neurons saw similar stimuli in the two visually-matching locations, we ran a model of receptive field responses on the sequences of images (a simulation of V1 complex cells). As expected, this model generated spatial modulation ratios close to 1 (0.97 ± 0.17, Extended Data Figure 3). We next asked whether the different responses seen in the two locations may be due to differences in images far outside the receptive field: particularly the end (grey) wall of the corridor. To test this, we placed two additional mice in a modified VR environment, where the two sections of the corridor were

pixel-to-pixel identical (Extended Data Figure 4). The spatial modulation ratio was again overwhelmingly <1 (0.62 ± 0.26; $p < 10^{-81}$; n = 1,044 cells), confirming that spatial modulation of V1 responses could not be explained by distant visual cues.

Spatial modulation of V1 responses also could not be explained by running speed, deviations in pupil position and diameter, or reward. Given that V1 neurons are influenced by running speed and visual speed[8,9], their different responses in visually-matching segments of the corridor could reflect speed differences. To control for this, we stratified the data according to three running speed ranges (low, medium, or high; Extended Data Figure 5). Even within a group (medium speed), the spatial modulation ratio was substantially <1 (0.47 ± 0.22; $p < 10^{-33}$). Moreover, the spatial ratio of responses was identical at low and high speeds (Figure 1**h**). The preference for position also could not be explained by reward or deviations in pupil position and size, as the spatial modulation ratio was markedly <1 even in sessions where the animals ran without a reward (0.57 ± 0.37; $p < 10^{-14}$), or when there were no changes in pupil size (0.63 ± 0.33, $p < 10^{-45}$) or pupil position (0.63 ± 0.33, $p < 10^{-27}$; Extended Data Figure 6). To assess the joint contribution of visual, task-related, and position variables, we developed three prediction models (Extended Data Figure 7). The first depended only on the visual scenes (which repeat twice), and on trial onset and offset (which introduce transients). The second additionally depended on running speed, reward times, pupil size, and eye position (non-spatial model). The third, additionally allowed responses to differ in amplitude in the matching segments. Only the last model could fit the activity of cells with unequal peaks, thus matching the spatial modulation ratios seen in the data (Extended Data Figure 7**d**). By contrast, the first two models predicted spatial modulation ratios closer to 1 (Figure 1h; Extended Data Figure 7**c-d**).

Having established that V1 responses are modulated by spatial position, we next asked whether the underlying modulatory signals reflect the spatial position encoded in the brain's navigational systems (Figure 2). We recorded simultaneously from V1 and hippocampal area CA1 using two 32-channel electrodes (Figure 2**a**). To gauge a mouse's estimate of position, we trained the mice to lick a spout for water reward upon reaching a specific region of the corridor (Figure 2**b**; Supplementary Video 1, Extended Data Figure 8). All four mice (wild-type) learned to perform this task with >80% accuracy and strongly relied on vision: performance persisted when we changed the gain relating wheel rotation to progression in the corridor[3,10] and performance decreased when we lowered visual contrast (Extended Data Figure 8).

Many neurons in both visual cortex and hippocampus had place-specific response profiles, thus encoding the mouse's spatial position (Figure 2**c-f**). Consistent with our observations with two-photon imaging, V1 neurons responded more strongly in one of the two visually-matching segments of the corridor (Figure 2**c**, Extended Data Figure 9**c**). In turn, hippocampal CA1 neurons exhibited place fields[3,10,11], responding in a single corridor location (Figure 2**d**, Extended Data Figure 9**a-c**). Therefore, responses in both V1 and CA1 encoded the position of the mouse in the environment, with no ambiguity between the two visually-matching segments. Indeed, an independent Bayes decoder was able to read out the animal's position from the activity of neurons recorded from V1 (33 ± 17 neurons/session, n = 8 sessions; Figure 2**e**) or from CA1 (42 ± 20 neurons/session; n = 8 sessions; Figure 2**f**).

Furthermore, when visual cortex and hippocampus made errors in estimating the mouse's position, these errors were correlated with each other (Figure 2**g-h**). The distributions of errors in position decoded from V1 and CA1 peaked at zero (Figure 2**g**) but were significantly correlated (Figure 2**h**; ρ = 0.125, p = 0.0129, two-sided t-test, n = 8). In principle, correlation could arise from a common modulation of both regions by behavioural factors such as running speed, which affects responses of both visual cortex 8,9 and hippocampus12–14. To isolate the effect of speed, we shuffled the data between time-points while preserving the relationship between speed and position (Supplementary Methods). After shuffling, the correlation between V1 and CA1 decoding errors decreased substantially from 0.125 to 0.022 (p = 0.0115; Figure 2**g, h**). Moreover, when we subtracted the shuffled distribution from the original joint distributions, the residual decoding errors were distributed along the diagonal (Figure 2**i, j**), indicating that V1 and CA1 representations are more correlated than expected from common speed modulation. This correlation could also not be explained by common encoding of behavioural factors such as licking (Extended Data Figure **9d-f**). Finally, a prediction of V1-encoded position from all external variables (true position, running speed, licks and rewards), could still be improved by the position decoded from CA1 activity (Extended Data Figure 10).

We next asked whether the spatial position encoded by V1 and CA1 relates to the animal's subjective estimate of position (Figure 3**a-f**). CA1 activity is influenced by the performance of navigation tasks15–18, and may reflect the animal's subjective position more than actual position15,17,19. We assessed a mouse's subjective estimate of position from the location of its licks. We divided trials into three groups: *Early* trials when too many licks (usually 4-6) occurred before the reward zone, causing the trial to be aborted; *Correct* trials when one or more licks occurred in the reward zone; and *Late* trials when the mouse missed the reward zone and licked afterwards. To understand how V1 and CA1 spatial representations related to this behaviour, we trained the Bayesian decoder on the activity measured in *Correct* trials, and analysed the likelihood of decoding different positions in the three types of trials. Decoding performance in *Early* and *Late* trials showed systematic deviations: in *Early* trials, V1 and CA1 overestimated the animal's progress along the corridor (deviation above the diagonal, Figure 3**a**, **d**), whereas in *Late* trials they underestimated it (deviation below the diagonal, Figure 3**b**, **e**). Accordingly, in both CA1 and V1, the probability of being in the reward zone peaked before the reward zone in *Early* trials and after it in *Late* trials (Figure 3**c, f**). These consistent deviations suggest that the representations of position in V1 and CA1 correlate with the animal's decisions to lick and thus reflect its subjective estimate of position**.**

The licks provide an opportunity to gauge when the mouse's subjective estimate of position lies in the reward zone**.** If activity in V1 and CA1 reflects subjective position, it should place the animal in the reward zone whether the animal correctly licked in that zone or incorrectly licked earlier or later. To test this prediction, we decoded activity in V1 and CA1 at the time of licks. By definition, the distributions of licks in *Early*, *Correct*, and *Late* trials were spatially distinct (Figure 3**g, h** and **j, k**). However, when plotted as a function of decoded position these distributions came into register over the reward zone, whether the decoding was done from V1 (Figure 3**g-i**) or from CA1 (Figure 3**j-l**). Thus, regardless of the animal's

position, when animals licked for a reward, the activity of both V1 and CA1 indicated a position in the reward zone.

Taken together, these results indicate that visual responses in V1 are modulated by the same spatial signals represented in hippocampus, and that these signals reflect the animal's subjective position. This modulation may become stronger as environments become familiar[6,7], perhaps contributing to the changes observed in V1 as animals learn behavioural tasks[20–22]. The correlation between V1 and CA1 representations may be due to feed-forward signals from vision or feedback signals from navigational systems. While V1 and CA1 are not directly connected, they could share spatial signals through indirect connections[23,24], which may involve retrosplenial, parietal, or prefrontal cortices, which are known to carry spatial information[25,26]. Further insights into the nature of these signals could be obtained by modulating the relationship between actual position and distance run[3,10] or time[27], and investigating more natural, 2-D environments[28–30]. In such environments, however, it would be difficult to control and repeat visual stimulation. This ability proved essential in our study. Our results show that signals related to an animal's own estimate of position appear as early as in a primary sensory cortex. This observation suggest that the mouse cortex does not keep a firm distinction between navigational and sensory systems: rather, spatial signals may permeate cortical processing.

Methods

All experiments were conducted according to the UK Animals (Scientific Procedures) Act, 1986 under personal and project licenses issued by the Home Office following ethical review.

For simultaneous recordings in V1 and CA1, we used four C57BL/6 mice (all male, implanted at 4-8 weeks of age). For calcium imaging experiments, we used double or triple transgenic mice expressing GCaMP6 in excitatory neurons (5 females, 1 male, implanted at 4-6 weeks). The triple transgenics expressed GCaMP6 fast[31](Emx1- Cre;Camk2a-tTA;Ai93, 3 mice). The double transgenic expressed GCaMP6 slow[32] (Camk2a-tTA;tetO-G6s, 3 mice). Because Ai93 mice may exhibit aberrant cortical activity[33], we used the GCamp6 slow mice to validate the results obtained from the GCaMP6 fast mice. Additional tests[33] confirmed that none of these mice displayed the aberrant activity that is sometimes seen in Ai93 mice. No randomisation or blinding was performed in this study.

Virtual Environment and task

The virtual reality environment was a corridor adorned with a white noise background and four landmarks: two grating stimuli oriented orthogonal to the corridor and two plaid stimuli (Figure 1a). The corridor dimensions were 100 x 8 x 8 cm, and the landmarks (8 cm wide) were centred 20, 40, 60 and 80 cm from the start of the corridor. The animal navigated the environment by walking on a custom-made polystyrene wheel (15 cm wide, 18 cm diameter). Movements of the wheel were captured by a rotary encoder (2400 pulses/rotation, Kübler, Germany), and used to control the virtual reality environment presented on three monitors surrounding the animal, as previously described[12]. When the animal reached the end of the corridor, it was placed back at the start of the corridor after a 3-5 s presentation of

a grey screen. Trials longer than 120 s were timed out and were excluded from further analysis.

Mice used for simultaneous V1-CA1 recordings (n = 4 animals, 8 sessions) were trained to lick in a specific region of the corridor, the reward zone. This zone was centred at 70 cm and was 8 cm wide. Trials in which the animals were not engaged in the task, i.e. when they ran through the environment without licking, were excluded from further analysis. The animal was rewarded for correct licks with ~2 µl water using a solenoid valve (161T010; Neptune Research, USA), and licks were monitored using a custom device that detected breaks in an infrared beam.

Mice used for calcium imaging (n = 6 animals, 25 sessions) ran through the corridor, with no specific task. Two of the mice were motivated to run with water rewards. One animal received rewards at random positions along the corridor. The other received rewards at the end of the corridor. To control for the effect of the reward on V1 responses, no reward was delivered in some sessions (n = 8 sessions; 2 animals).

