Article

Frontal cortex neuron types categorically encode single decision variables

https://doi.org/10.1038/s41586-019-1816-9

Received: 8 May 2017

Accepted: 15 October 2019

Published online: 04 December 2019

Junya Hirokawa[1,2,5], Alexander Vaughan[1,5], Paul Masset[1,3,4], Torben Ott[1] & Adam Kepecs[1]*

Individual neurons in many cortical regions have been found to encode specific, identifiable features of the environment or body that pertain to the function of the region[1–3]. However, in frontal cortex, which is involved in cognition, neural responses display baffling complexity, carrying seemingly disordered mixtures of sensory, motor and other task-related variables[4–13]. This complexity has led to the suggestion that representations in individual frontal neurons are randomly mixed and can only be understood at the neural population level[14,15]. Here we show that neural activity in rat orbitofrontal cortex (OFC) is instead highly structured: single neuron activity co-varies with individual variables in computational models that explain choice behaviour. To characterize neural responses across a large behavioural space, we trained rats on a behavioural task that combines perceptual and value-guided decisions. An unbiased, model-free clustering analysis identified distinct groups of OFC neurons, each with a particular response profile in task-variable space. Applying a simple model of choice behaviour to these categorical response profiles revealed that each profile quantitatively corresponds to a specific decision variable, such as decision confidence. Additionally, we demonstrate that a connectivity-defined cell type, orbitofrontal neurons projecting to the striatum, carries a selective and temporally sustained representation of a single decision variable: integrated value. We propose that neurons in frontal cortex, as in other cortical regions, form a sparse and overcomplete representation of features relevant to the region's function, and that they distribute this information selectively to downstream regions to support behaviour.

The brain represents the external world in patterns of neural activity that guide adaptive behaviour. In many regions, individual cortical neurons respond to features, such as visual edges[1], objects[2] or spatial locations[3], that reflect regional function. When examining frontal areas engaged in decision making, however, one is struck most of all by the complexity and diversity of their neuronal responses[4–14,16–18]. The difficulty in identifying structure in frontal cortical representations probably reflects the fact that cognitive variables are more challenging to define than simpler features such as visual edges. It is further challenging to design behavioural tasks that engage the specific cognitive functions of frontal cortical neurons to sufficiently probe the relevant feature space. As a result, it is unclear whether frontal cortex representations are comprehensible in single neurons or instead require neural population analysis (for example, 'random mixed selectivity'[14,15], but see ref. [19]).

Behavioural task engaging orbitofrontal cortex

We trained Long Evans rats on a complex reward-biased psychometric olfactory discrimination task (Fig. 1a) that requires integration of decision confidence and reward value. This task separately varies perceptual uncertainty across trials and reward expectations across blocks. Choice accuracy varied systematically with stimulus difficulty; changing reward size induced rapid and sustained changes in choice biases and reaction times[20] (Fig. 1b, c and Extended Data Fig. 1a–e; see Methods), confirming that animals' strategy maximized total reward (Extended Data Fig. 1f, g; see Methods). A model encoding key decision variables accounts for overall choice strategy (Fig. 1d, e) and trial-by-trial biases arising from previous outcomes (Extended Data Fig. 1h–l).

We recorded 485 neurons from lateral orbitofrontal cortex (OFC) in three rats (Extended Data Fig. 2) and analysed the post-choice epoch when rats await an uncertain reward ('reward anticipation', Fig. 1a). We identified OFC neurons that encode canonical variables in our model, including decision confidence, with activity proportional to the evidence supporting the choice, but not influenced by expected reward size[8,21] (Fig. 2a, b). We also observed representations of other canonical decision variables, including anticipated reward size (assuming a correct choice) and integrated value (that is, probability of reward for a given choice multiplied by reward size; Extended Data Fig. 3a).

[1]Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, USA. [2]Graduate School of Brain Science, Doshisha University, Kyotanabe, Kyoto, Japan. [3]Watson School of Biological Sciences, Cold Spring Harbor, NY, USA. [4]Present address: Department of Molecular and Cellular Biology, Centre for Brain Science, Harvard University, Cambridge, MA, USA. [5]These authors contributed equally: Junya Hirokawa, Alexander Vaughan. *e-mail: kepecs@cshl.edu

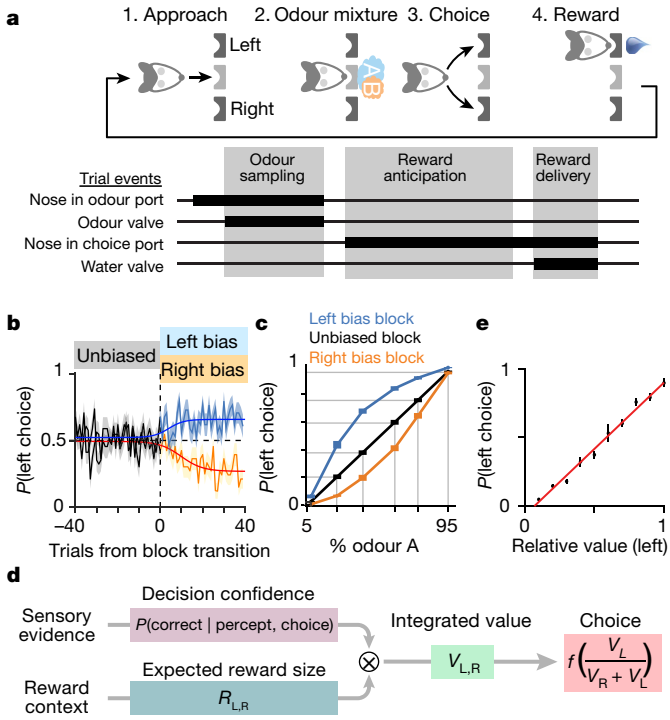


Fig. 1 | Reward-biased psychometric odour discrimination task to probe decision-variable integration. a, Task design, single trial. We primarily analysed the 1 s reward anticipation window. **b**, Behavioural performance is modified by block-wise changes in reward size. Choice biases emerged rapidly across blocks (5.45 ± 0.48 trials to bias shift, mean ± s.e.m., $n = 67$ sessions from 3 rats). **c**, Average psychometric functions in unbiased, left- and right-bias blocks reveal systematic bias (30 sessions from one rat (C068)). **d**, Decision-variable model of choice behaviour in this task integrates variables from reward context (choice, size of potential reward, presence of reward) and sensory evidence to generate internal variables representing expected reward value and decision confidence. These combine to estimate integrated value, which determines choice probability. **e**, The relative value derived from the model explains choice behaviour (mean ± s.e.m., $n = 67$ sessions from 3 rats).

Clustered single neuron representations

We next considered whether these OFC neurons were representative or outliers from a diverse population (Fig. 2b–g). We generated response profiles for each neuron across 42 task contingencies combining odour stimulus, reward size, behavioural choice, and previous trial outcome (Fig. 2b and Extended Data Fig. 3b–d). We developed two improved statistical tests to examine deviations from random mixed selectivity (ePAIRS and eRP) and showed that OFC neurons do not encode random mixtures of available information[22,23] (Extended Data Fig. 4a–l; see Methods). Instead, similarly tuned neurons are more common than expected (Fig. 2d, e and Extended Data Fig. 4e–m), and neuronal activity does not uniformly fill the space of available representations (Extended Data Fig. 3e, f).

We used a model-free approach to identify clusters of OFC neurons with similar response profiles. To alleviate the challenges of a high-dimensional response space, we used nonlinear spectral clustering with bootstrap validation to identify clusters as strongly connected subcomponents of a neighbourhood graph (Fig. 2c; Methods). This analysis identified ~9 robust clusters (Fig. 2f), each encoding one 'categorical' representation (Fig. 2g).

As controls, we confirmed that clustering did not arise from segregation of neurons from different rats (Extended Data Fig. 5a, b), spatial patterning in the OFC (Supplementary Note 1 and Extended Data Fig. 5c–h), or data pre-processing (Extended Data Fig. 5i, j; see Methods). Additionally, our approach was biased against finding such clusters in several ways (see Supplementary Discussion).

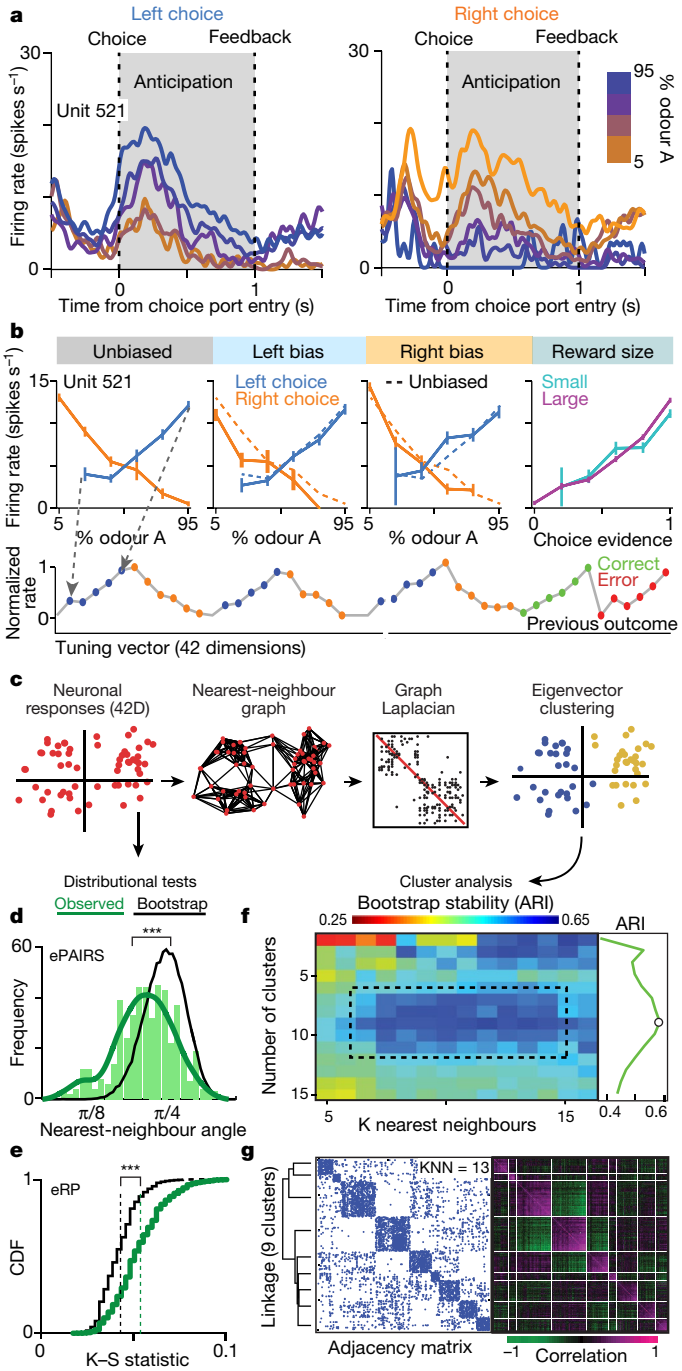


Fig. 2 | OFC neurons form nine discrete clusters. a, **b**, Response during reward anticipation for one neuron (unit 521). Activity in the unbiased block reflected evidence supporting animal's choice (**a**), and matched expected response profile of representation of decision confidence (varying with choice evidence, but not reward size) (**b**). **c**, Analysis pipeline for cluster analysis. **d**, **e**, Distribution of nearest-neighbour angles (**d**) and projection angles (**e**) reveal non-random distribution in OFC population with angle magnitudes suggesting clustering (***$P < 0.001$ for both tests; Methods, Extended Data Fig. 4). **f**, Spectral clustering reveals nine robust groups of neurons, with high cluster stability across a range of hyperparameters (dashed box; adjusted rand index, ARI; see Methods for details of cross-validation). **g**, Sorted adjacency and correlation matrices reveal strong within-cluster similarity and between-cluster antagonism. CDF, cumulative distribution function; K-S, Kolmogorov–Smirnov; KNN, K nearest neighbours.

To confirm reproducibility, we repeated our complete analysis on an independent cohort of four rats, and observed categorical selectivity in OFC neurons with only minor differences (639 neurons; Extended

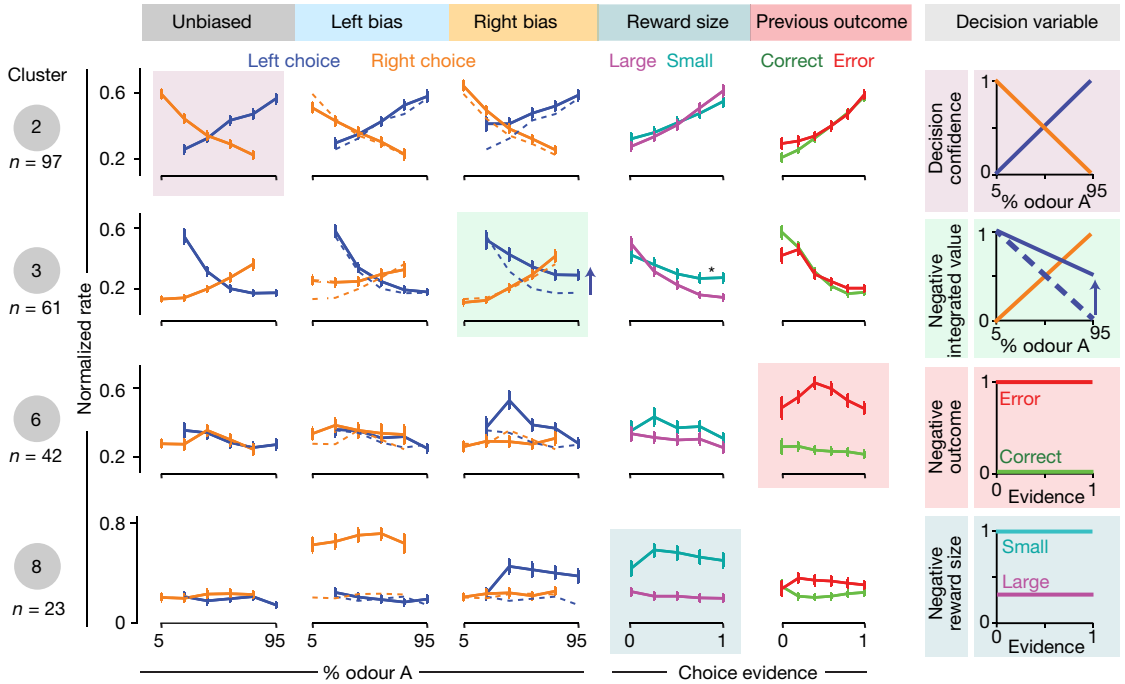


Fig. 3 | OFC neuron clusters represent putative decision variables. Average response profiles for 4 out of 9 clusters (see Extended Data Fig. 7e for others). Each row represents a cluster, with average normalized firing rate plotted across behavioural conditions. The first three conditions represent combinations of stimulus and choice in different reward blocks (unbiased, left, right bias). Reward size column represents tuning as a function of choice evidence and reward size blocks (left and right bias). Previous outcome column represents tuning as a function of choice evidence and previous trial outcome (correct rewarded, error not rewarded). Tuning profiles suggest specific decision variables represented. For each cluster, we highlight the expected shape of response profile that indicates which putative decision variable is represented, including potential reward-size biases that differ from the non-biased condition (dashed line). Error bars show s.e.m. *$P < 0.01$, *t*-test.

Data Fig. 4m, Extended Data Fig. 6 and Supplementary Note 2 'Cohort analysis').

Decision variable coding

Although clusters were identified without model input, the average response profile of each cluster resembles a putative decision variable in the behavioural model (Fig. 3), including decision confidence (confidence$^{(+)}$, cluster 2), integrated value (value$^{(-)}$, cluster 3), previous trial outcomes (previous outcome$^{(-)}$, cluster 6), reward size (cluster 8), and others (Extended Data Fig. 7).

We confirmed this correspondence by analysing three clusters—confidence$^{(+)}$, confidence$^{(-)}$ and integrated value$^{(-)}$—for which we had quantitative predictions. Confidence$^{(+)}$ neuron (cluster 2) responses matched key expectations of statistical decision confidence[8,21,24] (Fig. 4a and Extended Data Fig. 8a). Neural activity increased with stimulus contrast for correct choices but decreased for errors (Fig. 4a, bottom left), and firing rates predicted choice accuracy regardless of stimulus identity (Fig. 4a, bottom right). The confidence$^{(-)}$ cluster showed the same characteristics, sign-reversed (Extended Data Fig. 8b). Integrated value$^{(-)}$ neurons (cluster 3, Fig. 4b) also matched predictions, representing confidence and reward size in population averages (Fig. 4b, bottom left and Extended Data Fig. 8c) and correlates strongly with model estimates in individual neurons (Extended Data Fig. 8c (panels ii and vi)). Critically, firing rate tracked behavioural accuracy below 50%, reflecting outcome probability for decisions made under reward bias (Fig. 4b, bottom right).

We next considered how well the OFC population reflects decision variables without specifying a particular model. We generated a canonical regression model of elementary task variables (for example, odour stimulus, choice side, expected reward) that combine to produce decision variables (for example, confidence, integrated value; Fig. 4c, Methods and Extended Data Fig. 8e). Individual OFC neurons and cluster-averaged responses were well fit by this elementary model (clusters, $P < 1 \times 10^{-64}$; neurons, $P < 1 \times 10^{-10}$; *t*-test, Fig. 4d), revealing that each neuron can be represented in a space spanned by task-relevant variables.

Nevertheless, responses in this linear model may still represent arbitrary mixtures of task variables, rather than identifiable decision variables. As OFC neurons tend to represent coherent variables like integrated value$^{(+)}$ (combining confidence$^{(+)}$ and reward size$^{(+)}$; Extended Data Fig. 8d) rather than incoherent combinations (for example, confidence$^{(+)}$ and reward size$^{(-)}$), we tested whether representations of canonical decision variables were enriched in OFC. To do so, we compared regressions of OFC responses using the canonical model against a library of models containing the same variables randomly mixed by basis set rotation (Fig. 4e; see Methods). Although both models can represent any combination of variables, only the canonical model represents model-based decision variables sparsely. As expected, LASSO regression of cluster-averaged response profiles using the canonical model had higher sparsity (smaller L1 penalty, $P < 0.00042$) without penalizing fit ($P > 0.05$; Fig. 4f–g and Extended Data Fig. 8f). We observed similar results for individual neurons (Fig. 4h and Extended Data Fig. 8g). Thus, OFC representations identified by model-free clustering correspond to the canonical elements of a decision-variable model.

We observed similarly robust clustering in other epochs of the behavioural task, albeit with fewer clusters arising from the smaller number of behavioural variables in the stimulus and feedback epochs (Extended Data Fig. 7). Additionally, we observed structured transitions in decision-variable coding across epochs: neurons within a cluster were likely to co-cluster also in other epochs (Fig. 5a, Extended Data

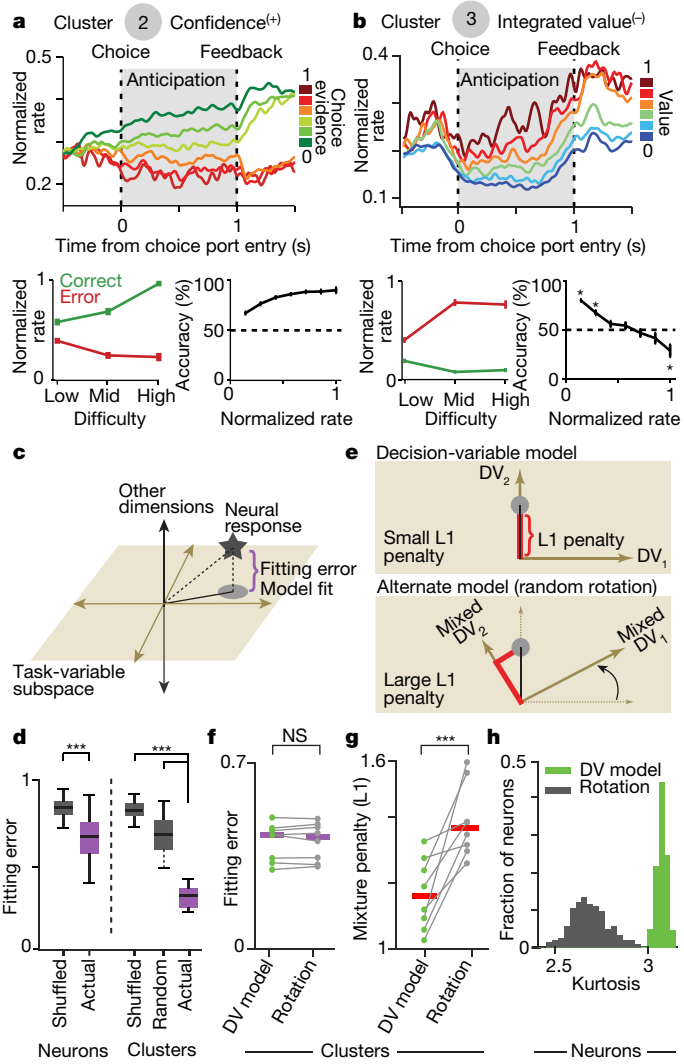


Fig. 4 | OFC clusters quantitatively encode the full decision-variable model. a, Response profile of neurons in cluster 2 correspond to decision variable confidence(+). Top panel shows peri-stimulus time histogram of normalized firing rate, grouped by evidence supporting choice. Bottom left panel shows firing rate as a function of stimulus difficulty and choice. Bottom right panel shows choice accuracy as a function of firing rate. **b**, Same convention as panel **a** for cluster 3, corresponding to decision variable integrated value(−). **c**, Cartoon of decision-variable model regression fit. **d**, Mean response profiles of single neurons and averaged clusters were fit: error was lower for actual neurons than for trial-shuffled control, and dramatically lower for fits of cluster averages than for single neurons, clusters of trial-shuffled neurons, or random clusters. Error distributions were compared under cross-validation by bootstrap test. Error bars show s.e.m., with horizontal links showing significant differences. ***$P < 0.001$. **e**, Least absolute shrinkage and selection operator (LASSO) regression analysis of individual clusters. We used two sets of decision-variable models to fit average cluster response profile: canonical model using elementary task variable such as stimulus and choice side, and alternate models in which these elementary task variables are randomly mixed (for example, choice side + reward block type − stimulus evidence). **f**, **g**, Fits of canonical decision-variable model (DV model) to clusters required smaller mixture penalty (L1, ***$P < 0.001$, sign rank test) than rotated variables (that is, mixtures of decision variables), with similar fitting error. NS, not significant; horizontal lines show tested comparisons across groups. **h**, Individual neurons are also better fit to canonical decision variables than their rotated versions: distribution of regression coefficients derived from fitting individual neuronal responses has higher kurtosis when fit with the canonical decision-variable model than with rotated models ($P < 0.001$; bootstrap test).

Fig. 7 and Supplementary Note 3). An analogous clustering procedure focused on temporal features of neural activity revealed robust structure in neuronal dynamics, with most neurons sparsely activated in a few task epochs and clustering into ~8 temporal profiles (Fig. 5b and Extended Data Fig. 9a–e).

Cell-type-specific coding

Seeking an anatomical substrate for these response profiles, we examined OFC neurons projecting to striatum (OFC–STR), a pathway important for reversal learning and choice value updating[25,26]. We used retrograde viruses to target ChR2 to OFC–STR neurons and identified 24 photo-tagged neurons based on signatures of direct light activation (Fig. 5c–e and Extended Data Fig. 10). Notably, their behavioural tuning and time course matched specific representations identified by model-free clustering (Fig. 5b). OFC–STR neurons significantly encoded trial outcome (23 out of 24, 5 positive, 18 negative), with activity sustained beyond the feedback period until the next trial. Negatively tuned OFC–STR neurons reduced their activity during the anticipation period (Fig. 5f, g), encoding negative integrated value. They increased their firing after negative outcomes during the feedback epoch, and sustained firing throughout the self-paced inter-trial interval, often for many seconds. This neuronal profile matched the dynamics of one temporal cluster, with the same transition from negative integrated value to sustained negative outcome coding (Fig. 5h, $n = 96$ neurons, cluster B in Fig. 5b). These results, including a similar pattern for positively tuned OFC–STR neurons (Extended Data Fig. 9f, g), suggest that the temporally structured and decision-variable-specific representations in OFC are supported, at least in part, by cell-type-specific circuit organization.

Representational logic of frontal cortex

We used a behavioural task that combines perceptual and value-guided decision making to demonstrate that OFC representations are highly structured: encoding a small set of categorical representations that correspond to coherent decision variables in specific task epochs. The functionally homogeneous encoding of decision variables by OFC–STR neurons further suggests that OFC response diversity is partly due to cell-type-specific organization. Sustained firing of OFC–STR neurons across trials encoded value, potentially a neural correlate of a temporal credit assignment mechanism. Alongside neuron-type-specific recordings in cortical and subcortical areas during behaviour[27–29], this result exposes an intimate connection between the functional and anatomical organization of cortex supporting computation.

The internal decision variables guiding behaviour (for example, value) can only be indirectly determined, like identifying an object from its shadow, by conditioning neural activity on measurable task variables (for example, stimulus, choice). Consequently, categorical representations of internal variables like decision confidence can appear mixed when examined as a function of external variables (Fig. 3). Our unbiased, model-free approach revealed that many hypothesized decision variables have a privileged but possibly task-dependent representation, constraining models of OFC function and clarifying its representational logic.

This structured representation of decision variables has strong analogies to the framework of sparse and overcomplete representations providing efficient sensory encoding[30]. OFC shares many similarities with these regions: sparse activation (Extended Data Fig. 3e, f), redundant encoding (Figs. 2 and 3) and representational sparsity (Fig. 4). We propose that this architecture is a fundamental feature of frontal cortex, with distinct cell types in OFC specializing in different computational functions to support adaptive behaviour.