To ensure that the spatial modulation of V1 responses could not be explained by the end wall of the corridor being more visible in the second half than in the first half, two of the mice used for calcium imaging were further trained in a modified version of the corridor, where visual scenes were strictly identical 40 cm apart. In this environment, mice ran the same distance as before (100 cm) and were also placed back at the start of the corridor after 3-5 s presentation of a grey screen. The same four landmarks were also centred in the same positions as before. However, the corridor was extended to 200 cm length, repeating the same sequence of landmarks (Extended Data Figure 4). The virtual reality software was modified to render only up to 70 cm ahead of the animal, ensuring the visual scenes were strictly identical in the sections between 10 - 50 cm and 50 - 90 cm; the white noise background also repeated with same 40 cm periodicity. Prior to recording in the 200cm corridor, mice were first exposed to 5 sessions in the 100 cm corridor, then placed in the 200 cm corridor and allowed to habituate to the new environment for another two or three sessions before the start of recordings.

Surgery and training

The surgical methods are similar to those described previously[9,34]. Briefly, a custom head-plate with a circular chamber (3-4 mm diameter for electrophysiology; 8 mm for imaging) was implanted on 4-8 week mice under isoflurane anaesthesia. For imaging, we performed a 4 mm craniotomy over left visual cortex by repeatedly rotating a biopsy punch. The craniotomy was shielded with a double coverslip (4 mm inner diameter; 5 mm outer diameter). After 4 days of recovery, some mice were water restricted (> 40 ml / kg / day) and were trained for 30-60 min, 5-7 days/week.

Mice used for simultaneous V1-CA1 recordings were trained to lick selectively in the reward zone using a progressive training procedure. Initially, the animals were rewarded for running past the reward location on all trials. After this, we introduced trials where the animal was rewarded only when it licked in the rewarded region of the corridor. The width of the reward region was progressively narrowed from 30 cm to 8 cm across successive days

of training. To prevent the animals from licking all across the corridor, trials were terminated early if the animal licked more than a certain number of times before the rewarded region. We reduced this number as the animals performed more accurately, typically reaching a level of 4-6 licks by the time recordings were made. Once a sufficient level of performance was reached, we controlled on some (random) trials that the animal performed the task visually by measuring the performance when decreasing the visual contrast or changing the distance to the reward zone (Extended Data Figure 8). Training was carried out for 3-5 weeks. Animals were light shifted (9 am light off, 9 pm light on) and experiments were performed during the day.

Widefield calcium imaging

For widefield imaging we used a standard epi-illumination imaging system[35,36] together with an SCMOS camera (pco.edge, PCO AG). A Leica 1.6x Plan APO objective was placed above the imaging window and a custom black cone surrounding the objective was fixed on top of the headplate to prevent contamination from the monitors' light. The excitation light beam emitted by a high-power LED (465 nm LEX2-B, Brain Vision) was directed onto the imaging window by a dichroic mirror designed to reflect blue light. Emitted fluorescence passed through the same dichroic mirror and was then selectively transmitted by an emission filter (FF01-543/50-25, Semrock) before being focused by another objective (Leica 1.0 Plan APO objective) and finally detected by the camera. Images of 200 x 180 pixels, corresponding to an area of 6.0 x 5.4 mm were acquired at 50 Hz.

To measure retinotopy we presented a 14°-wide vertical window containing a vertical grating (spatial frequency 0.15 cycles/°), and swept[37,38] the horizontal position of the window over 135° of azimuth angle, at a frequency of 2 Hz. Stimuli lasted 4 s and were repeated 20 times (10 in each direction). We obtained maps for preferred azimuth by combining responses to the 2 stimuli moving in opposite direction, as previously described[37].

Two-photon imaging

Two-photon imaging was performed with a standard multiphoton imaging system (Bergamo II; Thorlabs) controlled by ScanImage4[39]. A 970 nm laser beam, emitted by a Ti:Sapphire Laser (Chameleon Vision, Coherent), was targeted onto L2/3 neurons through a 16x water-immersion objective (0.8 NA, Nikon). Fluorescence signal was transmitted by a dichroic beamsplitter and amplified by photomultiplier tubes (GaAsP, Hamamatsu). The emission light path between the focal plane and the objective was shielded with a custom-made plastic cone, to prevent contamination from the monitors' light. In each experiment, we imaged 4 planes set apart by 40 μm. Multiple-plane imaging was enabled by a piezo focusing device (P-725.4CA PIFOC, Physik Instrumente), and an electro-optical modulator (M350-80LA, Conoptics Inc.) which allowed adjustment of the laser power with depth. Images of 512x512 pixels, corresponding to a field of view of 500x500 μm, were acquired at a frame rate of 30 Hz (7.5 Hz per plane).

Pre-processing of raw imaging movies was done using the Suite2p pipeline[40] and involved: 1) image registration to correct for brain movement, 2) ROI extraction, i.e. cell detection and

3) correction for neuropil contamination. For neuropil correction, we used an established method[41,42]. We used Suite2p to determine a mask surrounding each cell's soma, the 'neuropil mask'. The inner diameter of the mask was 3 μm and the outer diameter was < 45 μm. For each cell we obtained a correction factor, α, by regressing the binned neuropil signal (20 bins in total) from the 5th percentile of the raw binned cell signal. For a given session, we obtained the average correction factor across cells. This average factor was used to obtain the corrected individual cell traces, from the raw cell traces and the neuropil signal, assuming a linear relationship. All correction factors fell within 0.7 and 0.9.

To manually curate Suite2p's output we used two criteria: one anatomical and one activity-dependent criterion. One of the anatomical criteria in Suite2p is 'area', i.e. mean distance of pixels from ROI centre, normalised to the same measure for a perfect disk. We used this criterion (area < 1.04) to exclude ROIs likely to correspond to dendrites rather than somata. The activity-related criterion is the standard deviation of the cell trace, normalised to the standard deviation of the neuropil trace. We used this criterion to exclude ROIs whose activity was too small relative to the corresponding neuropil signal (typically with std(neuropil corrected trace) / std(neuropil signal) < 2). We finally excluded cells with extremely seldom firing (once or twice within a 20 mins session).

Pupil tracking

We tracked the eye of the animal using an infrared camera (DMK 21BU04.H, Imaging Source) and a zoom lens (MVL7000, Navitar) at 25 Hz. Pupil position and size were calculated by fitting an ellipsoid to the pupil for each frame using a custom software. X and Y positions of the pupil were derived from the centre of mass of the fitted ellipsoid.

Electrophysiological recordings

On the day prior to the first recording session, we made two 1 mm craniotomies, one over CA1 (1.0 mm lateral, 2.0 mm anterior from lambda), and a second one over V1 (2.5 mm lateral, 0.5 mm anterior from lambda). We covered the chamber using KwikCast (World Precision Instruments) and the animals were allowed to recover overnight. The CA1 probe was lowered until all shanks were in the pyramidal layer, which was identified based on the increase in theta power (5-8Hz) of the local field potential and an increase in the number of detected units. We waited ~30 min for the tissue to settle before starting the recordings. In two animals, we dipped the probes in red-fluorescent DiI (Figure 2a). In these animals, we had only one recording session. The other animals had two and four recording sessions.