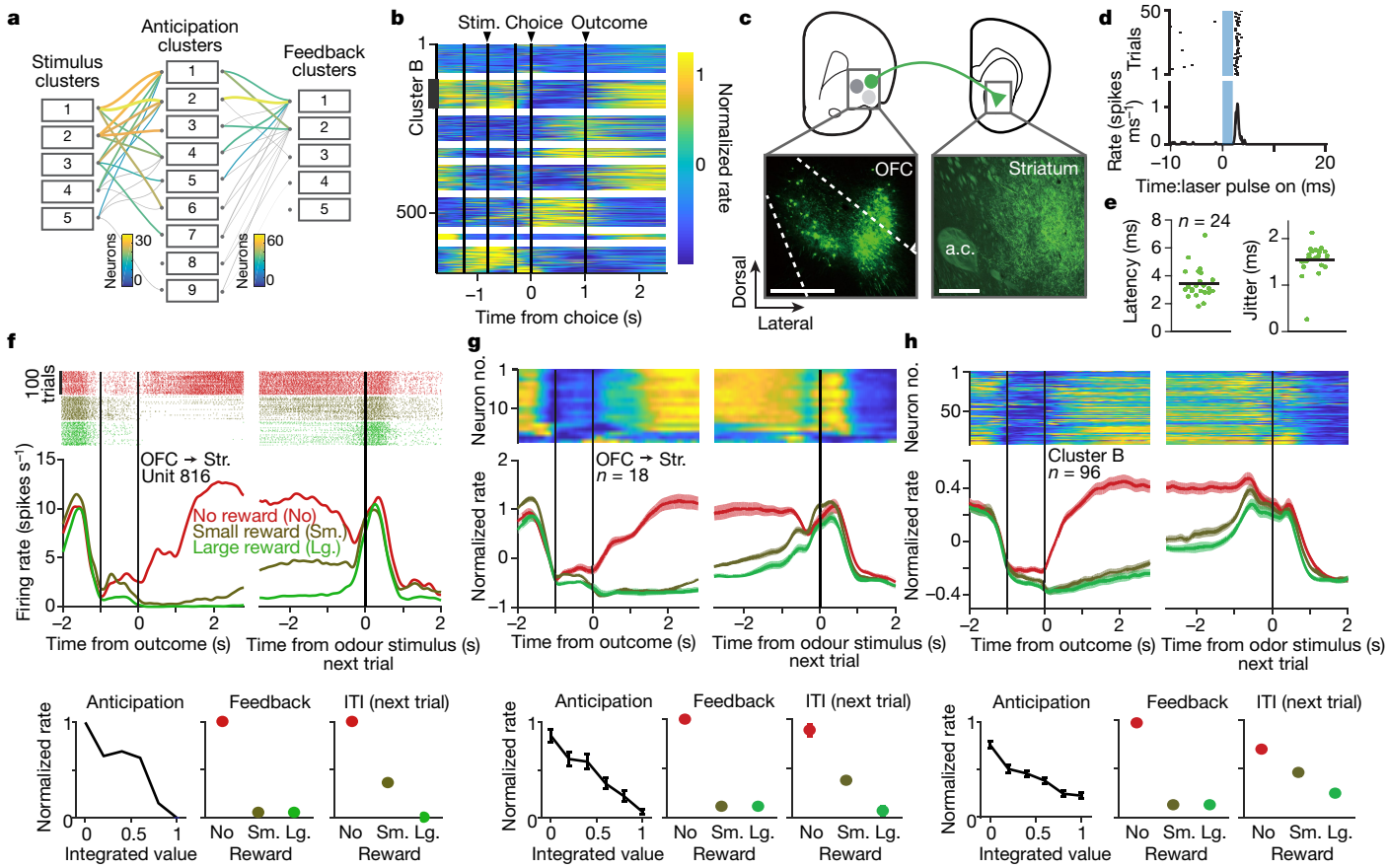


Fig. 5 | OFC-to-striatum projection neurons encode integrated value across task epochs. a, Cluster transitions across epochs: each box represents a neuron cluster in one epoch; lines between them represent neurons that belong to the same cluster in two epochs. Transitions in cluster identities across trial epochs are highly structured, with two neurons in a given cluster likely to remain in the same clusters over time. **b**, Dynamics of response time course for eight temporal clusters (unrewarded trials, see Extended Data Fig. 9d). **c**, Retrograde targeting of channelrhodopsin-2 (ChR2) to striatum-projecting neurons in lateral OFC for optogenetic identification (white dashed lines indicate approximate border for lateral OFC), scale 500 µm. **d**, Laser pulse (blue band) aligned responses of an optogenetically identified neuron. **e**, Latency and timing jitter of spikes from laser pulse onset for identified OFC–STR neurons was short and low (horizontal bars show averages). **f**, Spike raster and average activity in single striatum-projecting neuron aligned to outcome (reward or error sound, left) or to stimulus delivery in the next trial (right). Lower panels show analysis of firing rates in three different epochs: anticipation, for which firing rate varies with integrated value (left); feedback, for which firing rate reflects trial outcome with negative valence (middle); inter-trial interval (ITI), for which firing rates reflect graded outcomes with negative valence (right; NR, no reward; SR, small reward; LR, large reward). **g**, Average activity in error trials for optogenetically identified OFC–STR projection neurons (negative outcome selective, 18 of 24, area under the curve (AUC) < 0 with $P < 0.01$ permutation test) and their average peri-stimulus time histogram (PSTH) grouped by error, correct small and large reward trials. **h**, Average dynamics and tuning of neurons in temporal cluster B (panel **b**) match those of optogenetically identified striatum-projecting neurons in **f**, **g**. To avoid double counting, identified OFC–STR projection neurons were excluded from cluster B. Error bars represent mean ± s.e.m. in **g** and **h**.

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/s41586-019-1816-9.

1. Hubel, D. H. & Wiesel, T. N. Receptive fields of single neurones in the cat's striate cortex. *J. Physiol.* **148**, 574–591 (1959).
2. Bruce, C., Desimone, R. & Gross, C. G. Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque. *J. Neurophysiol.* **46**, 369–384 (1981).
3. O'Keefe, J. & Dostrovsky, J. The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat. *Brain Res.* **34**, 171–175 (1971).
4. Abe, H. & Lee, D. Distributed coding of actual and hypothetical outcomes in the orbital and dorsolateral prefrontal cortex. *Neuron* **70**, 731–741 (2011).
5. Feierstein, C. E., Quirk, M. C., Uchida, N., Sosulski, D. L. & Mainen, Z. F. Representation of spatial goals in rat orbitofrontal cortex. *Neuron* **51**, 495–507 (2006).
6. Roesch, M. R., Taylor, A. R. & Schoenbaum, G. Encoding of time-discounted rewards in orbitofrontal cortex is independent of value representation. *Neuron* **51**, 509–520 (2006).
7. Kennerley, S. W. & Wallis, J. D. Evaluating choices by single neurons in the frontal lobe: outcome value encoded across multiple decision variables. *Eur. J. Neurosci.* **29**, 2061–2073 (2009).
8. Kepecs, A., Uchida, N., Zariwala, H. A. & Mainen, Z. F. Neural correlates, computation and behavioural impact of decision confidence. *Nature* **455**, 227–231 (2008).
9. Machens, C. K., Romo, R. & Brody, C. D. Functional, but not anatomical, separation of "what" and "when" in prefrontal cortex. *J. Neurosci.* **30**, 350–360 (2010).
10. McGinty, V. B., Rangel, A. & Newsome, W. T. Orbitofrontal cortex value signals depend on fixation location during free viewing. *Neuron* **90**, 1299–1311 (2016).
11. Morrison, S. E. & Salzman, C. D. The convergence of information about rewarding and aversive stimuli in single neurons. *J. Neurosci.* **29**, 11471–11483 (2009).
12. Padoa-Schioppa, C. Neuronal origins of choice variability in economic decisions. *Neuron* **80**, 1322–1336 (2013).
13. Padoa-Schioppa, C. & Assad, J. A. Neurons in the orbitofrontal cortex encode economic value. *Nature* **441**, 223–226 (2006).
14. Rigotti, M. et al. The importance of mixed selectivity in complex cognitive tasks. *Nature* **497**, 585–590 (2013).
15. Mante, V., Sussillo, D., Shenoy, K. V. & Newsome, W. T. Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature* **503**, 78–84 (2013).
16. Stalnaker, T. A., Cooch, N. K. & Schoenbaum, G. What the orbitofrontal cortex does not do. *Nat. Neurosci.* **18**, 620–627 (2015).
17. Steiner, A. P. & Redish, A. D. Behavioral and neurophysiological correlates of regret in rat decision-making on a neuroeconomic task. *Nat. Neurosci.* **17**, 995–1002 (2014).

18. Sul, J. H., Kim, H., Huh, N., Lee, D. & Jung, M. W. Distinct roles of rodent orbitofrontal and medial prefrontal cortex in decision making. *Neuron* **66**, 449–460 (2010).
19. Xie, J. & Padoa-Schioppa, C. Neuronal remapping and circuit persistence in economic decisions. *Nat. Neurosci*. **19**, 855–861 (2016).
20. Zariwala, H. A., Kepecs, A., Uchida, N., Hirokawa, J. & Mainen, Z. F. The limits of deliberation in a perceptual decision task. *Neuron* **78**, 339–351 (2013).
21. Hangya, B., Sanders, J. I. & Kepecs, A. A Mathematical framework for statistical decision confidence. *Neural Comput*. **28**, 1840–1858 (2016).
22. Raposo, D., Kaufman, M. T. & Churchland, A. K. A category-free neural population supports evolving demands during decision-making. *Nat. Neurosci*. **17**, 1784–1792 (2014).
23. Cuesta-Albertos, J. A., Cuevas, A. & Fraiman, R. On projection-based tests for directional and compositional data. *Stat. Comput*. **19**, 367–380 (2009).
24. Sanders, J. I., Hangya, B. & Kepecs, A. Signatures of a statistical computation in the human sense of confidence. *Neuron* **90**, 499–506 (2016).
25. Morris, L. S. et al. Fronto-striatal organization: Defining functional and microstructural substrates of behavioural flexibility. *Cortex* **74**, 118–133 (2016).
26. Burguière, E., Monteiro, P., Feng, G. & Graybiel, A. M. Optogenetic stimulation of lateral orbitofronto-striatal pathway suppresses compulsive behaviors. *Science* **340**, 1243–1246 (2013).
27. Economo, M. N. et al. Distinct descending motor cortex pathways and their roles in movement. *Nature* **563**, 79–84 (2018).
28. Kvitsiani, D. et al. Distinct behavioural and network correlates of two interneuron types in prefrontal cortex. *Nature* **498**, 363–366 (2013).
29. Namboodiri, V. M. et al. Single-cell activity tracking reveals that orbitofrontal neurons acquire and maintain a long-term memory to guide behavioral adaptation. *Nature Neurosci*. **22**, 1110–1121 (2019).
30. Olshausen, B. A. & Field, D. J. Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature* **381**, 607–609 (1996).

Methods

Rats

Nine male Long Evans rats (~300 g) were used for the study. Seven rats were purchased from Taconic Biosciences (NY, USA) and two rats were obtained from Shimizu Laboratory Supplies (Kyoto, Japan). The rats were pair-housed and maintained on a reverse 12 h light/dark cycle and tested during their dark period. Food was available ad libitum, and rats had scheduled access to water with daily body weight monitoring to ensure that body mass remained above 85% of initial weight. Rats received water during behavioural sessions, and ad libitum in the following 30 min in their home cage. All procedures involving animals were carried out in accordance with National Institutes of Health standards and were approved by the Cold Spring Harbour Laboratory Institutional Animal Care and Use Committee and by the Animal Research Committee of Doshisha University.

Behavioural apparatus

The behavioural setup consisted of three ports equipped with infrared photodiodes; interruption of the infrared beam signalled port entry[8]. Odours were mixed with pure air to produce a 1:10 dilution at a flow rate of 1 l min^{-1} using a custom-built olfactometer (Island Motion; AALBORG). Delivery of odours and water reinforcement was controlled using Pulse Pal[31] and Bpod (J. Sanders and A.K.) with custom software written in MATLAB. Water rewards were delivered from gravity-fed reservoirs regulated by solenoid valves 1 s after the subject entered the choice port.

Reward biased odour discrimination task

Rats were trained and tested on a reward-biased two-alternative forced choice (2AFC) odour discrimination task as follows. Rats self-initiated each trial with a nose-poke into the central port to receive the odour stimulus. After a variable delay (uniform distribution of 0.2–0.6 s), a binary mixture of two pure odorants (*S*(+)-2-octanol and *R*(-)-2-octanol) was delivered at one of six concentration ratios ranging from 5% to 95% in pseudorandom order within a session. The specific odour ratios delivered were varied during training to produce graded accuracy levels from chance to near perfect performance.

After a variable odour sampling time, rats responded by withdrawing from the central port, thus terminating odour delivery, and moved to the left or right choice ports. Choices were rewarded according to the dominant component of the mixture, that is, at the left port for mixtures A/B > 50/50 and at the right port for A/B < 50/50. Reward amount was set to 0.025 ml in the control blocks and was reduced to 0.3–0.5 of the original amount for either right or left in left-biased reward block or right-biased reward block, respectively. The specific reward bias was determined to encourage graded responses to the reduced reward-size side and avoid excessive bias towards the larger reward side. Error choices resulted in water omission and were signalled by white noise (Fig. 1a, b; except for cohort 2 rats, see Supplementary Note 2 'Cohort analysis'). Two additional rats (cohort 3) were tested in a version of the task without reward size manipulation. In each cohort, randomization was not required as all rats were treated similarly.

In order to maximally utilize the psychometric response space, however, we intentionally modified the set of odour concentrations used in each session to generate a linearized set of choice probabilities (approximately [0, 0.2, 0.4, 0.6, 0.8, 1.0]) in the control block (that is, when reward sizes are equal; Fig. 1c). This linearized odour representation is reported using the axis legend '% odour A' in relevant figures.

Training

Rats were trained over the course of 4–6 weeks, with progressive introduction of each aspect of the task: imperative trials (not conditional on stimulus) followed by choice trials (conditional on the stimulus). Both trial difficulty and the delay before reward delivery were gradually increased until behaviour was stable; subsequently, reward bias was gradually introduced by expanding reward amount differences from baseline.

Surgery

Anaesthesia was induced with inhalation of 2.5% isoflurane and retained with intraperitoneal injections of ketamine (50 mg kg^{-1}) and medetomidine (0.4 mg kg^{-1}) at the onset of the surgery and supplemented as necessary based on the hind leg reflex. Body temperature was maintained using a heating pad (HoMedics). Rats were stereotactically implanted with custom-made microdrives in the left orbitofrontal cortex (targeted 1.5 mm above OFC (AP +3.7, ML ±3.2, DV +3.0; AP, anterior–posterior; ML, medial–lateral; DV, dorsal–ventral). Following surgery, rats were administered ketoprofen (Fort Dodge Animal Health) as an analgesic (5 mg kg^{-1}). Rats were monitored during their recovery from surgery at least 7 days before recordings began.

Electrophysiological recordings

Custom-built light-weight microdrives were constructed for deep brain recording and optogenetic stimulation with an optic fibre and 8 tetrodes. 3D printed microdrive bodies housed moveable shuttles that converted screw rotation into vertical motion advancing the shuttle and with it the attached tetrode and/or optical fibre. Individual tetrodes consisted of four twisted polyimide-coated nichrome wires (Precision Fine Tetrode Wire, Sandvik; single wire diameter 12 µm, gold plated to 0.25–0.5 MΩ impedance). Extracellular recordings were acquired on a DigitalLynx data acquisition system (Neuralynx) with a sampling rate of 31.25 kHz. Tetrodes were advanced daily (approximately 90 µm) after recording sessions so as to sample an independent population of cells across sessions and all the recorded neurons were analysed. For cohort 3 (2 rats, HA56 and HA58), the OpenEphys system was used for recording with a sampling rate of 30 kHz.

Histology

Rats were anaesthetized (sodium pentobarbital; overdose) and then transcardially perfused with saline and 4% paraformaldehyde. The brains were removed and 100 µm serial coronal sections were prepared with a vibratome. Recording sites were marked by coating electrodes with fluorescent dye (Vybrant DiI, Invitrogen) or by electrolytic lesions.