Offline spike sorting was carried out using the KlustaSuite[43] package, with automated spike sorting using KlustaKwik[44], followed by manual refinement using KlustaViewa[43]. Hippocampal interneurons were identified based on their spike time autocorrelation and excluded from further analysis. Only time points with running speeds greater than 5 cm/s were included in further analyses.
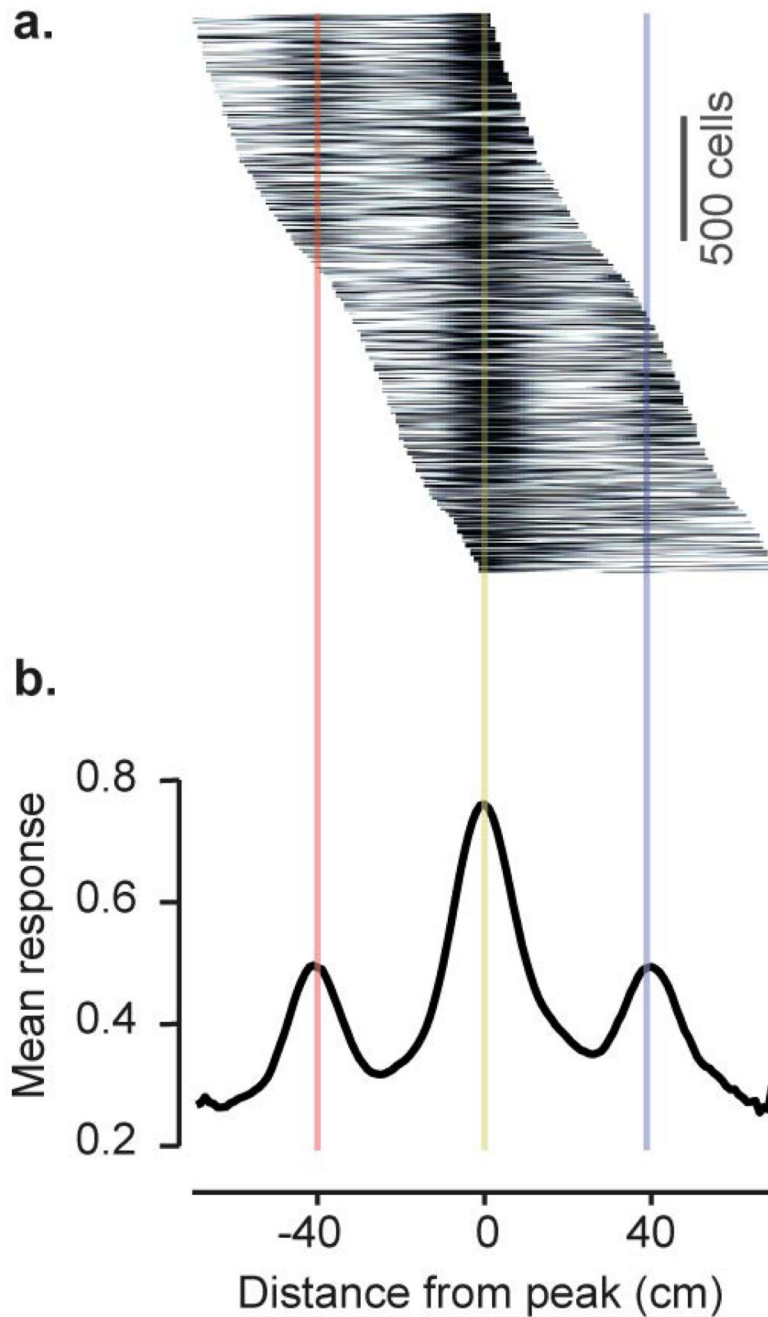
Extended Data
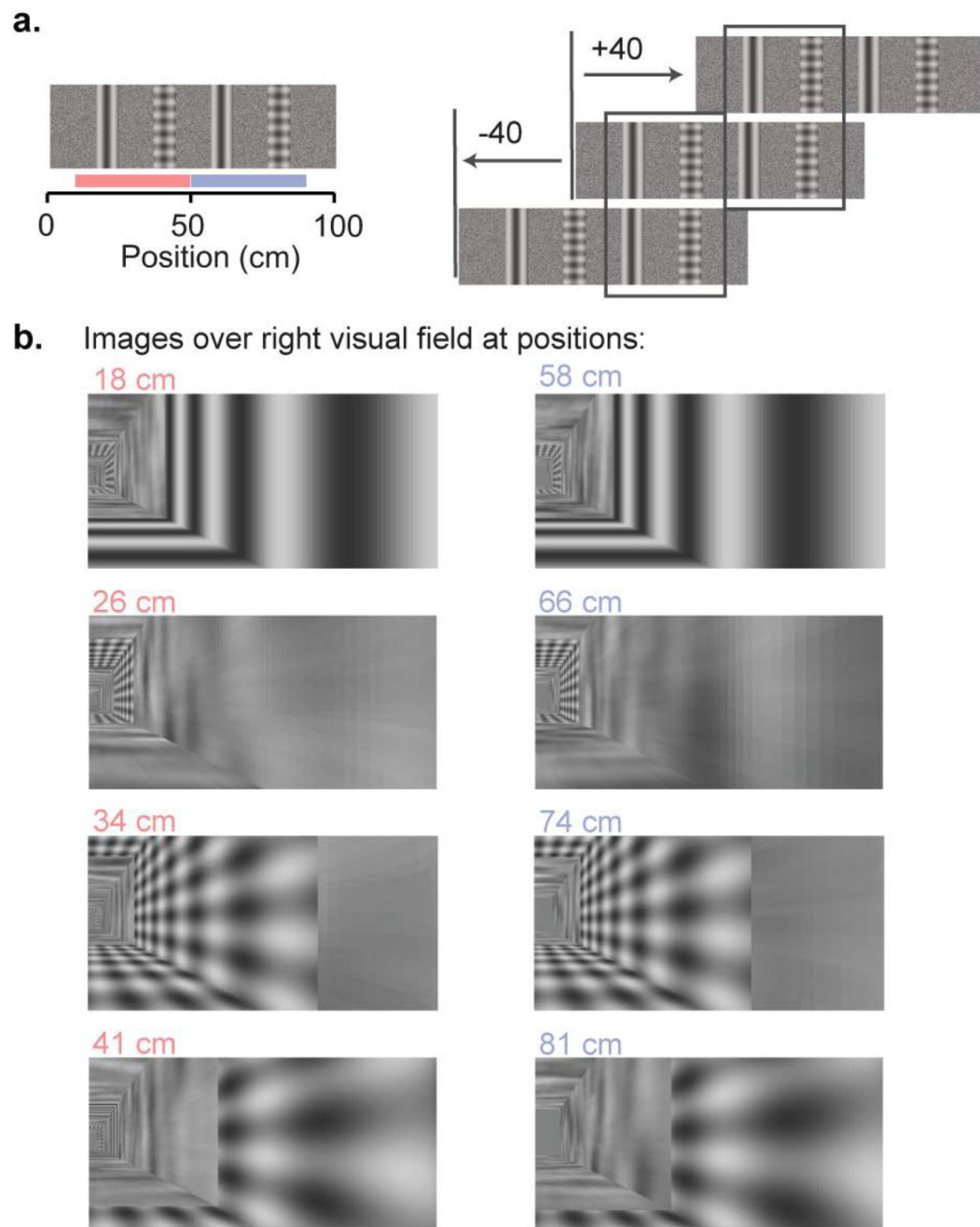
Extended Data Figure 1. Design of virtual environment with two visually matching segments.
a, The virtual corridor had four prominent landmarks. Two visual patterns (grating and plaid) were repeated at two positions, 40 cm apart creating two visually-matching segments in the room, from 10 cm to 50 cm and from 50 cm to 90 cm (indicated by *red* and *blue* bars in the *left panel*), as illustrated in the *right panel*.

b, Example screenshots of the right visual field displayed in the environment when the animal is at different positions. Each row displays screen images at positions approximately 40 cm apart.


Extended Data Figure 2. Spatial averaging of visual cortical activity confirms the difference in response between visually matching locations.
a, Mean response of V1 neurons as a function of the distance from the peak response location (2,422 cells with peak response between 15 and 85 cm along the corridor). To ensure that the average captured reliable spatially-specific responses, the peak response

location for each cell was estimated from only odd trials, while the mean response was computed from even trials.

b, Population average of responses shown in **a.** Lower values of the side peaks compared to central peak indicates strong preference of V1 neurons for one segment of the corridor over the other visually-matching segment (40 cm away from peak response).

Complex cell simulations
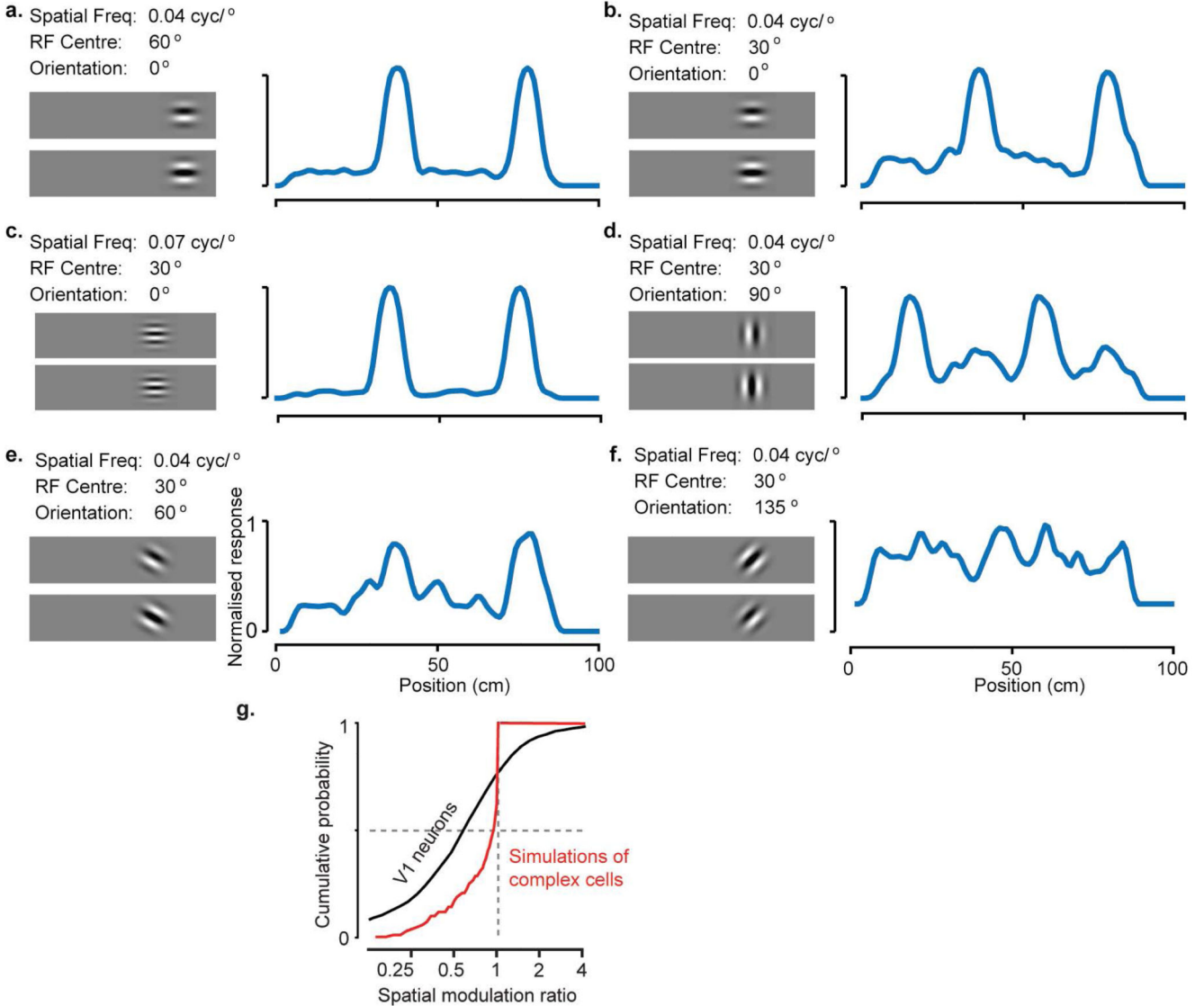
Extended Data Figure 3. Simulation of purely visual responses to position in VR.
a-f, Responses of 6 simulated neurons with purely visual responses, produced by a complex cell model with varying spatial frequency, orientation, or receptive field location. The images on the left of each panel show the quadrature pair of complex cell filters; traces on the right show the cell's simulated response as a function of position in the virtual environment. Simulation parameters matched those that are commonly observed in mouse

V1 (spatial frequency: 0.04, 0.05, 0.06 or 0.07 cyc/°; orientation: uniform between 0° and 90° but with twice more cells for cardinal orientations; receptive field positions 40°, 50°, 70°, 80°, similar to the V1 neurons we considered for analysis. In rare cases (like in **f**) when the receptive fields do not match the features of the environment, we get little selectivity along the corridor. These cases lead to lower value for the response ratios. **g,** The spatial modulation ratios calculated for the complex cell simulations are close to 1 (0.97 ± 0.17), and different from the ratios calculated for V1 neurons.
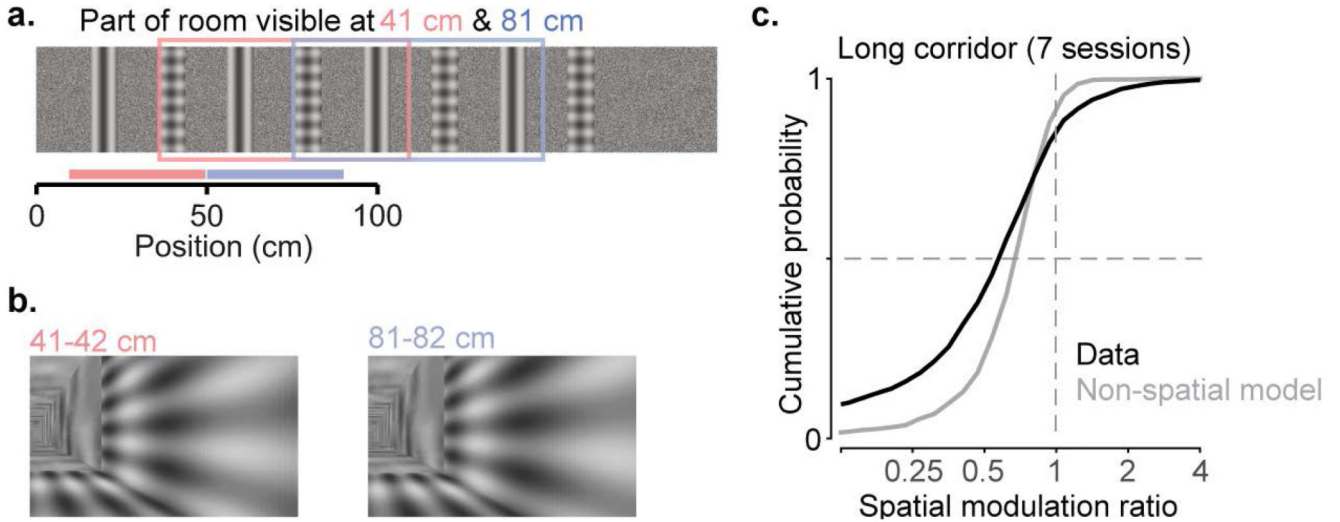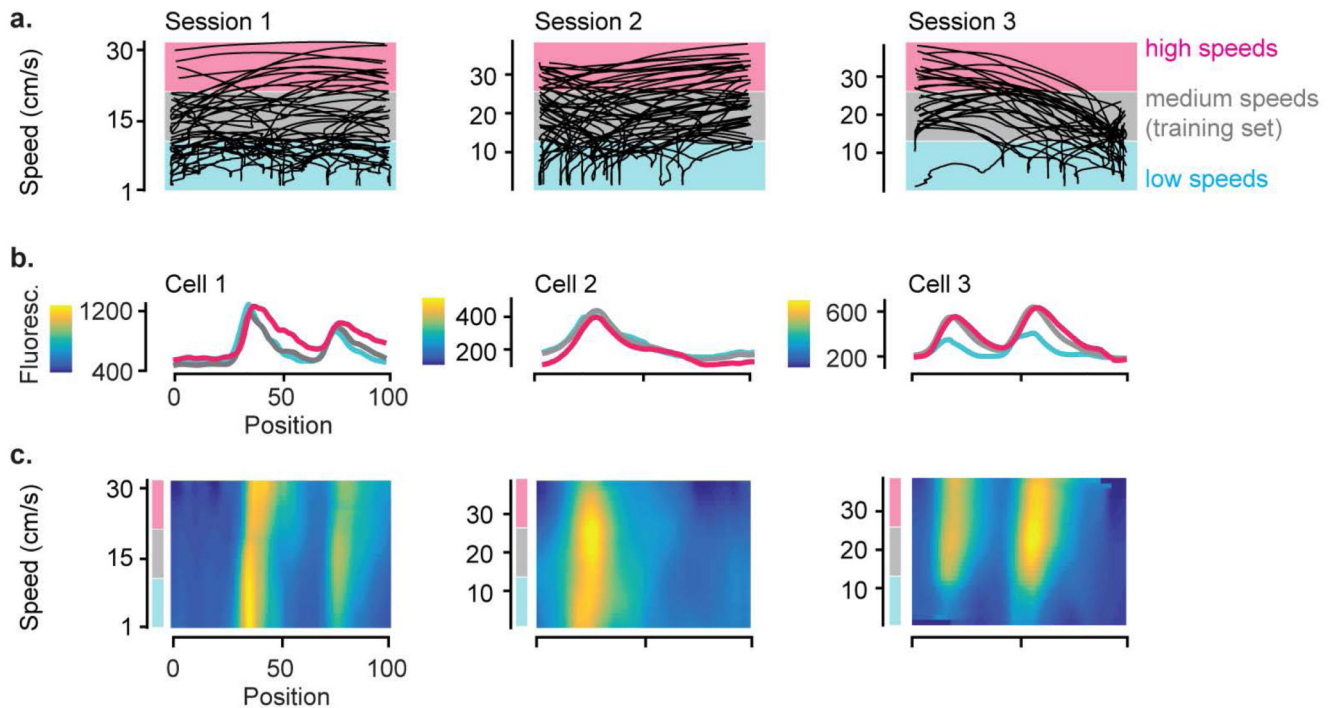

Extended Data Figure 4. The spatial modulation of V1 responses is not due to end-of-corridor visual cues

a, Diagram of the 200 cm virtual corridor, containing the same grating and plaid as the regular corridor, repeated 4 times instead of twice.

b, Visual scenes from locations within the first 100 cm of the extended corridor, separated by 40 cm, are visually (pixel-to-pixel) identical.

c, Cumulative distribution of the spatial modulation ratio across the two mice that were placed in the long corridor (7 sessions, 2 mice; median ± m.a.d: 0.62 ± 0.26; *black line*). *Grey line* shows the spatial modulation ratio predicted by the "non-spatial" model (which predicts activity from the visual scene, trial onset and offset, speed, reward, pupil size and displacement from the central position of the eye, see Extended Data Figure 7 below, non-spatial model). The two distributions are significantly different (two-sided Wilcoxon rank sum test; $p<10^{-14}$).

Extended Data Figure 5. The spatial modulation of V1 responses cannot be explained by speed.
a, Speed-position plots for all single-trial trajectories in three example recording sessions.
b, Response profile of example V1 cells in each session as a function of position in the corridor, stratified for three speed ranges corresponding to the shading bands in **a**.
c, Two-dimensional response profiles of the same example neurons showing activity as a function of position and running speed for speeds higher than 1 cm/s.
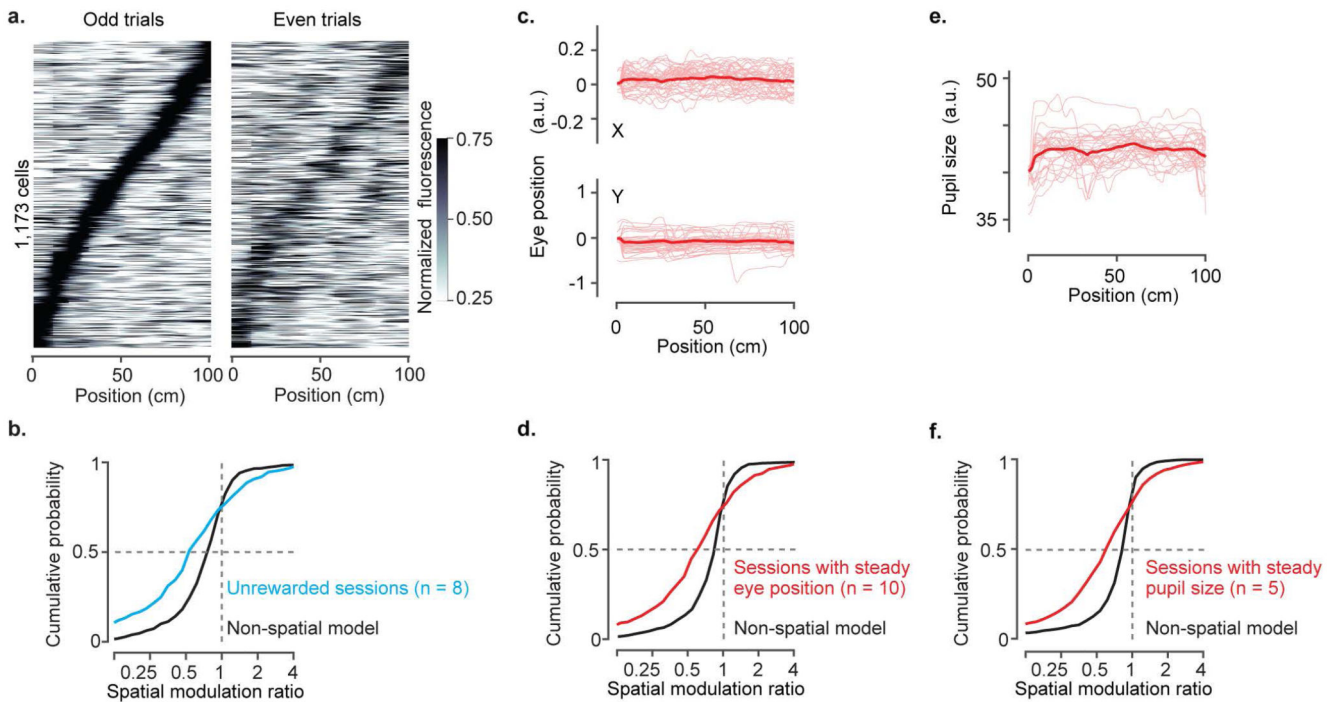
Extended Data Figure 6. The spatial modulation of V1 responses cannot be explained by reward, pupil position or diameter.

a, Normalised response as a function of position in the virtual corridor, for sessions without reward (1,173 cells). Data come from 2 out of 4 mice that ran the environment without reward (8 sessions, 2 mice). Responses in even trials (*right*) are ordered according to the position of maximum activity measured in odd trials (*left*).

b, Distribution of spatial modulation ratio for unrewarded sessions (8 sessions; median ± m.a.d.: 0.57 ± 0.37; *cyan*) and for modelled ratios obtained from the 'non-spatial' model on the same sessions (*black*). The two distributions are significantly different (two-sided Wilcoxon rank sum test; $p<10^{-8}$).

c, Pupil position as a function of location in the virtual corridor, for an example session with steady eye position. Sessions with steady eye positions were defined as those with no significant difference in eye positions between visually matching positions 40 cm apart (with unpaired t-test $p < 0.01$). Thin red curves: position trajectories on individual trials; thick curves, average. Top and bottom panels: x- and y-coordinates of the pupil.

d, Distribution of spatial modulation ratio for sessions with steady eye position (10 sessions; median ± m.a.d.: 0.63 ± 0.33; *red*) and for modelled ratios obtained from the 'non-spatial' model on the same sessions (*black*). The two distributions are significantly different (two-sided Wilcoxon rank sum test; $p<10^{-14}$).

e, Pupil size as a function of position for an example session with steady pupil size.