Viral injection

To target striatum-projecting OFC neurons we used a retrograde labelling approach. For rats CO91, S11 and HA56, we used adeno-associated virus (AAV) 2/9 serotype (5E12 pp ml^{-1} (physical particle ml^{-1}) UNC Vector Core Facility) carrying EF1a-DIO-ChR2-EYFP[32] or hSyn-DIO-$\{mCAR\}_{off}\{ChR2\}_{on}$(ref. [33]) and CAV2-Cre 4.1E12 pp ml^{-1} (ref. [34]) (Montpellier vectorology platform), injected on the same day into 4-week-old rats in OFC and striatum, respectively. For rat HA58 AAV retro[35], AAV-Syn-ChR2(H134R)-GFP (AddGene) was injected in the striatum. Rats were anaesthetized using 2.5% isoflurane and their eyes were protected with ophthalmic lubricant (Puralube Vet Ointment, Dechra Pharmaceuticals). Rats were placed in a stereotaxic apparatus (David Kopf Instruments, or Narishige) and the skull was levelled along both the antero-posterior and medio-lateral axis to allow the target coordinates: OFC (AP, +3.7 mm; ML, 3.2 mm; DV, 3.0 and 2.7 mm) and striatum (AP, +1.2 mm; ML, 2.3 mm; DV, 6.3 and 6.0 mm). At each dorso-ventral level, 150 nl virus was injected slowly for 4 min via a glass pipette pulled (P-97 Flaming/Brown Micropipette Puller, Sutter Instruments) from borosilicate capillaries (5 µl; tip diameter 20 µm). Injections were carried out by delivering brief pulses of pressure using Picospritzer II (Parker) or a micropump (UltraMicroPump III, WPI). After injections, the pipette was slowly pulled out after a 5 min waiting time. Rats were housed in their home cage and training commenced once they reached 14 weeks of age.

Optical stimulation

Optical stimulation was performed as previously described[28]. In brief, a multimode optical fibre (55 µm diameter NA = 0.22, Polymicro Technologies) was coupled via a modified LC–LC type connector to a multimode fibre (126 µm diameter NA = 0.27, CablesPlus), which collected light from a blue laser (473 nm; 50 mW; Ultralasers). Maximal power at the tip of the fibre ranged from 6 to 15 mW of total output. The light stimulation protocol (15–30 min) for optogenetic tagging was performed at the end of each recording session consisting of varying frequencies (10, 20 and 40 Hz with 1 ms or 3 ms light pulses) and intensities (0.1–10 mW) to enable reliable identification of directly light-activated neurons (Extended Data Fig. 10).

Behavioural data analysis

Each trial was defined by the stimulus presented, the animal's choice, and the reward associated with each choice. We calculated the odour sampling duration (OSD) as the difference between odour valve actuation and the odour port exit, with 100 ms subtracted to account for the delay from valve opening to odour reaching the nose (Fig. 1a and Extended Data Fig. 1e, l).

Daily behavioural sessions consisted of 821 ± 11 (mean ± s.e.m.) trials for cohort 1 (67 sessions), 893 ± 33 for cohort 2 (42 sessions) and 683 ± 19 for cohort 3 (photo-tagged rats, 47 sessions). For each session, we calculated behavioural accuracy as the fraction of correct choices, excluding trials in which OSD was less than 100 ms (0.47% of all trials) and trials in which no choice was made before the trial was terminated (0.82% of trials). Error bars are mean ± s.e.m. (n across rats) or mean ± s.d. (n across sessions).

The effect of odour contrast on accuracy or OSD was tested using one-way ANOVA with pairwise comparisons between different mixture contrast ratios at a significance level of $P < 0.0125$ (that is, adjusted for multiple comparisons).

Behavioural choice model

We developed a computational model for our behavioural task to describe the integration of sensory and reward information by the animal. In doing so, we aimed to predict choice behaviour for both biased and unbiased blocks, under four assumptions: (1) we assumed that animals rapidly learned and maintained a constant reward size estimate for each choice option within a given reward block. This is justified by the observation that animals adapted to new bias blocks within ~5 trials (Fig. 1b). (2) We assumed that animals engaged in our sensory discrimination task exploited the odour mixture to generate a choice (estimating the dominant mixture) but also to generate an internal estimate of outcome likelihood (correct or error) for each choice option (left or right)[8]. (3) We assumed that relative reward estimates, R, and outcome likelihood, $L = P(\text{correct} \mid \text{percept, choice})$, associated with each choice are integrated multiplicatively to generate an estimate of integrated value, $V = R \times L$. (4) Rats chose the option with the larger value. Under this assumption, the choice probability of each option is proportional to the relative value of the available choices.

Reward size: to formalize this variable, we defined the expected relative reward size, R_A, for each option in a particular block as follows:

$$R_A = \frac{W_A}{W_{control}}$$

where W_A is water reward amount for choice option A (that is, left or right) in reward-biased blocks and W_A is water reward amount in control block ($W_{control}$). Animals showed approximately linear dependence on relative reward size during the task.

Outcome likelihood: animals typically showed a sigmoid psychometric response profile across a range of odour cues. We estimated the subjective likelihood, L_L, of outcome for the left choice for a particular stimulus based on the psychometric function itself. For example, if an animal responded to odour pair A/B = (55%/45%) by choosing the left port on 60% of trials, then the animal's estimated probability of success (L_L) on that odour pair would be 0.6. Note that we used the estimated outcome likelihood to explain choice probability under reward size manipulation (or previous outcome effect).

Choice value: the value of each choice was estimated as a product of outcome likelihood and relative reward size:

$$V_X = L_X(R_X w)$$

where w is the weighting coefficient that determines reward size sensitivity and X denotes the choice side (that is, left or right).

Choice probability: Choice probability for choosing left was proportional to the relative reward value:

$$P_L = \frac{V_L}{V_L + V_R}$$

and was matched to the actual behavioural choice probability function (Fig. 1d) by fitting one free parameter, w. Note that w was set to 1 for the choices associated with normal reward size (so w does not cancel out in the function above) and therefore values for left and right choices are always asymmetric in the bias block. The coefficient of determination R^2 was calculated by squaring the correlation between the model-derived estimate of P_L and actual choice data.

Previous outcome bias: we extended this model to explain choice bias following correct or error trials in the control block. Since rats could differentially weigh the value of previously rewarded and unrewarded trials (beyond the sign difference), we independently estimated subjective values after previous correct and error choices. To estimate subjective value for previously correct and error trials, we used $V_X = L_X w_C$, for previously correct and $V_X = L_X w_E$, for previously error trials. For the unchosen side, we set $V_X = L_X$ as a reference. L_X represents the outcome likelihood and X denotes the choice side (that is, left or right). Then, we fit the choice probability as a function of relative integrated value as described above to find the w_C and w_E that minimized the squared error. See Extended Data Fig. 1h–l.

As noted above, the linearized odour stimulus representation is reported using the axis legend '% odour A'.

Behavioural strategies

In the reward-biased psychometric task used here, animals' performance is limited by two factors: (1) the difficulty of the psychometric task, which produces probabilistic rewards on many trials; and (2) the reward bias across trials, which provides inferior rewards in response to some choices. Within the limits of an animal's psychometric performance, the optimal reward-maximization strategy is to generate choices bases on an integrated estimate of choice value—estimated here as a multiplicative integration of outcome likelihood (confidence) and reward size associated with each choice.

To demonstrate that rats do indeed perform such optimal integration, we first modelled several inferior strategies (Extended Data Fig. 1f, g). Psychometric-only model: we modelled a strategy that relies on baseline psychometric performance to generate an estimate of the total reward available if an animal's choices only depended on outcome likelihood (ignoring reward size information). Reward-only model: we generated a model in which animals always choose the port associated with the larger reward (ignoring sensory evidence). Integrative model: for each session, expected reward amount was calculated as:

$$\text{Expected reward amount} = \sum_{n=1}^{T}(L_n R_n)$$

where L_n is reward likelihood for trial n, R_n is reward size in trial n. Reward likelihood was approximated based on the psychometric function across sessions, and T is total number of trials in a session. For psychometric-only and reward-only models, either R_n or L_n was fixed to the average values for appropriate to any given block.

Extended Data Fig. 1f, g shows that animals outperformed either the psychometric-only or the reward-only model. In addition, they showed hallmarks of reward-biases psychometric performance as described above, including the observations that the magnitude of reward bias is highest under maximal uncertainty (Fig. 1c), and that the choice probability is strongly correlated with the model (Fig. 1e). On the basis of this we infer that rats are performing a reward-maximizing integration of sensory evidence and reward contingencies, with their overall performance limited by the available sensory evidence within the difficult psychometric task.

Spike detection and sorting

Spikes were manually sorted into single-unit clusters (presumptive single neurons) off-line based on peak amplitude and waveform energy using the MClust software (A. D. Redish). Clusters were considered as single units only when the following criteria were met: (1) refractory period violations were less than 1% of all spikes; and (2) the isolation distance, which was estimated as the distance from the centre of identified cluster to the nearest cluster on the basis of the Mahalanobis distance, was greater than 20[36]. Units were sorted blind to any other criteria.

Optogenetic identification of tagged units

To identify photo-tagged neurons we used the Stimulus-Associated spike Latency Test (SALT; see supplementary note 1 of ref. [28] for a detailed description). SALT is a statistical test to determine whether optogenetic activation caused a significant change in the timing of spikes after stimulation onset. Specifically, the distribution of first spike latencies relative to the light pulse, assessed in a 10 ms window after light stimulation, was compared to epochs of the same duration in the stimulus-free baseline period. The choice of 10 ms window size provided sufficient statistical power without limiting the number of detected neurons.

Light-evoked spikes were defined based on the peak of PSTH during the 10 ms after light onset. Well-isolated single units (see above for the criterion) with significant correlation of average waveforms between spontaneous and the light-evoked spikes (Pearson's $r > 0.85$) as well as reliable spike generation (probability >0.4, $P < 0.01$, SALT) were identified as striatal projecting neurons. The activated neurons formed a distinct cluster in the space defined by spike latency, jitter and probability of the first spike (Extended Data Fig. 10g). Identified neurons showed strong correlation of the light-evoked and spontaneous spike waveforms (median correlation coefficient, 0.99; s.e., 0.0068; range, 0.89 to 1.0; Extended Data Fig. 10f). The median reliability of light-evoked responses was 0.61 (s.e., 0.036; range, 0.1 to 0.9) for low-frequency stimulation (10 Hz for $n = 22$ neurons; 20 Hz for $n = 2$ neurons).

Spike train analysis

We analysed single unit data from 146 behavioural sessions from 9 rats. Unless otherwise stated, spike trains were smoothed by convolution with a Gaussian kernel ($\sigma = 15$ ms) to obtain a spike density function (SDF) for the analysis of the temporal profile of neuronal activity. For most analyses, we focused our analysis on the 'reward anticipation period' while rats remained at one of the choice ports.

For cohort 1 (rats C051, C052 and C068), neuronal responses were analysed for 485 well-isolated neurons that had a non-zero firing rate during the anticipation window. For cohort 2 (rats C091, S11, V03 and V05), neuronal responses were analysed for 639 well-isolated neurons that had a non-zero firing rate during the anticipation window. For cohort 3 (rats HA56 and HA58), neuronal responses were analysed for 383 well-isolated neurons. This cohort was used only for OFC to striatum projection neurons and because the rats were tested without reward size manipulation these were not included in any of the clustering analyses.

For cohort 1 and 2, response profiles were generated across 48 conditions (Fig. 2b) with six conditions dropped due to frequent missing values (Extended Data Fig. 5i, j). The matrix of response profiles was subject to de-noising and imputation of sparse missing values using probabilistic principal components analysis (pPCA). To retain full population diversity in downstream analyses, coefficients were truncated to retain 90% of between-neuron variance, resulting in a set of coefficients for each neuron in a 21-dimensional response space (Extended Data Fig. 3b).

Although PCA is notoriously susceptible to over-interpretation, we note that approximately the first seven (accounting for ~60% of observed variance) resemble somewhat distorted mixtures of common decision variables (the first three are shown as Extended Data Fig. 3d). The interpretability of such eigenvectors as dominant neuronal tuning curves is limited by the assumption that the data are distributed as a multivariate Gaussian, and the number of eigenvectors resembling decision variables can be smaller than the actual number of unique decision variables if those decision variables are not linearly independent. For example, a 2D space for which the basis vectors represent confidence and reward size could contain representations of three distinct decision variables (confidence, reward size and choice value). We therefore sought to understand whether neurons in such a space show random mixed selectivity, or instead form categorical representations that align to putative decision variables using more detailed analyses.

Tests for random mixed selectivity

The assumption of random mixed selectivity is pervasive in the interpretation of cortical areas, but is rarely quantified directly. Here, we outline an improved set of tests that are appropriate for multi-dimensional neuronal data, drawing on the idea that random mixed selectivity is observed when neuronal responses can be represented as a multiple Gaussian (or similar) distribution.

Random mixed selectivity in this population was assessed using two novels methods: the 'elliptical projection angle index of response similarity' test (ePAIRS) and the 'elliptical random projection' test (eRP). In brief, these non-parametric tests assess whether the set of neuronal response profile are evenly distributed throughout the representational space, while accounting for differences in dimension size and tolerating changes in the magnitude of each response.

We formulate our approach as follows. We will first recapitulate a fairly standard approach to analysis of neuronal populations, and then discuss a deviation from some of its assumptions. We assume that we have recorded a set of n q-dimensional response profiles $X^{n,q}$. Each column of X corresponds to a single stimulus or task condition, and each row corresponds to a single neuron. This set of response profiles is usually z-scored, and de-noised to d dimensions using principal components analysis, retaining only a truncated set of d-dimensional coefficients for each neuron, along with a set of d q-dimensional orthogonal loading vectors.

For a set of z-scored response profiles as a row of X, PCA follows as $X'X = U\Sigma U'$ where columns of U encode the set of n loading vectors, and scores or coefficients are recovered as $S = XU$. The loading vectors are ranked by their contribution to the overall between-neuron variance of the neuronal population, represented by an associated eigenvalue on the diagonal of Σ. For the truncated decomposition using only the first d columns of U, S can be considered a de-noised rotation of X. As such, the first columns of B represent the axes of maximum population variance.

PCA has an intrinsic assumption of normality, and a stronger assumption that each dimension of X is independent. If this is met, loadings of S follow a multivariate Gaussian distribution, with the size of each dimension representing the between-neuron variance in the underlying variation.

Under the assumption of normality, the coefficients in S for each neuron are drawn independently for each neuron—if (and only if) this is true of the population as a whole, it can be said to show random mixed selectivity.

As an example, we present the population of synthetic neurons in Extended Data Fig. 4a–d. In each panel we show the coefficients of neurons in two dimensions; these neurons are also split into two subpopulations (blue and red). The population in the top panels (Extended Data Fig. 4a, b) have equal variance in each dimension, where those in the bottom panels (Extended Data Fig. 4c, d) have unequal variance. All of these populations may show simple mixed selectivity as long as the dimensions of variance (that is, loading vectors contained in columns of U) arise as combinations of relevant task variables (labelled here as Dimension 1 and Dimension 2).

However, this is not equivalent to random mixed selectivity, which is a statement of how the representation of each dimension is distributed across the population. The populations shown in the left panels (Extended Data Fig. 4a, c) show independent coefficients for each dimension, and thus show random mixed selectivity.

Neurons in the right panels (Extended Data Fig. 4b, d) have the same overall variance structure across the entire population: that is, the size of each dimension in the right panels matches the size of each dimension in the left panels. However, they do not show random mixed selectivity, because each neuron belongs to a subpopulation in which the coefficients in each dimension are no longer independent.

Many tests of random mixed selectivity begin from the intuition of symmetry: if the coefficients of a distribution are drawn independently and identically distributed (i.i.d), then the overall distribution should be rotation invariant, and show spherical symmetry[23]. This intuition is correct in the case where dimensions are of equal size (that is, spherical distributions; Extended Data Fig. 4a, b), but not where dimensions are of unequal size (that is, elliptical distributions; Extended Data Fig. 4c, d). Notably, essentially all neuronal datasets are elliptical for the simple reason that some representations are more prevalent than others.

Here we adapt and validate two tests for spherical symmetry that can be used for elliptical distributions. We define these tests based on two recently reported non-parametric tests for rotational invariance that are applicable to spherical distributions: the projection angle index of response similarity test (PAIRS[22]), and a modified random projection test (RP[23]).

The PAIRS test was first presented by ref. [22]. PAIRS approaches the problem as follows. (1) Given a d-dimensional dataset $X(n \times d)$, calculate for each row in X the cosine distance to its nearest neighbour. Calculate the empirical median of these nearest-neighbour distances as $\tilde{e}$. (2) Generate a set of m bootstrap distributions $Y(\mathbf{Y}_1 \ldots \mathbf{Y}_m)$, as spherical d-dimensional Gaussians. For each $\mathbf{Y}^*$, calculate the distribution of nearest-neighbour angles, and pool these measurements to generate a bootstrap distribution $\mathbf{B}(b_1 \ldots b_m)$. (3) Calculate the likelihood that the dataset $\mathbf{X}$ shows significant clustering as the empirical expectation $P = \mathrm{E}(\tilde{e} < B)$.

The intuition here is that, for any plausible form of clustered data, median nearest neighbour distances are likely to be smaller than those for uniformly distributed data. This assumption does not hold under some pathological but non-random distributions (for example, points that are regularly spaced).

The RP test was first presented by ref. [23] as the RPk test, with our implementation based on that of B. Lau (https://github.com/brian-lau/highdim). The RP test is similar to PAIRS in that it derives from a comparison of angle distributions. Specifically, for a spherical d-dimensional dataset X, we (1) generate a set of k random vectors as rows of a matrix $Z(k \times d)$, drawn uniformly across d dimensions. By default, the dimensions are of equal size (but see below). (2) Calculate the distribution of projection angles for each row in Z onto each row in X. Each of these k distributions is notated as a vector $\mathbf{e}_k$. (3) Generate a set of m bootstrap distributions $Y(\mathbf{y}_1 \ldots \mathbf{y}_m)$, as spherical d-dimensional Gaussians. Calculate the distribution of projection angles for each point in each distribution Y^* onto each vector in Z. Each such distribution is notated as $B_{\mathrm{m,k}}$. (4) For each k, calculate the Kolmogorov–Smirnov statistic between the empirical distribution $\mathbf{e}_\mathrm{k}$ and each bootstrap sample $B^*_{,\mathrm{k}}$. The median Kolmogorov–Smirnov statistic is designated as the empirical observation $\tilde{e}$. (5) For each k, calculate Kolmogorov–Smirnov statistics between all $m(m-1)$ pairs $B_{m,k}$. Pool all such Kolmogorov–Smirnov statistics as the bootstrap distribution $\mathbf{b}$. (6) Compare the median empirical-versus-bootstrap observation $\tilde{e}$ to the bootstrap-versus-bootstrap distribution $\mathbf{b}$. As larger Kolmogorov–Smirnov statistics reflect larger deviations between distributions, we define the probability that X is spherically uniform as $P = E(\tilde{e} <= \mathbf{b})$.

We note that this formulation is distinct from the RPk test proposed by Cuesta-Albertos and colleagues[23], who suggested a direct comparison of Kolmogorov–Smirnov P values with Bonferroni correction. We find that method to have low specificity and to be unstable under minor noise levels, and suggest the Monte Carlo procedure used here as a more reliable alternative.

As validation of our concern about sphericity, we note that both PAIRS and RP will correctly categorize the spherical uniformity of Extended Data Fig. 4a, and correctly detect the deviation from spherical uniformity in Extended Data Fig. 4b, d. However, both tests will also report that the distribution in Extended Data Fig. 4c is not spherically uniform (Extended Data Fig. 4e, i). This is at odds with the intuition that the elliptical distribution in Extended Data Fig. 4c still shows random mixed selectivity.

As a compensation for this effect, we can simply alter both the RP test and PAIRS test to use elliptical bootstrap distributions that match the variance structure of the test data. This is done by drawing bootstrap distributions Y from elliptical normal distributions that match the distribution sizes λ_i observed in X.

$$Y = N(0, 1, D)$$

$$Y = \{\mathbf{y}_i \times \lambda_i\} \text{ for each dimension size } \lambda$$

The resulting set of vectors can be used to generate a bootstrap distribution, because although it is not uniform on the unit hypersphere, it does arise from a matching spherically uniform Gaussian distribution. We refer to the resulting tests as 'elliptical' tests: ePAIRS and eRP.

To validate our approach, we conducted simulations using a set of point distributions that matched the variance structure observed in our OFC data (Extended Data Fig. 3c), but were otherwise either distributed uniformly or contained synthetic clusters generated as von Mises distributions (Extended Data Fig. 4e, f) and eRP (Extended Data Fig. 4i, j). PAIRS and ePAIRS tests were conducted with 1,000 bootstrap samples, across 30 replicates. The RP and eRP tests were conducted with 20 bootstrap distributions and 500 samples in each bootstrap distribution.

As expected, we observed that both the spherical PAIRS and RP tests produced frequent false-positive results for data that were known to be drawn from a uniform elliptical distribution (Extended Data Fig. 4e, i). The improved eRP and ePAIRS tests, however, generally correctly distinguish between clustered and un-clustered data (Extended Data Fig. 4f, j). Generally, the eRP test appears to have better sensitivity in identifying non-uniformity in the form of von Mises distributions, but we have not attempted to generalize this result to other models of non-uniformity.

This approach gives rise to several empirical results. First, we observe that the spherical RP and PAIRS tests are subject to extremely high type 1 (false positive) errors when presented with data from an elliptical Gaussian distribution. Nevertheless, this result does not necessarily invalidate the approach of ref. [22], for which the non-corrected test is simply the most conservative case for the result they ultimately observed (that is, apparent uniformity and random mixed selectivity).

After correction, we observe that the elliptical random projection (eRP) and elliptical PAIRS (ePAIRS) tests correctly fail to reject the hypothesis of spherical uniformity for the uniform Gaussian distribution, but correctly detect the deviation for the clustered von Mises distributions. In general, the eRP test appears to have greater sensitivity for weakly clustered von Mises distributions.

An implicit variable in much of this analysis is the dimensionality of the observed dataset $\boldsymbol{d}$. Our approach throughout this paper has been to only minimally de-noise our dataset and retain 21 dimensions that account for 90% of population variance. However, analyses presented in Extended Data Fig. 4e, f, i, j account for variations in $\mathbf{d}$ and show that our main result is invariant across $\mathbf{d}$.

Spectral clustering

Spectral clustering was performed using standard techniques on a binary adjacency matrix[37,38], including Shi and Malik normalization of the Laplacian followed by k-means clustering of the small eigenvectors (Extended Data Fig. 3f).

The adjacency matrix was generated by identifying the k-nearest neighbours under a cosine or correlation distance metric. Hyperparameters, including number of nearest neighbours (k), the number of clusters (c), were selected by maximizing the adjusted Rand Index (ARI). For ARI analysis, 100 bootstrap samples were generated for each hyperparameter combination, with a sub-sampling fraction of 0.9. Similar results were observed for both cosine and correlation distance metrics; the results of a correlation metric are shown in Fig. 2e, f. Peak ARI in this dataset (ARI ~ 0.65) was significantly higher than expected from trial-shuffled datasets (ARI ~ 0; $P < 0.01$; t-test), or from datasets in which each dimension of the response profile was independently shuffled (ARI ~ 0; $P < 0.01$; t-test).

Regression analysis

For the decision-variable model, we first defined a set of five putative decision variables: (1) the choice that the rat just made (left or right); (2) stimulus uncertainty (that is, stimulus difficulty conditional on choice); (3) expected reward size for each trial; (4) outcome of the previous trial; and (5) overall block type (biased or unbiased). Each variable was parameterized as a z-scored vector corresponding to the conditions in which it was relevant, with irrelevant conditions masked by setting them to zero. This set of vectors—appropriately masked and z-scored—was used as the model or design matrix for the regressions that follow.

Using this design matrix, we fit the z-scored response profile of each neuron, or the z-scored average response profile arising from each cluster of neurons. Because some conditions may be missing for each neuron, these were imputed using probabilistic PCA (see 'Spike train analysis').

For the regression in Fig. 4c, d, these variables were used in a non-penalized linear regression. For the penalized regression in Fig. 4e–h, these variables were used in a LASSO regression using the glmnet package. We used the traditional LASSO representation

$$\min_x \frac{1}{2}||y-(Ax+x_0)|| + \lambda||x||_1$$

where y is the response profile of a neuron or cluster, A is the design matrix, and x is the set of recovered coefficients.

The regularization parameter λ was selected to minimize the error of tenfold cross-validation. The overall error of each fit was measured as root-mean square error (RMSE).

To generate each of the null models, we applied a random rotation of the design matrix, equivalent to rotating the axes of the LASSO cartoon depicted in Fig. 4e. In a non-penalized regression, this rotation results in no change in overall fitting error. In an L1-penalized regression, however, the L1 penalty will rise, but only if the output variable (that is, the neuronal response profile) does in fact represent a sparse combination of the elements of the design matrix (that is, the decision variables).

For each regression, we computed the reconstruction error and mixture penalty (L1 penalty, or sum of absolute value of all coefficients). We observed significant increase in the mixture penalty when the regression was performed with the non-canonical null models. The average values for RMSE and L1 penalty are shown in Extended Data Fig. 8g.

Note that in Fig. 4f, g we omitted the cluster corresponding to 'conditional updating' for this analysis, because all of the neurons in this cluster arose from a single rat and because the response profile of this cluster was a poor fit to any hypothesized set of decision variables.

We performed the same regression on the response profiles of individual neurons. For some conditions, we had relatively few trials to enter into the average response (<5); therefore, we limited our approach to a twofold cross-validation, repeated 20 times. We examined the RMSE and L1 penalty of fits using the canonical decision variable model as well as a set of 20 random rotations of that model. In total, we performed 400 regressions for each of 485 neurons, scanning across a set of 10,000 lambda values. We observed that the L1 penalty was significantly lower for regressions using the canonical model as opposed to rotated models ($P = 0.002$; sign rank test; Extended Data Fig. 8f) with no difference in overall error.

In addition, we repeated this analysis after removing neurons that were poorly fit by the regression (that is, lacking a local minimum in reconstruction error, or for which all regression coefficients were less than 1×10^{-3}) and observed similar results (retained 319 cells; $P = 0.01$, sign rank test).

To provide a complementary analysis that did not rely on cross-validation, we performed a similar regression in which the neuronal response profile of all neurons was fit via a multi-response Gaussian with no penalization. This approach minimizes the Frobenius norm of the full error matrix:

$$\min_x \frac{1}{2}\sum ||y-(Ax+x_0)||_F^2$$

which has the convenient property that it is invariant under rotation (that is, simultaneous rotation of the neuronal response matrix and the design matrix).

This analysis produces 2,425 coefficients total (arising from 485 neurons and 5 coefficients per neuron). We treated this as a single distribution, and considered that the kurtosis of this distribution would be highest when the decision variables contained in the design matrix were maximally aligned with individual neuronal responses (Fig. 4g). This intuition is analogous to that underpinning independent components analysis, which directly seeks the axes on which data deviate from a Gaussian distribution[39].

We therefore compared the kurtosis of the coefficients distribution under the canonical decision variable model to that observed for fits under a set of random rotations (Fig. 4h). As expected, we observed significantly higher kurtosis under the canonical model than expected from random rotations ($P = 0.000266$, sign rank test; $P < 0.002$, bootstrap test). Together with the results of the standard LASSO analysis, these results demonstrate that individual neuronal response profiles are sparsely loaded onto the decision variable model.

Epoch analysis

We repeated the same clustering procedure as described above for two additional task epochs (the stimulus epoch and the feedback epoch). For the stimulus epoch, we included all spikes from the start of odour delivery until the rat left the centre port. For the feedback epoch, we included all spikes in the 500 ms following the revelation of the outcome. See Supplementary Note 3 'Epoch analysis' for further details.