f, Distribution of spatial modulation ratio for sessions with steady pupil size (5 sessions; median ± m.a.d.: 0.63 ± 0.33; *red*) and for modelled ratios obtained from the 'non-spatial' model on the same sessions (*black*). The two distributions are significantly different (two-sided Wilcoxon rank sum test; $p<10^{-13}$).
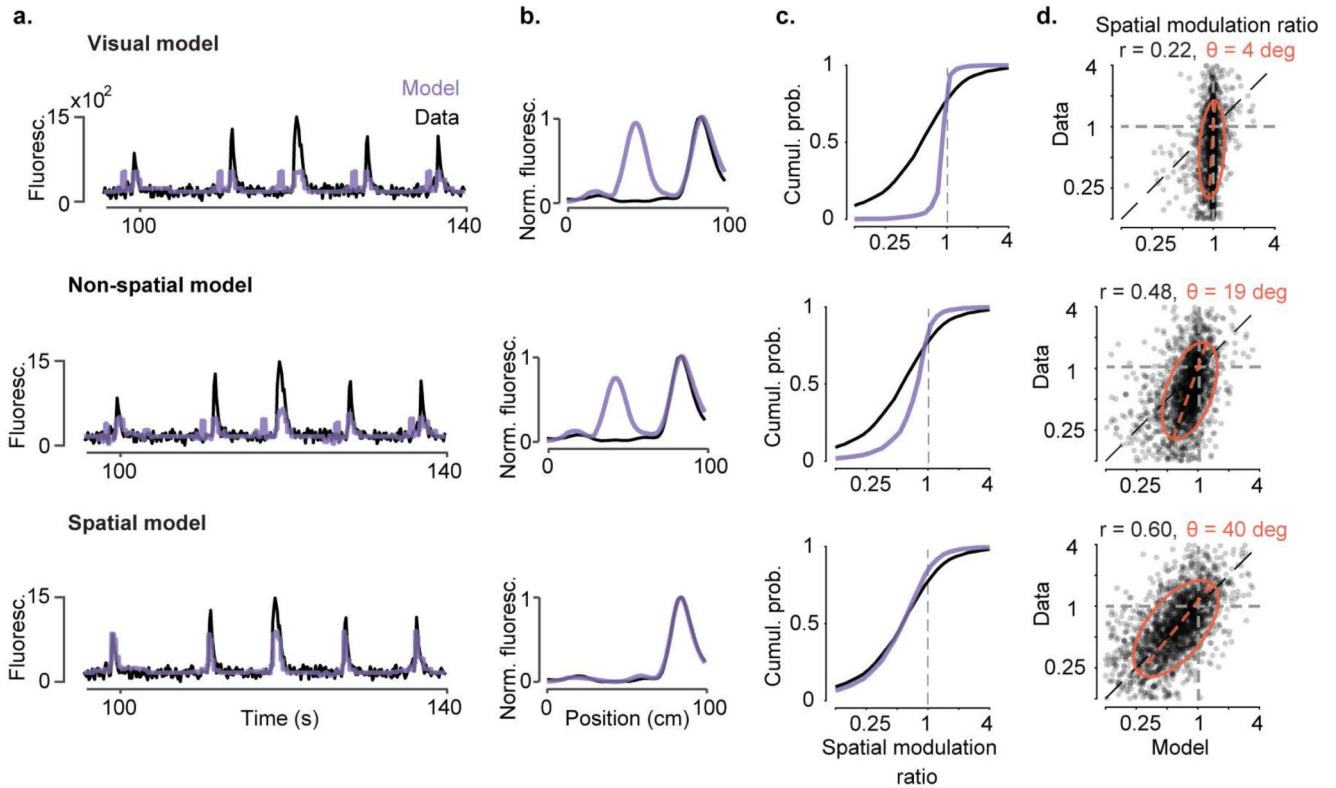
Extended Data Figure 7. Observed values of spatial modulation ratio can only be modelled using spatial position.

a-b, We constructed three models to predict the activity of individual V1 neurons from successively larger sets of predictor variables. In the simplest, the 'visual' model, activity is required to depend only on the visual scene visible from the mouse's current location, and is thus constrained to be a function of space that repeats in the visually matching section of the corridor. The second 'non-spatial' model also includes modulation by behavioural factors that can differ within and across trials: speed, reward times, pupil size, and eye position. Because these variables can differ between the first and second halves of the track, modelled responses need no longer be exactly symmetrical; however, this model does not explicitly use space as a predictor. The final 'spatial' model extends the previous model by also allowing responses to the two matching segments to vary in amplitude, thereby explicitly including space as a predictor. Example single-trial predictions are shown as a function of time in (a), together with measured fluorescence. Spatial profiles derived from these predictions are shown in (b).

c, Cumulative distributions of spatial modulation ratio for the three models (*purple curves*). For comparison, the *black curve* shows ratio of peaks derived from the data (even trials) (median ± m.a.d: visual model: 0.99 ± 0.03; two-sided Wilcoxon rank sum test: $p < 10^{-40}$, non-spatial model: 0.83 ± 0.18; $p < 10^{-40}$; spatial model: 0.60 ± 0.27; p = 0.09, n = 2,422 neurons)

d, Measured spatial modulation ratio versus prediction of the 3 models. Each point represents a cell; *red ellipse* represents best fit Gaussian, *dotted line* measures its slope. The purely visual model (*top*) does poorly, and is only slightly improved by adding predictions

from speed, reward, pupil size, and eye position (*middle*). Adding an explicit prediction from space provides a much better match to the data (*bottom*). r: Pearson's correlation coefficient, n = 2,422 neurons; θ: orientation of the major axis of the fitted ellipsoid.
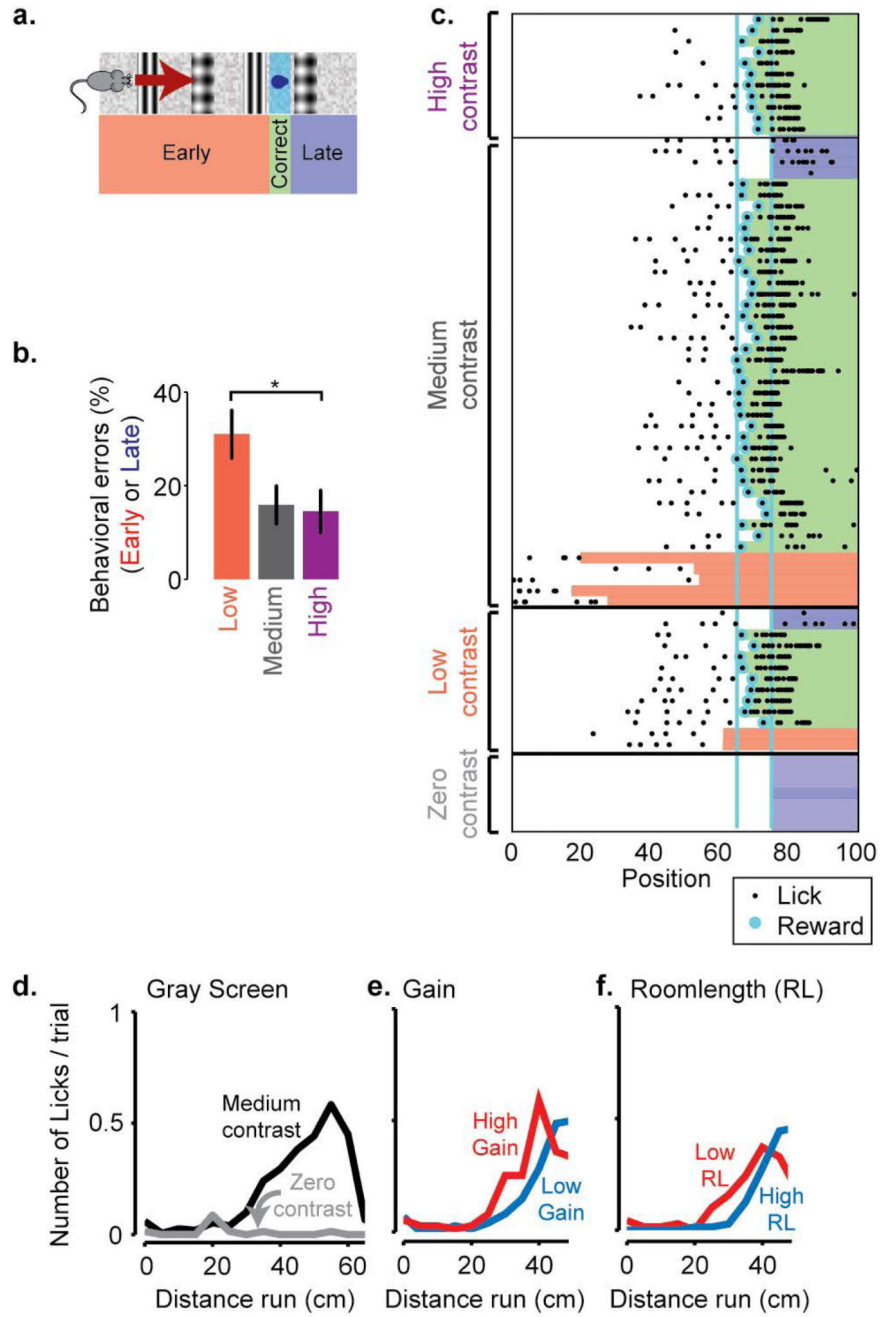
**Extended Data Figure 8. Behavioural performance in the task.**
a, Illustration of the virtual reality environment with four prominent landmarks, a reward zone, and the zones that define trial types: *Early*, *Correct* and *Late*.

b, Percentage of trials during which the animal makes behavioural errors, by licking either too early or too late at three different contrast levels: 18% (low), 60% (medium) or 72% (high).

c, Illustration of performance on all trials of one example recording session. Each row represents a trial, *black dots* represent positions where the animal licked, and *cyan dots* indicate the delivery of a water reward. *Coloured bars* indicate the outcome of the trial, *red*: *Early*, *green*: *Correct*, *blue*: *Late*.