Temporal dynamics clustering

For temporal clustering, we first constructed a feature vector for each neuron by calculating the average firing rate in 10 ms consecutive, non-overlapping time bins smoothed with a Gaussian kernel (sigma = 50 ms) and aligned to several task events (that is, centre poke, stimulus delivery, side poke, outcome). To allow for averaging firing rates across trials despite variable task epochs, we divided each task epoch into a fixed number of bins (one bin on average corresponds to 10 ms) and calculated the firing rate for each bin (pre-poke epoch, 50 bins; pre-stimulus epoch, 45 bins; stimulus epoch, 50 bins; movement epoch, 30 bins; anticipation epoch, 100 bins; feedback epoch, 150 bins; see figure 10 in ref. [40]). We then averaged the firing rate for each bin across trials to construct a trial-averaged, time-normalized feature vector separately for all correct and all error trials. Next we generated the full feature matrix by aligning the *z*-scored firing rates of all recorded neurons (cohort 1 and Cohort 2 with at least 1 Hz firing rate and sufficiently long pre-stimulus period), concatenating correct and error trial averages to yield a 850-dimensional feature matrix with $n = 679$ neurons. We reduced the number of feature dimensions (that is, time points) using PCA retaining 90% of variance (14 PCs). This feature space was used for spectral clustering following the same procedure as above. Spectral clustering yielded robust clusters validated by bootstrap stability using the adjusted rand index and the most stable clustering was obtained for hyperparameters $n = 8$ clusters and $k = 13$ nearest neighbours.

No statistical methods were used to predetermine sample size.

Reporting summary

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

Data availability

The datasets generated during and/or analysed during the current study are available from the corresponding author upon request.

Code availability

Software for ePAIRS and eRP is available at https://github.com/KepecsLab/EllipticalClustering.

31. Sanders, J. I. & Kepecs, A. A low-cost programmable pulse generator for physiology and behavior. *Front. Neuroeng.* **7**, 43 (2014).
32. Gradinaru, V. et al. Molecular and cellular approaches for diversifying and extending optogenetics. *Cell* **141**, 154–165 (2010).
33. Li, S. J., Vaughan, A., Sturgill, J. F. & Kepecs, A. A viral receptor complementation strategy to overcome CAV-2 tropism for efficient retrograde targeting of neurons. *Neuron* **98**, 905–917.e5 (2018).
34. Soudais, C., Laplace-Builhe, C., Kissa, K. & Kremer, E. J. Preferential transduction of neurons by canine adenovirus vectors and their efficient retrograde transport in vivo. *FASEB J.* **15**, 2283–2285 (2001).
35. Tervo, D. G. R. et al. A designer AAV variant permits efficient retrograde access to projection neurons. *Neuron* **92**, 372–382 (2016).
36. Schmitzer-Torbert, N., Jackson, J., Henze, D., Harris, K. & Redish, A. D. Quantitative measures of cluster quality for use in extracellular recordings. *Neuroscience* **131**, 1–11 (2005).
37. Von Luxburg, U. A tutorial on spectral clustering. *Stat. Comput.* **17**, 395–416 (2007).
38. Shi, J. & Malik, J. Normalized cuts and image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **22**, 888–905 (2000).
39. Comon, P. Independent component analysis, A new concept? *Signal Process.* **36**, 287–314 (1994).
40. Kobak, D. et al. Demixed principal component analysis of neural population data. *eLife* **5**, e10989 (2016).

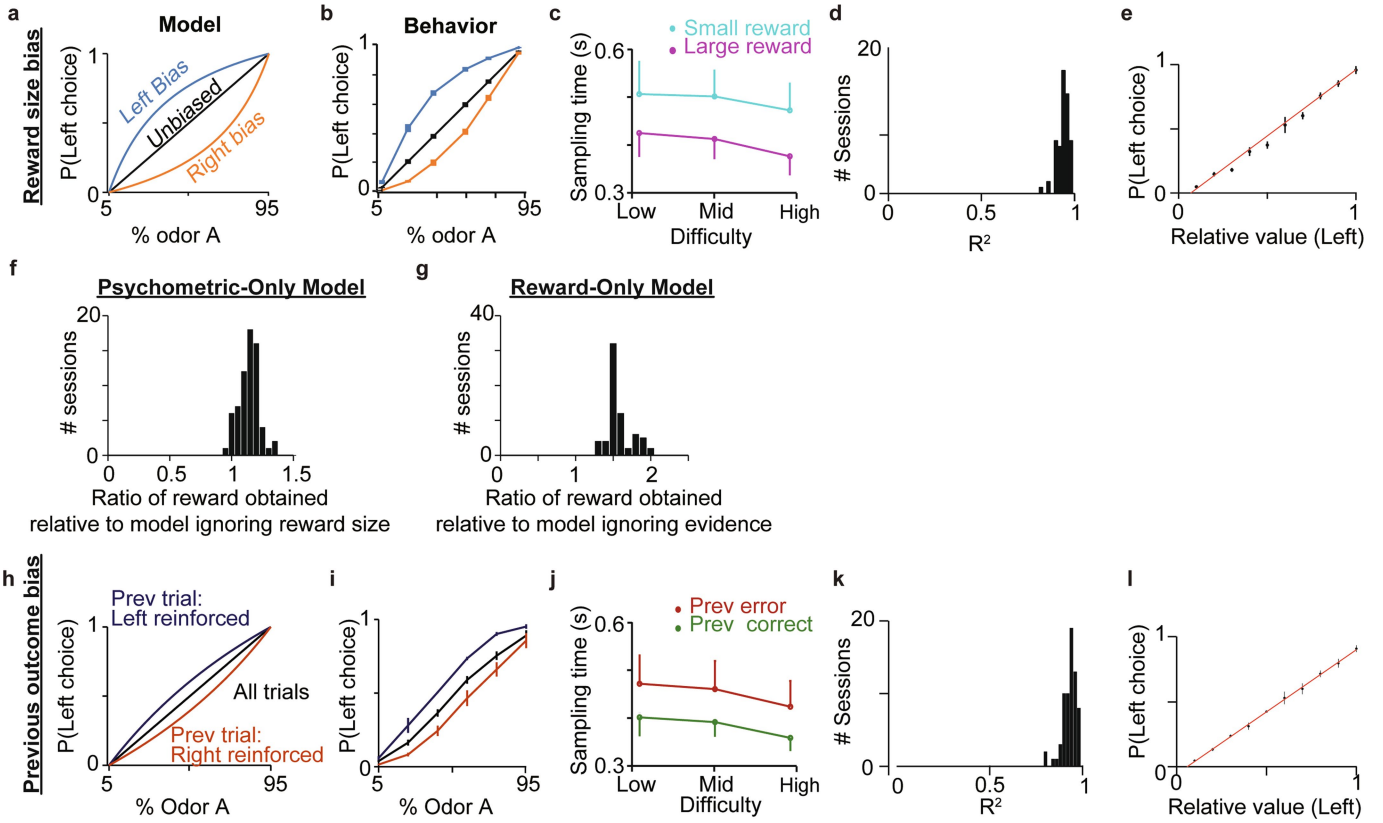
Acknowledgements We would like to thank past and present members of the Kepecs Laboratory and B. Mensh for many valuable discussions; J. Sanders and B. Hangya for help with experimental setup; and B. Mensh, T. Gouvea, M. Kaufman and A. Lak for comments on an earlier version of this paper. This study was funded by the grants from the Klingenstein, Alfred P. Sloan, Swartz, Whitehall Foundations and NIH grants R01DA038209 and R01MH097061 (A.K.), and KAKENHI 16K18380, 16H02061, 19H05028 (J.H.).

Author contributions J.H. designed and performed the experiments. J.H. and A.V. designed and performed the primary analyses, and T.O. and P.M. conducted additional analyses. A.K. designed the experiments, analyses and supervised the project. All authors contributed to writing the manuscript.

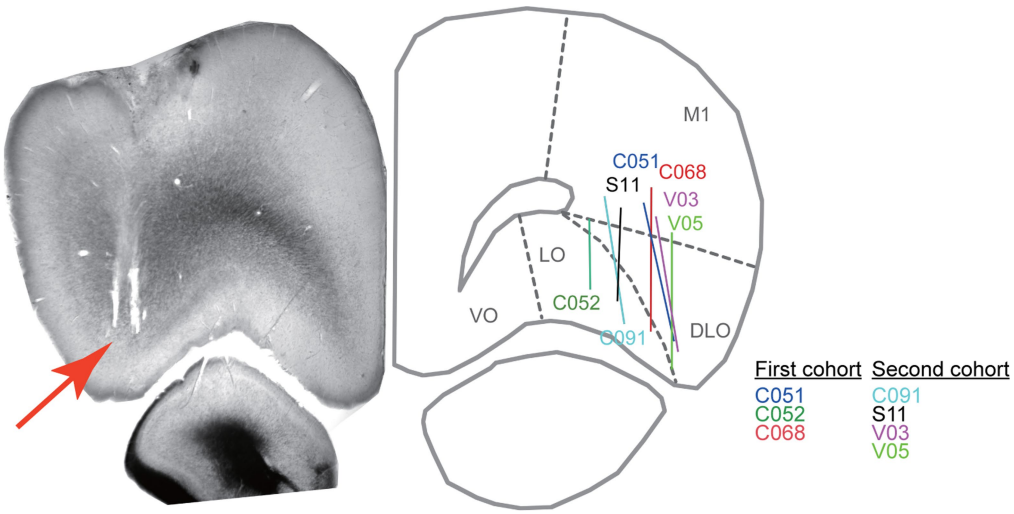
Competing interests The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-019-1816-9.

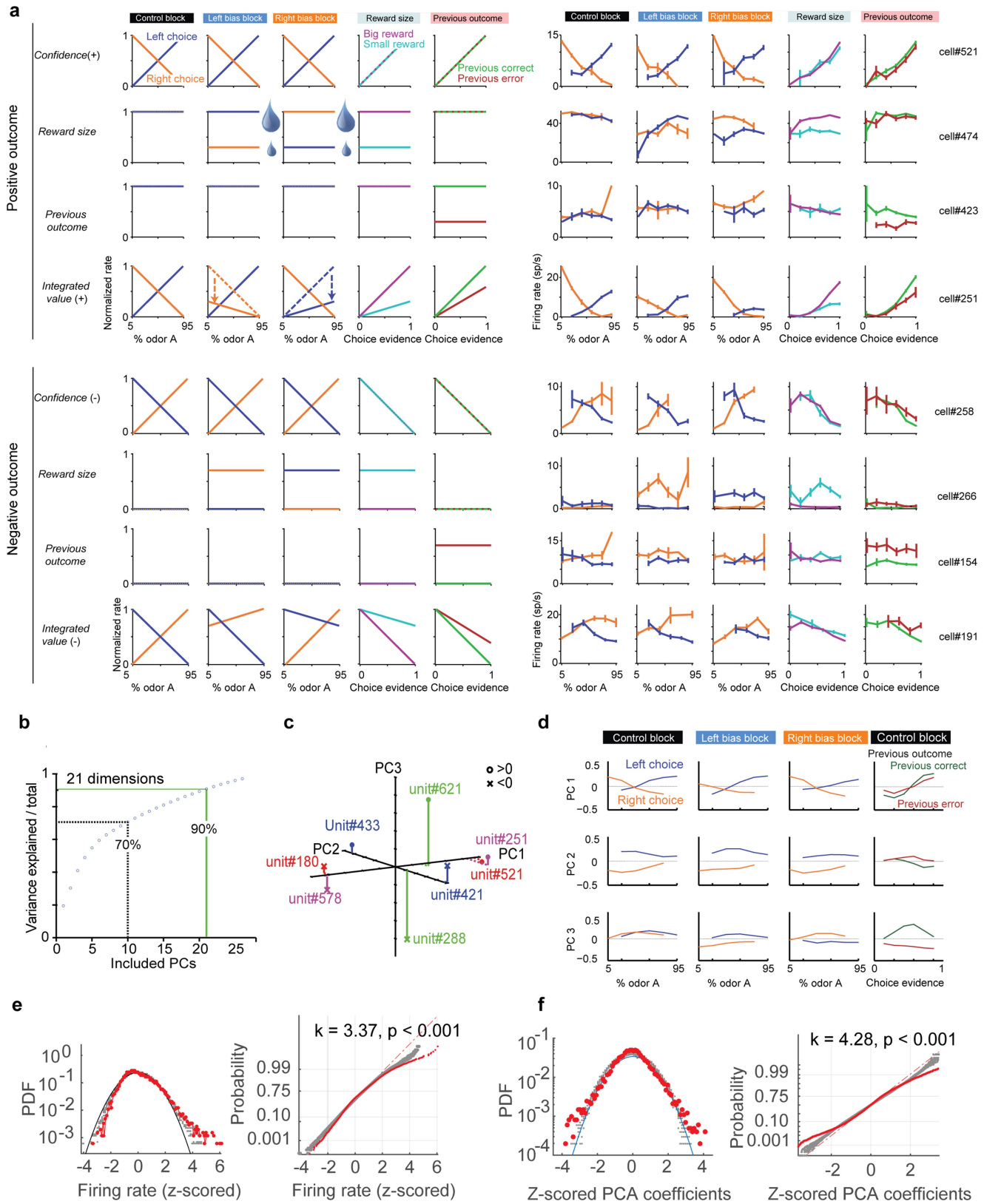
Correspondence and requests for materials should be addressed to A.K.



Extended Data Fig. 1 | Rat behaviour reflects an integration of evidence and reward-size. a–**e**, Average psychometric functions in unbiased, left and right bias blocks as a function of the odour percept ($n = 67$ sessions). The decision-variable model (**a**) and actual data from the same data (**b**) (replication of Fig. 1c) are shown. **c**, Odour sampling time was larger for small rewards than for large rewards. Errors are shown as mean ± s.e.m. ($n = 3$ rats). **d**, The model provides an excellent fit of choice patterns for each session. **e**, Choice driven by the relative value of left choice (replication of Fig. 1e). **f**, Histogram across sessions, comparing the ratio of actual reward obtained to a model relying on odour stimuli but ignoring reward size. **g**, Histogram across sessions, comparing the ratio of actual reward obtained to a model relying on reward size but ignoring odour stimuli. **h**–**l**, Same convention as **a**–**e** but reporting the bias arising from the outcome of the previous trial during the control block without reward bias ($n = 67$ sessions). **h**, **i**, The probability for left choice as a function of odour percept for all trials or separated by which choice (left/right) was rewarded in the previous trial. 'Left reinforced' indicates that rats are rewarded (correct) on the left side in the previous trial or not rewarded (error) on the right side, regardless of the stimulus conditions used in previous trials. The decision-variable model accurately predicts changes in choice probability (**h**, **i**) arising due to previous outcome. **j**, Odour sampling time was larger after a previously unrewarded choice than for a previously rewarded choice. Errors are shown as mean ± s.e.m. ($n = 3$ rats). The model provides an accurate fit across sessions (**k**) driven by the relative value of the left choice (**l**).