d-f, Successful performance relies on vision: **d,** The animal does not lick when the room was presented at zero contrast. **e,** On some trials, we changed the gain between the animals' physical movement and movement in the virtual environment, thus changing the distance to the reward zone (high gain resulting in shorter distance), while visual cues remained in the same place. When plotted as a function of the distance run, the licks of the animal shifted, indicating that the animal was not simply relying on the distance travelled from the beginning of the corridor. **f,** If the position of the visual cues was shifted forward or back (high/low room length (RL)), the lick position shifted accordingly, indicating that the animals rely on vision to perform the task.
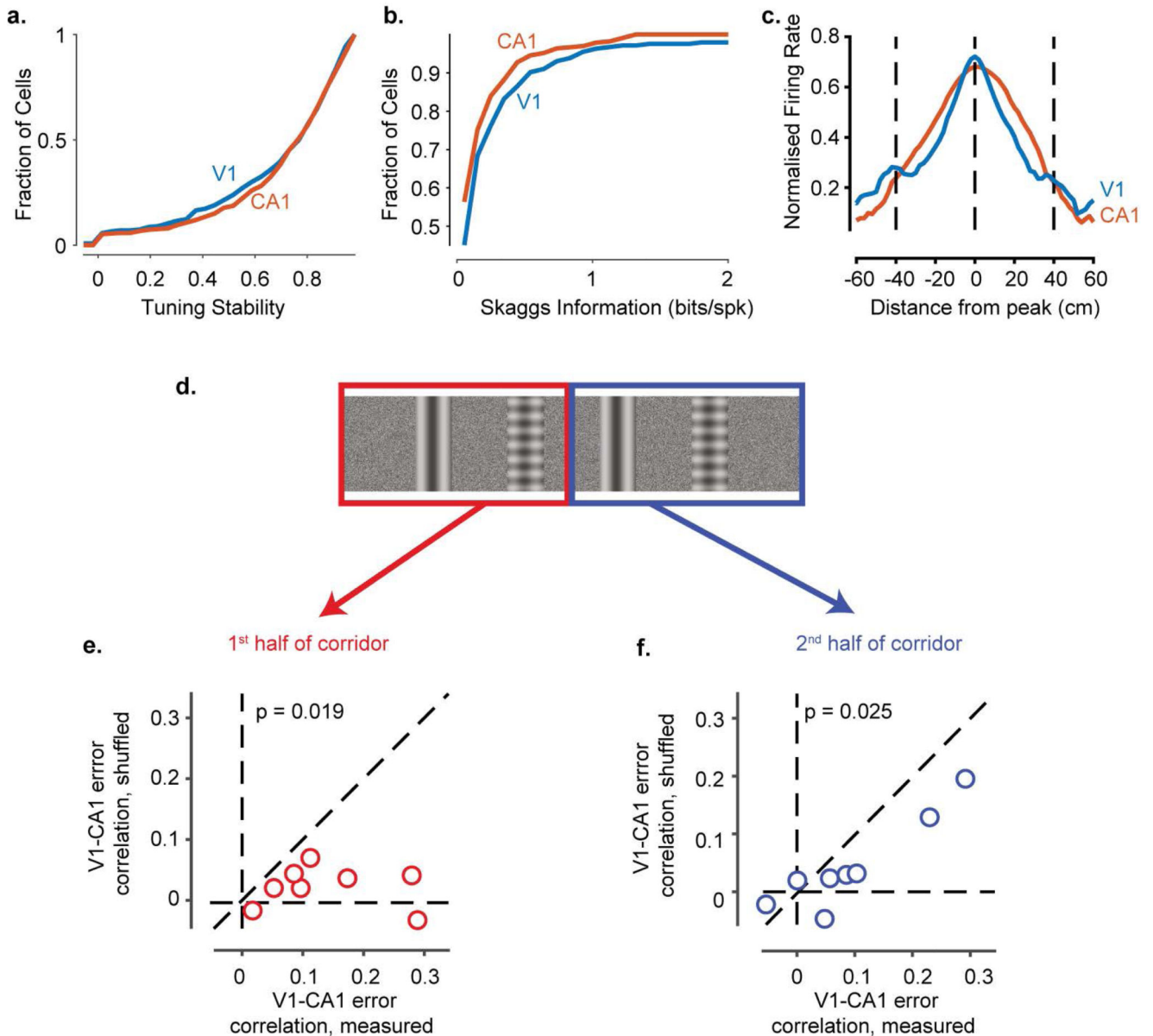
Extended Data Figure 9. Comparison of response properties between V1 and CA1 neurons and correlation of V1 and CA1 errors in the two halves of the environment
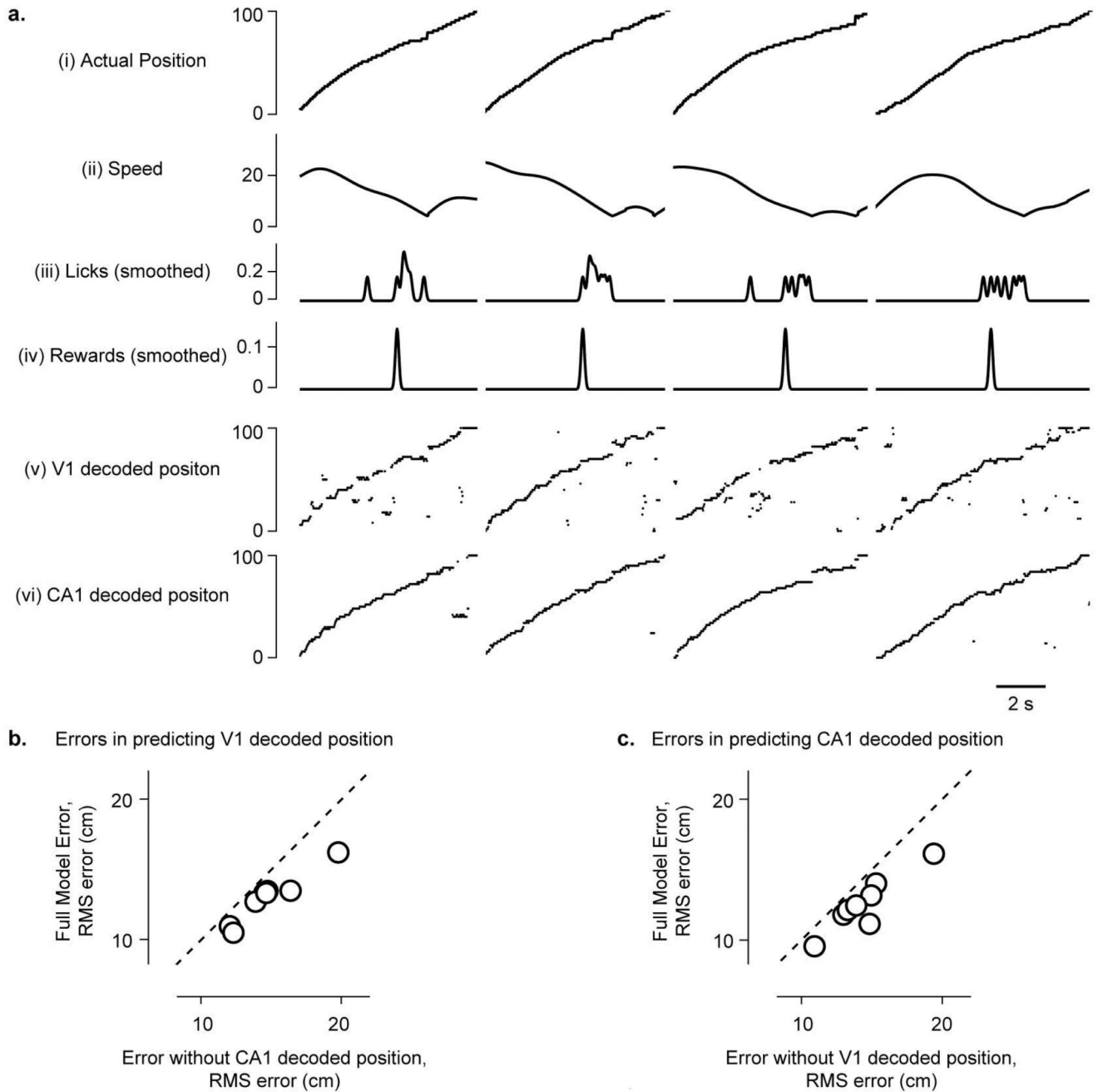
a, Cumulative distribution of the stability of V1 and CA1 response profiles: tuning stability - the stability of responses was computed as the correlation between the spatial responses measured from the first half and the second half of the trials. V1 and CA1 responses were highly stable within each recording session: the tuning stability was > 0.7 for more than 60% of neurons in both V1 and CA1.

b, Cumulative distribution of the Skaggs information (bits/spike) carried by V1 and CA1 neurons. Note that while V1 and CA1 neurons had comparable amounts of spatial information, this does not suggest that V1 represents space as strongly as CA1 since the Skaggs information metric mixes the influence of vision and spatial modulation.

c, Normalised firing rate averaged across V1 or CA1 neurons as a function of distance from the peak response (Similar to Extended Data Figure **2b**). Unlike CA1, the mean response averaged from V1 neurons shows a second peak at +/- 40 cm, consistent with the repetition of the visual scene.

d-e, Pearson's correlation between position errors estimated from V1 and CA1 populations in the first half of the corridor. Each point represents a behavioural session (n = 8 sessions); x-axis values represent measured correlations; y-axis values represent correlations calculated after having shuffled the data within the times where the speed was similar (similar to Figure 2h). The occurrence of error correlations in the unshuffled data indicates that these correlations are not due to rewards (which did not occur in this half of the maze) or licks (which were rare, and a 100 ms period surrounding the few that occurred were removed from analysis). The significance of the difference between the measured and shuffled correlations was calculated using a two-sided two-sample t-test.

f, Similar to **e** for the 2nd half of the corridor.

a.

(i) Actual Position

(ii) Speed

(iii) Licks (smoothed)

(iv) Rewards (smoothed)

(v) V1 decoded positon

(vi) CA1 decoded positon

2 s

b. Errors in predicting V1 decoded position

c. Errors in predicting CA1 decoded position

Full Model Error, RMS error (cm)

Error without CA1 decoded position, RMS error (cm)

Error without V1 decoded position, RMS error (cm)

Extended Data Figure 10. Position decoded from CA1 activity helps predict position decoded from V1 activity (and vice versa).

a, To test whether the positions encoded in V1 and CA1 populations are correlated with each other beyond what would be expected from a common influence of other spatial and non-spatial factors, we used a Random Forests decoder (Tree Bagger implementation in MATLAB) to predict V1 or CA1 decoded positions from different predictors. We then tested if the model prediction was further improved when we added the position decoded from the other area as an additional predictor (i.e. using the positions decoded from CA1 to predict V1 decoded positions and vice versa).

b, Adding CA1 decoded position as an additional predictor improved the prediction of V1 decoded positions in every recording session (i.e. reduced the prediction errors). V1 and CA1 decoded positions are thus correlated with each other beyond what can be expected from a common contribution of position, speed, licks and reward to V1 and CA1 responses. **c,** Same as **b** for predicting CA1 decoded position.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

References

1. Muller RU, Kubie JL. The effects of changes in the environment on the spatial firing of hippocampal complex-spike cells. J Neurosci. 1987; 7:1951–1968. [PubMed: 3612226]

2. Wiener SI, Korshunov VA, Garcia R, Berthoz A. Inertial, Substratal and Landmark Cue Control of Hippocampal CA1 Place Cell Activity. Eur J Neurosci. 1995; 7:2206–2219. [PubMed: 8563970]

3. Chen G, King JA, Burgess N, O'Keefe J. How vision and movement combine in the hippocampal place code. Proc Natl Acad Sci U S A. 2013; 110:378–383. [PubMed: 23256159]

4. Geiller T, Fattahi M, Choi J-S, Royer S. Place cells are more strongly tied to landmarks in deep than in superficial CA1. Nat Commun. 2017; 8

5. Ji D, Wilson MA. Coordinated memory replay in the visual cortex and hippocampus during sleep. Nat Neurosci. 2007; 10:100–107. [PubMed: 17173043]

6. Haggerty DC, Ji D. Activities of visual cortical and hippocampal neurons co-fluctuate in freely moving rats during spatial behavior. Elife. 2015; 4

7. Fiser A, et al. Experience-dependent spatial expectations in mouse visual cortex. Nat Neurosci. 2016; 19:1658–1664. [PubMed: 27618309]

8. Niell CM, Stryker MP. Modulation of Visual Responses by Behavioral State in Mouse Visual Cortex. Neuron. 2010; 65:472–479. [PubMed: 20188652]

9. Saleem AB, Ayaz A, Jeffery KJ, Harris KD, Carandini M. Integration of visual motion and locomotion in mouse visual cortex. Nat Neurosci. 2013; 16:1864–1869. [PubMed: 24185423]

10. Ravassard P, et al. Multisensory Control of Hippocampal Spatiotemporal Selectivity. Science (80-.). 2013; 340:1342–1346.