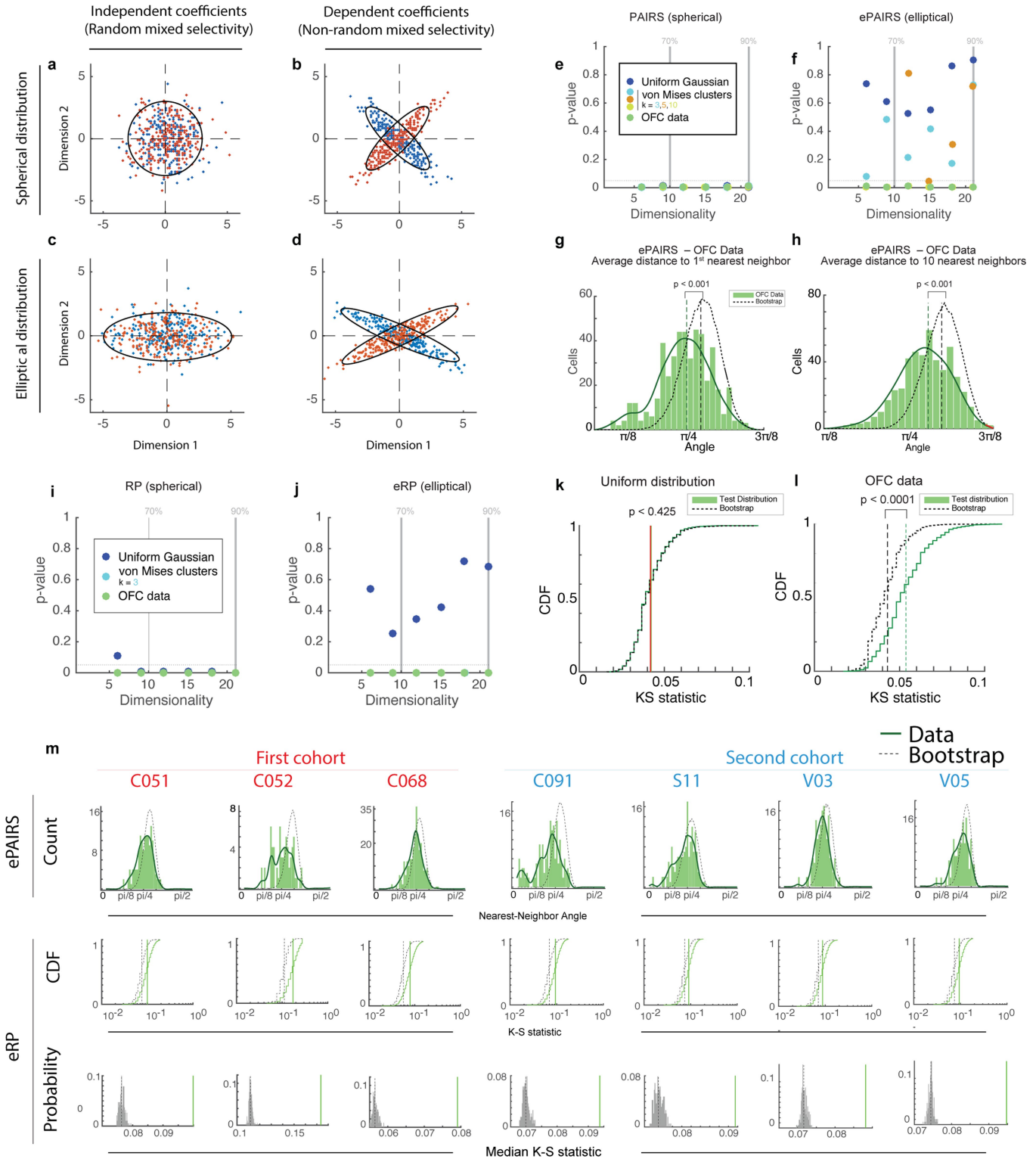


Extended Data Fig. 2 | Recording sites in lateral OFC and functional clustering across rats. Recording sites in lateral OFC for all seven rats from cohort 1 and cohort 2 are shown. Histological section shown in left (rat C068), in which the red arrow indicates the tip of the tetrode bundle.



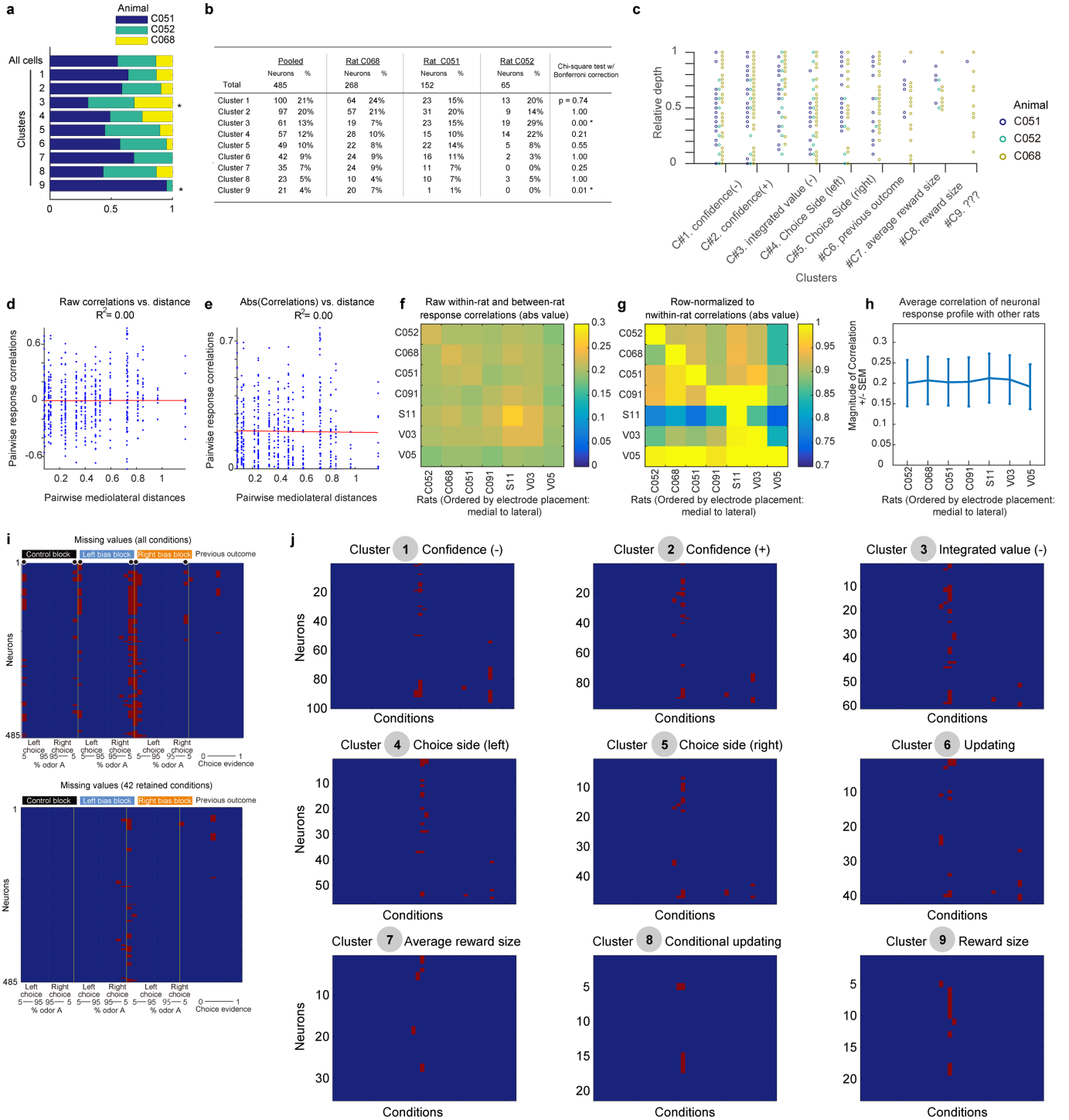
Extended Data Fig. 3 | See next page for caption.

Extended Data Fig. 3 | Profiles of individual neuron responses and population profiles in OFC. a, Example response profiles are shown for several individual neurons. In each case, a schematic tuning curve representing a plausible decision variable is shown in left panels, while a matching neuronal response profile is shown in right panels. **b**, PCA decomposition of the OFC dataset reveals high dimensionality, with 21 principal components required to retain 90% of response profile variability. **c**, Diversity of tuning vectors for several example neurons in the space of the first three principal components. **d**, The three dominant principal components arising from a probabilistic PCA decomposition of 485 OFC response profiles, which account for ~40% of population variance. **e**, Distribution of firing rates for all neurons across all conditions. Firing rates for observed data (red dots) show a right-tailed distribution, with strong activation of most neurons for only a small subset of conditions. This pattern of activation is significantly sparser than expected from a normal distribution (black line) or trial-shuffled data (grey dots). **f**, Coefficients arising from PCA analysis show a similar long-tailed distribution compared to a normal distribution (blue line) or trail-shuffled data (grey dots).

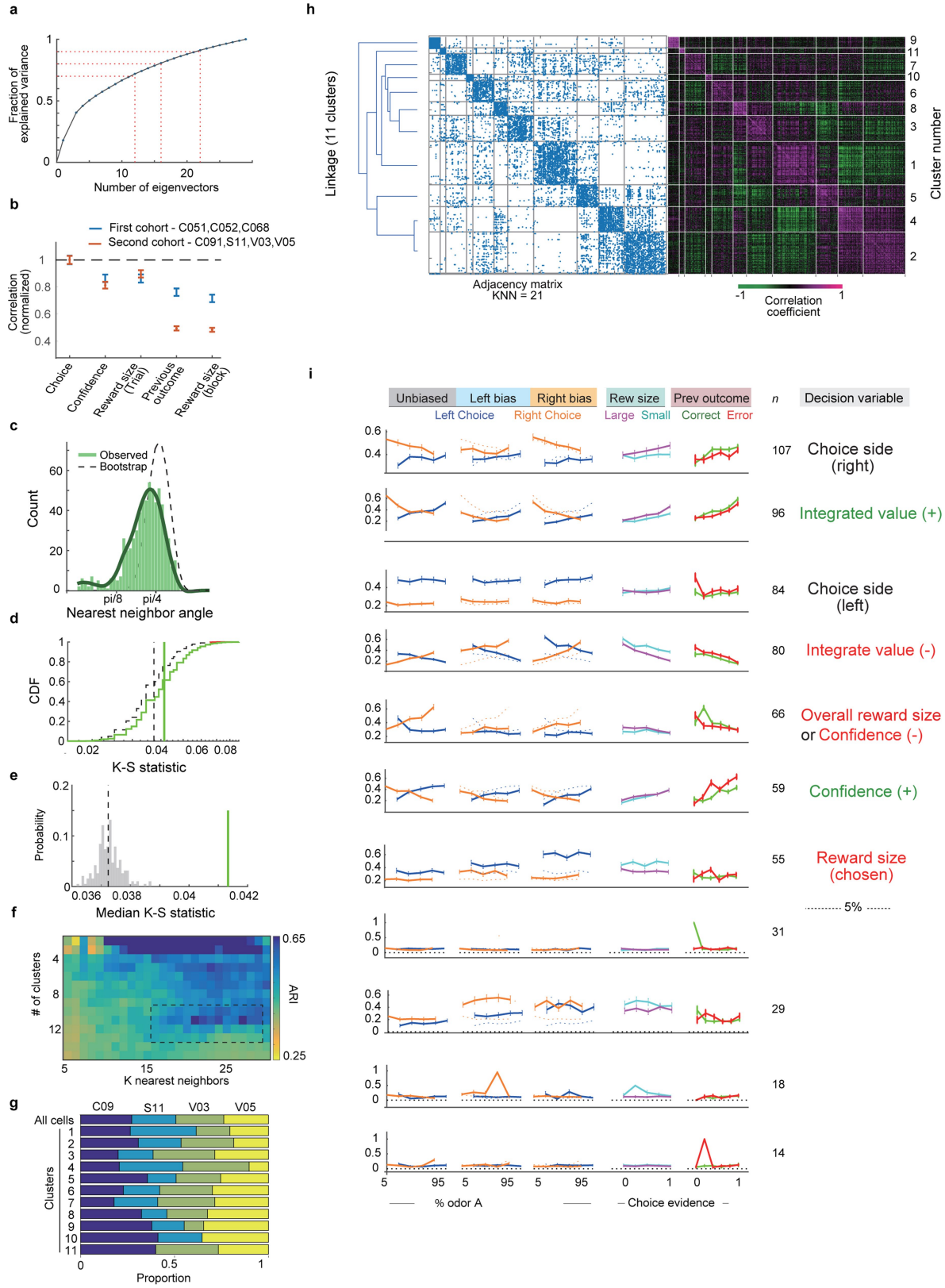


Extended Data Fig. 4 | See next page for caption.