11. Harvey CD, Collman F, Dombeck DA, Tank DW. Intracellular dynamics of hippocampal place cells during virtual navigation. Nature. 2009; 461:941–946. [PubMed: 19829374]

12. McNaughton BL, Barnes CA, O'Keefe J. The contributions of position, direction, and velocity to single unit activity in the hippocampus of freely-moving rats. Exp Brain Res. 1983; 52:41–49. [PubMed: 6628596]

13. Wiener SI, Paul CA, Eichenbaum HB. Spatial and behavioral correlates of hippocampal neuronal activity. J Neurosci. 1989; 9:2737–63. [PubMed: 2769364]

14. Czurkó A, Hirase H, Csicsvari J, Buzsáki G. Sustained activation of hippocampal pyramidal cells by 'space clamping' in a running wheel. Eur J Neurosci. 1999; 11:344–352. [PubMed: 9987037]

15. O'Keefe J, Speakman A. Single unit activity in the rat hippocampus during a spatial memory task. Exp Brain Res. 1987; 68:1–27. [PubMed: 3691688]

16. Lenck-Santini PP, Save E, Poucet B. Evidence for a relationship between place-cell spatial firing and spatial memory performance. Hippocampus. 2001; 11:377–390. [PubMed: 11530842]

17. Lenck-Santini P-P, Muller RU, Save E, Poucet B. Relationships between place cell firing fields and navigational decisions by rats. J Neurosci. 2002; 22:9035–9047. [PubMed: 12388610]

18. Hok V, et al. Goal-related activity in hippocampal place cells. J Neurosci. 2007; 27:472–482. [PubMed: 17234580]

19. Rosenzweig ES, Redish AD, McNaughton BL, Barnes CA. Hippocampal map realignment and spatial learning. Nat Neurosci. 2003; 6:609–615. [PubMed: 12717437]

20. Makino H, Komiyama T. Learning enhances the relative impact of top-down processing in the visual cortex. Nat Neurosci. 2015; 18:1116–1122. [PubMed: 26167904]

21. Poort J, et al. Learning Enhances Sensory and Multiple Non-sensory Representations in Primary Visual Cortex. Neuron. 2015; 86:1478–1490. [PubMed: 26051421]

22. Jurjut O, Georgieva P, Busse L, Katzner S. Learning Enhances Sensory Processing in Mouse V1 before Improving Behavior. J Neurosci. 2017; 37:6460–6474. [PubMed: 28559381]

23. Witter MP, et al. Cortico-hippocampal communication by way of parallel parahippocampal-subicular pathways. Hippocampus. 2000; 10:398–410. [PubMed: 10985279]

24. Wang Q, Gao E, Burkhalter A. Gateways of Ventral and Dorsal Streams in Mouse Visual Cortex. J Neurosci. 2011; 31:1905–1918. [PubMed: 21289200]

25. Moser EI, Kropff E, Moser M-B. Place Cells, Grid Cells, and the Brain's Spatial Representation System. Annu Rev Neurosci. 2008; 31:69–89. [PubMed: 18284371]

26. Grieves RM, Jeffery KJ. The representation of space in the brain. Behavioural Processes. 2017; 135:113–131. [PubMed: 28034697]

27. Eichenbaum H. Time cells in the hippocampus: A new dimension for mapping memories. Nature Reviews Neuroscience. 2014; 15:732–744. [PubMed: 25269553]

28. Cushman JD, et al. Multisensory control of multimodal behavior: Do the legs know what the tongue is doing? PLoS One. 2013; 8

29. Aronov D, Tank DW. Engagement of Neural Circuits Underlying 2D Spatial Navigation in a Rodent Virtual Reality System. Neuron. 2014; 84:442–456. [PubMed: 25374363]

30. Chen G, King JA, Lu Y, Cacucci F, Burgess N. Spatial cell firing during virtual navigation of open arenas by head-restrained mice. bioRxiv. 2018; doi: 10.1101/246744

31. Madisen L, et al. Transgenic mice for intersectional targeting of neural sensors and effectors with high specificity and performance. Neuron. 2015; 85:942–958. [PubMed: 25741722]

32. Wekselblatt JB, Flister ED, Piscopo DM, Niell CM. Large-scale imaging of cortical dynamics during sensory perception and behavior. J Neurophysiol. 2016; 115:2852–2866. [PubMed: 26912600]

33. Steinmetz NA, et al. Aberrant Cortical Activity in Multiple GCaMP6-Expressing Transgenic Mouse Lines. eneuro. 2017; 4

34. Ayaz A, Saleem AB, Schölvinck ML, Carandini M. Locomotion controls spatial integration in mouse visual cortex. Curr Biol. 2013; 23:890–894. [PubMed: 23664971]

35. Ratzlaff EH, Grinvald A. A tandem-lens epifluorescence macroscope: Hundred-fold brightness advantage for wide-field imaging. J Neurosci Methods. 1991; 36:127–137. [PubMed: 1905769]

36. Carandini M, et al. Imaging the Awake Visual Cortex with a Genetically Encoded Voltage Indicator. J Neurosci. 2015; 35:53–63. [PubMed: 25568102]

37. Kalatsky VA, Stryker MP. New paradigm for optical imaging: temporally encoded maps of intrinsic signal. Neuron. 2003; 38:529–45. [PubMed: 12765606]

38. Yang Z, Heeger DJ, Seidemann E. Rapid and precise retinotopic mapping of the visual cortex obtained by voltage-sensitive dye imaging in the behaving monkey. J Neurophysiol. 2007; 98:1002–14. [PubMed: 17522170]

39. Pologruto TA, Sabatini BL, Svoboda K. ScanImage: Flexible software for operating laser scanning microscopes. Biomed Eng Online. 2003; 2:13. [PubMed: 12801419]

40. Pachitariu M, et al. Suite2p: beyond 10,000 neurons with standard two-photon microscopy. bioRxiv. 2016; doi: 10.1101/061507

41. Peron SP, Freeman J, Iyer V, Guo C, Svoboda K. A Cellular Resolution Map of Barrel Cortex Activity during Tactile Behavior. Neuron. 2015; 86:783–99. [PubMed: 25913859]

42. Dipoppa M, et al. Vision and Locomotion Shape the Interactions between Neuron Types in Mouse Visual Cortex. Neuron. 2018; 98:602–615.e8. [PubMed: 29656873]

43. Rossant C, et al. Spike sorting for large, dense electrode arrays. Nat Neurosci. 2016; 19:634–641. [PubMed: 26974951]

44. Kadir SN, Goodman DFM, Harris KD. High-Dimensional Cluster Analysis with the Masked EM Algorithm. Neural Comput. 2014; 26:2379–2394. [PubMed: 25149694]
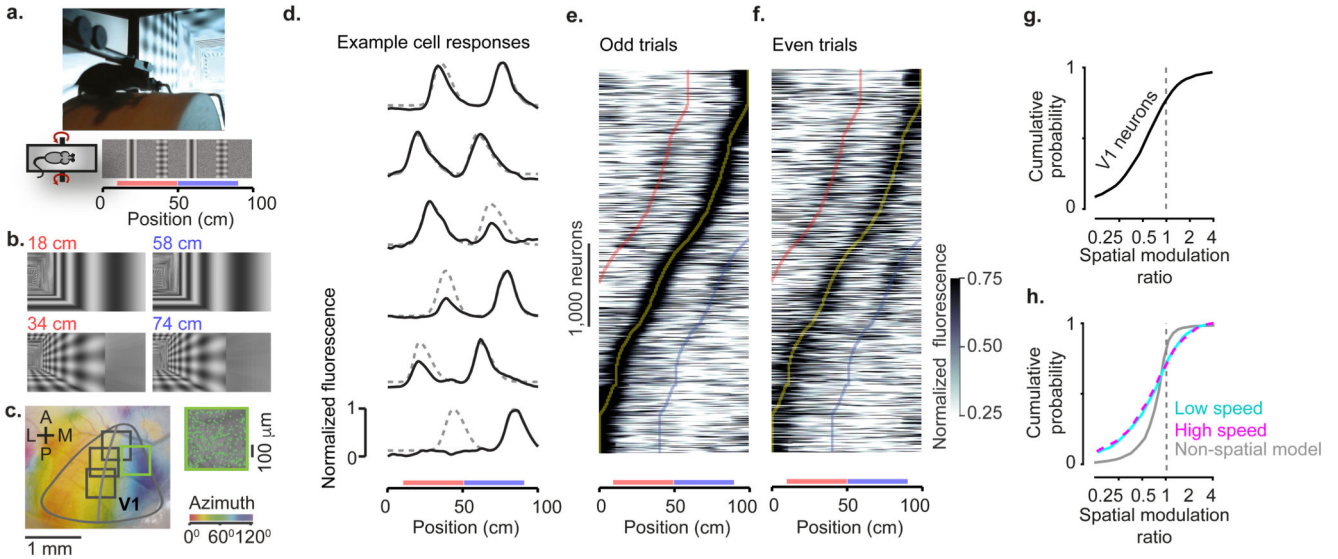
Figure 1. Responses in the primary visual cortex (V1) are modulated by spatial position.
a, Mice ran on a cylindrical treadmill to navigate a virtual corridor. The corridor had two landmarks that repeated after 40 cm, creating visually-matching segments (*red* and *blue bars*).
b, Screenshots showing the right half of the corridor at pairs of positions 40 cm apart.
c, Example retinotopic map of the cortical surface. *Grey curve* shows the border of V1. *Squares* denote the field of view in two-photon imaging sessions targeted to medial V1 (*inset* shows the field with *green frame*). We analysed responses from neurons with receptive field centre > 40° azimuth (*curve*).
d, Normalised response as a function of position in the corridor for six example V1 neurons. *Dotted lines* show predictions assuming identical responses in matching segments of the corridor.
e, Normalised response as a function of position, obtained from odd trials, for 4,958 V1 neurons. Neurons are ordered based on the position of their maximum response.
f, Same as **e**, for even trials. Neurons are ordered and scaled as in **e**. Curves indicate preferred position (*Yellow*) and preferred position ± 40 cm (*blue* and *red*).
g, Cumulative distribution of the spatial modulation ratio in even trials: response at non-preferred position (40 cm away from peak response) divided by response at preferred position for cells with responses within the visually matching segments.
h, Same as in **g** while stratifying the data by running speed and comparing to a model without spatial selectivity. The curves corresponding to low (*cyan*) and high (*purple*) speeds overlap and appear as a single dashed curve. *Grey* curve: spatial modulation ratios from a non-spatial model considering visual and behavioural factors (Extended Data Figure 7).
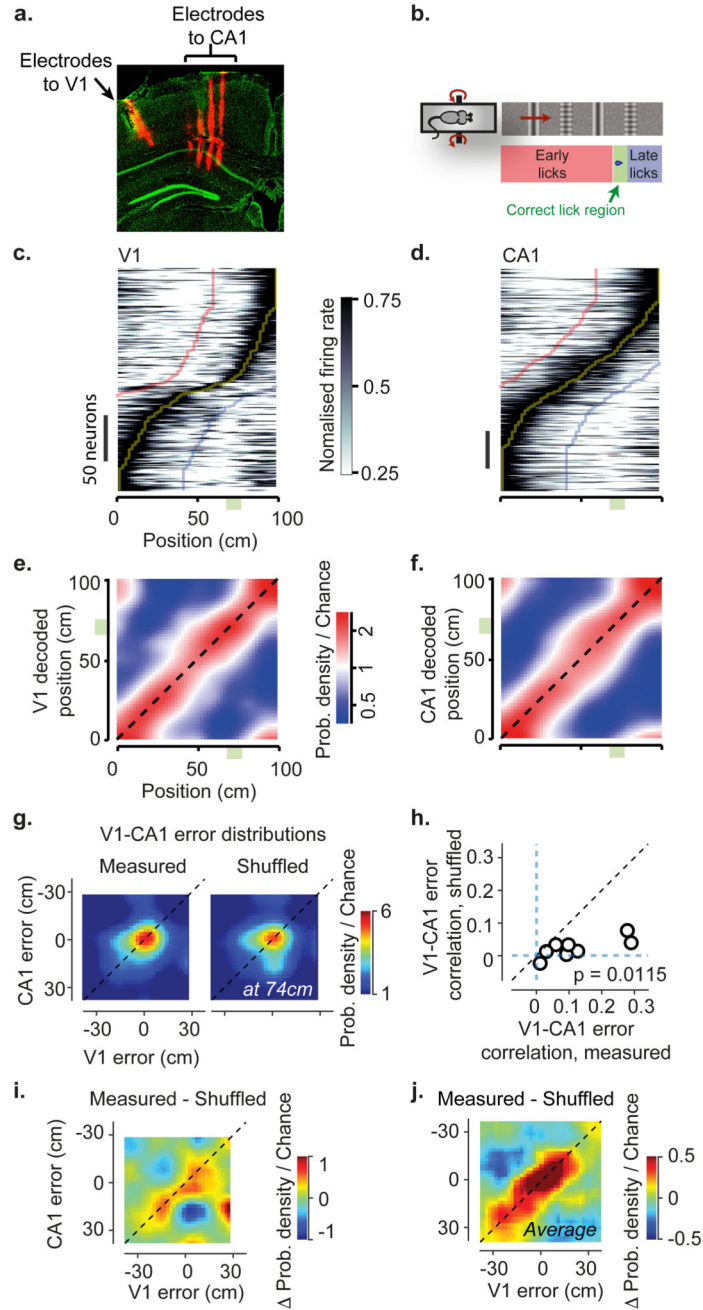
Figure 2. V1 and CA1 neural populations represent spatial positions in the virtual corridor and make correlated errors.

a, Example of reconstructed electrode tracks (*red*: DiI); *green* shows cells labelled by DAPI. *Panel* shows tracks from one array (four shanks) in CA1, and a second electrode (one shank) in V1.

b, In the task, water was delivered when mice licked in a reward zone (indicated by *green area*).

c, Normalised activity as a function of position in the corridor, for 226 V1 neurons (8 sessions). Neurons are ordered based on the position of their maximum response. Curves indicate preferred position (*Yellow*) and preferred position ± 40 cm (*blue* and *red*).

d, Similar plot for CA1 place cells (334 neurons; 8 sessions).

e, Density map showing the distribution of position decoded from the activity of simultaneously recorded V1 neurons (y-axis) as a function of the animal's position (x-axis), averaged across recording sessions (n = 8), and considering only correct trials. The *red diagonal stripe* indicates accurate estimation of position.

f, Similar plot for CA1 neurons.

g, Density map showing the joint distribution of position decoding errors from V1 and CA1 at one position (74 cm; *left*), together with a similar analysis on data shuffled preserving the correlation due to run speed and position (*right*).

h, Pearson's correlation coefficient of decoding errors in V1 and CA1 for each recording session (n = 3,800 - 21,000 time points), against similar analysis of shuffled data. Correlations are above shuffling control (p=0.0115, two-sided t-test, n = 8 sessions).

i, Difference between joint distribution of V1 and CA1 decoded position and shuffled control, for the example in **g**.

j, Difference between joint density map of V1 and CA1 decoded position, and shuffled control, averaged across positions (n = 50) and sessions (n = 8).
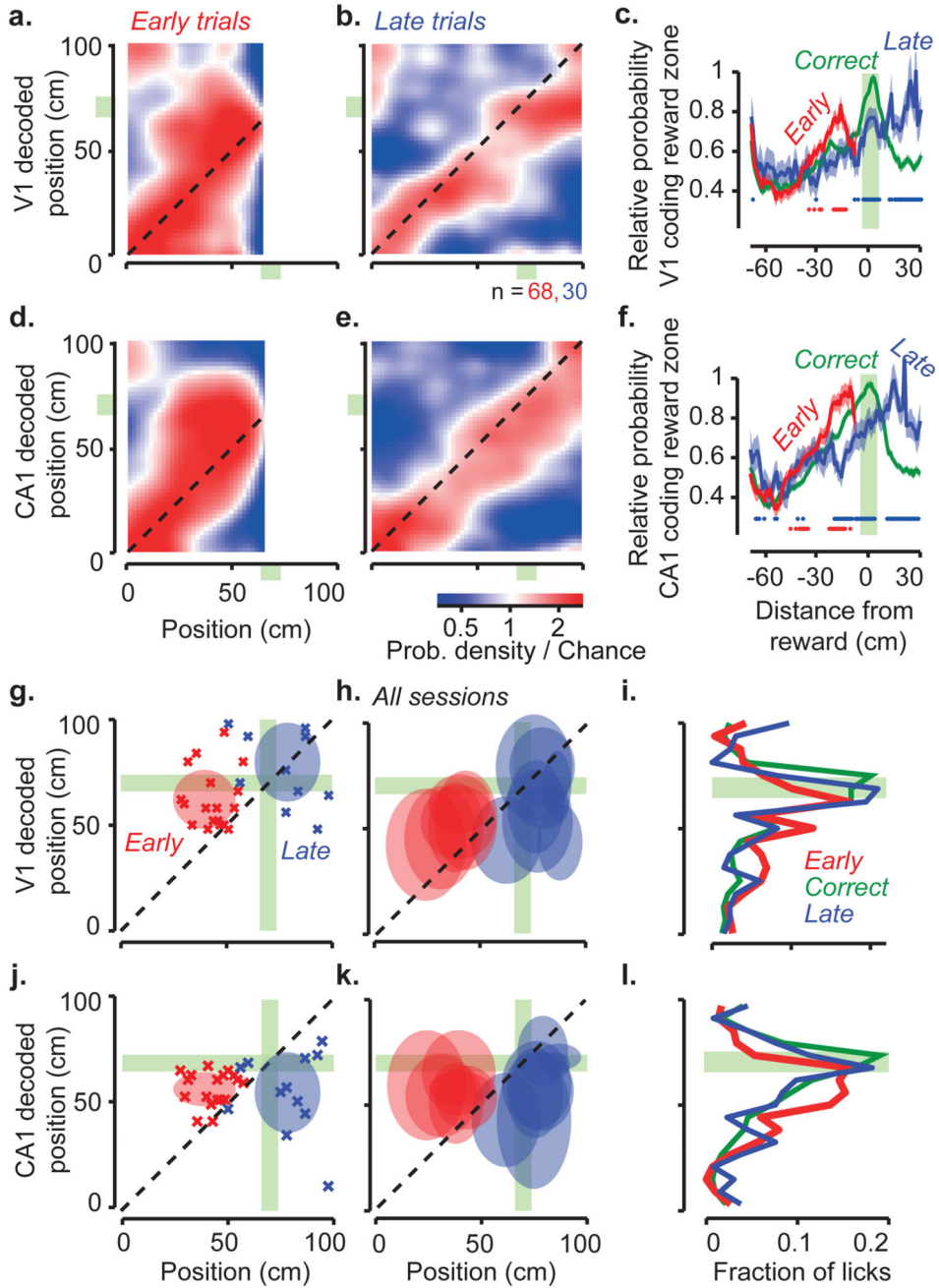
Figure 3. Positions encoded by visual cortex and hippocampus correlate with animal's spatial decisions.

a, Distribution of positions decoded from the V1 population, as a function of the animal's actual position, on trials where mice licked *early*. The decoder was trained on separate trials where mice licked in the correct position.

b, Same plot for trials where mice licked *late*.

c, The average decoded probability for the animal to be in the reward zone, as a function of distance from the reward. The curve for *early* trials (*red*) peaks before the reward zone,

while the curve for *late* trials (*blue*) peaks after, consistent with V1 activity reflecting subjective position, rather than actual position. Probabilities were normalised relative to the probability of being in the reward zone in the *correct* trials (*green*). *Red dots*: positions where the decoded probability of being in the reward zone differed significantly between *Early* and *Correct* trials ($p<0.05$, two-sample two-sided t-test). *Blue dots*: same, for *Correct* vs. *Late* trials. Shaded regions indicate mean ± s.e.m, n = 68 *Early* trials (*red*), 334 *Correct* trials (*green*), and 30 *Late* trials (*blue*).

d-f, Same as **a-c**, for decoding using the population of CA1 neurons.

g, Position decoded from V1 activity as a function of mouse position, in an example session. *Crosses* show positions when the animal licked on *Early* (*red*) or *Late* trials (*blue*). *Late* trials can include some early licks. These distributions (mean ± s.d.) are summarized as *shaded ovals* for *Early* (*red*, n = 20) and *Late* (*blue*, n = 12) trials. G*reen regions* mark the reward zone.

h, Summary distributions for all sessions (n = 8).

i, Fraction of licks as a function of distance from reward location in positions decoded from V1 activity.

j-l, Same as **g-i**, for CA1 neurons.