Extended Data Fig. 4 | Testing for random mixed selectivity. a–**d**, Exemplar distributions of toy model neurons showing mixed selectivity, presented using the coefficients for two dominant eigenvectors (dimension 1 and dimension 2). Two neuronal subpopulations are shown in red and blue. Populations in the top panels (**a**, **b**) have equal variance in both dimensions, whereas populations in the bottom panels (**c**, **d**) do not. Populations in the left panels (**a**, **c**) can be said to show random mixed selectivity, whereas distributions in the right panels (**b**, **d**) do not. Only the distribution shown in panel **a** can be said to show spherical symmetry, necessitating the development of modified tests. **e**–**h**, Comparison of the PAIRS to the modified ePAIRS test, which accounts for elliptical distributions. **e**, **f**, Sensitivity analysis of PAIRS and ePAIRS, tested across several datasets whose variance structure matched the OFC data (green), including a spherically uniform Gaussian (blue) and collections of five randomly oriented von Mises distributions with varying κ; blue, orange, yellow). Datasets were truncated at a given dimensionality, and mean P values are reported across 30 replicates. Results show that spherical PAIRS generates false-positive results when tested on non-spherical but otherwise i.i.d Gaussian data. The modified ePAIRS test successfully identifies the non-uniformity of strongly clustered data ($\kappa = 10$) but not weaker clustering. The dimensionality required to reconstruct 70% and 90% of the variance in the full dataset is shown (grey lines). **g**, **h**, The ePAIRS measure, nearest-neighbour angles, is shown for OFC data (green) and bootstrap distributions (black). OFC data showed smaller angles than expected for both 1 and 10 nearest neighbours, suggesting strong clustering. **i**–**l**, Comparison of the random projection (RP) test to the elliptical random projection (eRP) test (eRP). **i**, **j**, We compared the RP and eRP tests on several datasets including an elliptical Gaussian distribution (dark blue) and collections of five randomly oriented von Mises distributions with $\kappa = 3$ (light blue), and observed OFC data (green). Other parameters matched panels **e**, **f**. Results show that spherical RP generates false-positive results when tested on non-spherical but otherwise i.i.d Gaussian data (dark blue). The modified eRP test successfully identifies the non-uniformity of von Mises clusters, as well as OFC data, while rejecting spherically uniform Gaussian data. **k**, **l**, Cumulative distribution function (CDF) of Kolmogorov–Smirnov statistics arising in the eRP test. Results are shown for test distributions that are spherically uniform Gaussians (**k**), as well as for OFC data (**l**). **m**, Analysis of individual rats from cohort 1 and cohort 2. All rats showed significant deviation from uniformity for both ePAIRS and eRP tests. Top, histogram of nearest-neighbour angles for observed data (green) and bootstrap samples with similar elliptical distribution (black). All animals showed significant differences, assessed using a rank sum test. Middle, cumulative distribution of Kolmogorov–Smirnov test statistics from eRP for bootstrap samples and observed data. Here each observation is the K-S statistic derived from comparing the distribution of projected angles onto a single random vector between observed data and a matching elliptical Gaussian distribution. This procedure is repeated for a set of k random vectors to generate the plotted distribution. This calculation is performed both for observed data (green) and for a simulated spherically uniform bootstrap distribution with matching samples size and ellipticity (black). Bottom, comparison of the observed median K-S statistic from observed data (green), to the distribution of medians observed across several realizations of a bootstrap distribution (grey). For all eRP estimates, higher K-S statistics denote greater deviation from uniform distribution, and all rats showed significant differences using a bootstrap test.

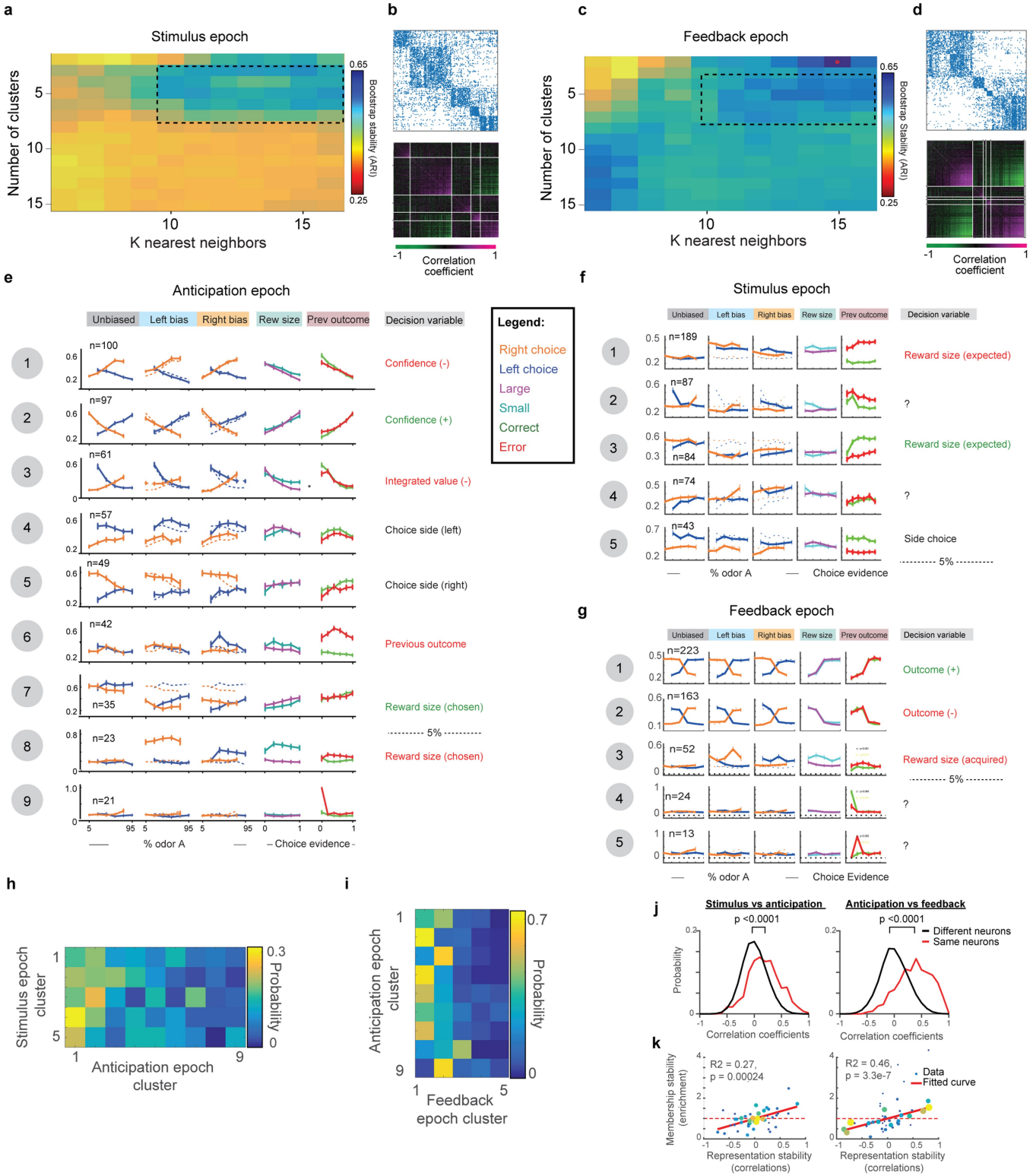


Extended Data Fig. 5 | Stability of cluster identification. a, Clusters are well distributed across animals from cohort 1. Cells in each cluster were generally drawn from all three animals, and rarely showed significant animal-specific bias. **b**, Table of the distribution of the neurons across rats from cohort 1 and across clusters identified through spectral clustering (see Fig. 2). **c**, Neurons in each cluster were plotted against the relative recording depths (normalized to 0–1 from recording starting point to end point shown in Extended Data Fig. 2) for each animal. **d**–**h**, Null relationships between spatial proximity and response correlation. **d**, For each pair of cells between all pairs of rats, we compared the pairwise mediolateral distances (*x* axis) to the pairwise correlations in response profiles (*y* axis). There is no significant relationship between mediolateral distance and magnitude of response correlation. **e**, This plot is identical to that in panel **d**, except that correlations are normalized for sign. **f**, Average magnitude of correlation between pairs of cells, analysed for all pairs of rats. Rats are ordered by electrode position, from medial to lateral. There is no obvious relationship between ML position order and average pairwise correlation of neuronal response profiles. **g**, Panel **g** is the same as panel **f** but normalized for each rat (that is, relative to within-rat response variability). **h**, Average magnitude of correlation between a given rat, and all other rats, with error bars representing s.e.m. This is equivalent to the row average of panel **f**. **i**, Missing values are shown for all 485 neurons in cohort 1. Top, missing values are common for some conditions, because animals rarely make errors against high-reliability cues, and are further biased against certain errors during bias blocks. Six such conditions are dropped due to excessive missing data (black dots). Bottom, after removing these conditions, 42 conditions remain for analysis, with sporadic missing values. **j**, Missing values are imputed during preprocessing using probabilistic PCA. If these missing values influence clustering, we expect to see a consistent pattern of missing values in certain clusters. Although there is variation in missing values across clusters, there is no obvious pattern of missing data.



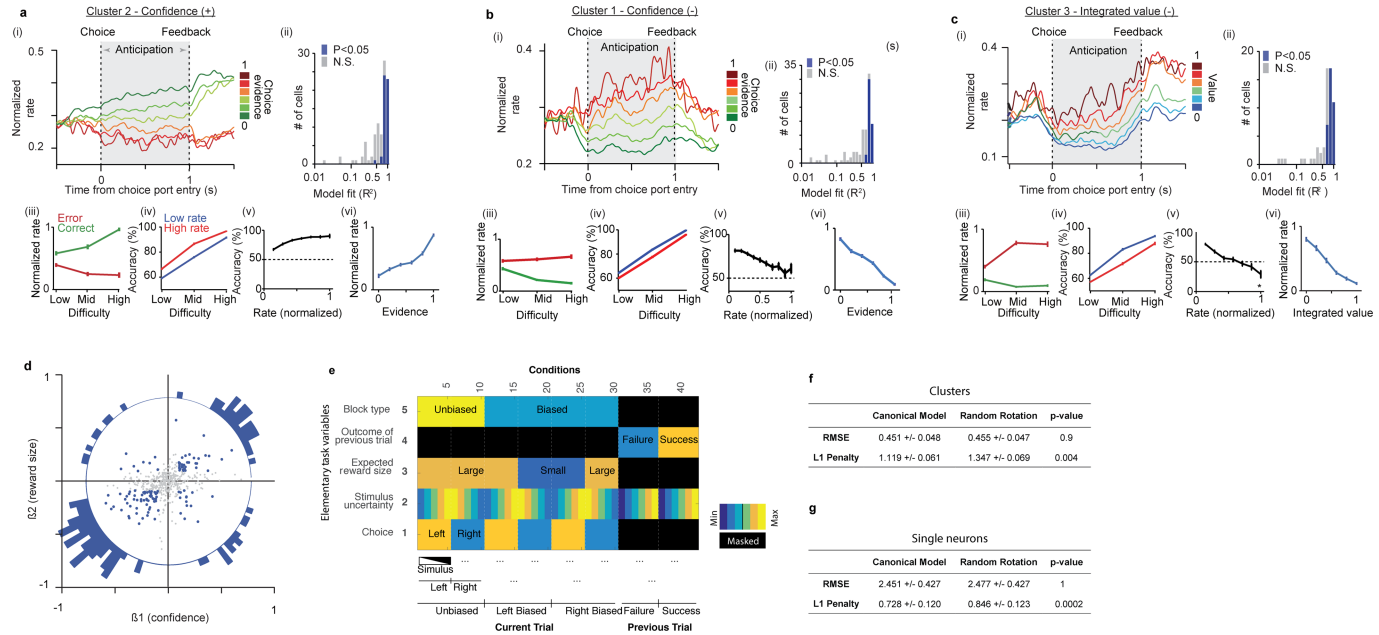
Extended Data Fig. 6 | See next page for caption.

Extended Data Fig. 6 | Response profiles of OFC neurons from cohort 2 replicate the results of cohort 1. a, Distribution of variance across the first 29 eigenvectors for cohort 2, see Extended Data Fig. 3b for corresponding panel for cohort 1. **b**, Average correlations of individual cell response profiles with a set of canonical response profiles corresponding to decision variables side choice, confidence, reward size (trial-by-trial), previous outcome, and reward size (block average). The sign of the correlation was discarded and normalized across cells by the strongest correlation (that is, side choice). Two representations (previous outcome and block-wise reward size) showed reduced representation in animals from cohort 2. **c**, ePAIRS test, showing the distribution of nearest-neighbour distances between observed data in cohort 2 (green) and a bootstrap distribution derived from simulated data with a matching elliptical Gaussian (black). $P < 0.001$, Rank sum test. **d**, Cumulative distribution function (CDF) for observed data (green) and a bootstrap distribution (black). Median values are shown with vertical lines. **e**, Distribution of median values for the K-S statistic across a set of bootstrap distributions (grey), compared to the median value for observed data (green). $P < 0.001$, bootstrap test. **f**, ARI for spectral clustering across k (k-nearest neighbours used to generate the adjacency matrix) and number of clusters, showing marked clustering around 11 clusters. **g**, Proportion of cells from each animal associated with each cluster. **h**, Left, dendrogram of inter-cluster distances. Middle, adjacency matrix, derived from $k = 21$ using a correlation distance. Right, between- and within-cluster correlations. **i**, Average response profiles for each cluster in cohort 2. The format of this figure matches Fig. 3 in the main text. Overall, we identified 11 clusters, of which the top 7 (each containing >5% of the cells in the dataset) correspond to separable representations of choice, confidence and value.



Extended Data Fig. 7 | See next page for caption.

Extended Data Fig. 7 | Analysis of OFC response structure across behavioural epochs (cohort 1). a, Clustering results and hyperparameter selection for responses in the stimulus epoch (most stable configuration: $c = 5$ clusters and $k = 16$ nearest neighbours). Each combination of parameters is evaluated for stability using the adjusted rand index (ARI) (see Methods and Fig. 2). **b**, Clustering results for the stimulus epoch. The relationship between the five clusters can be examined visually by observation of the nearest-neighbour graph (top) and the within-cluster and between-cluster correlation coefficient (bottom). **c**, Clustering results and hyperparameter selection for responses in the feedback epoch (most stable configuration: $c = 5$ clusters and $k = 16$ nearest neighbours). Each combination of parameters is evaluated for stability using the adjusted rand index (ARI) (see Methods and Fig. 2). **d**, Clustering results for the feedback epoch. The relationship between the five clusters can be examined visually by observation of the nearest-neighbour graph (top) and the within-cluster and between-cluster correlation coefficient (bottom). **e**–**g**, Full cluster response profiles for all three epochs. **e**, Average response profiles of each of the 9 identified response clusters in the anticipation epoch (compare to Fig. 2). For each cluster, the normalized firing rate is shown for all 42 behavioural conditions used to generate the clustering results (responses conditioned on stimulus and choice, unbiased, left bias, and right bias blocks; conditioned on outcome of the previous choice and the evidence supporting the current choice, previous outcome). In addition, normalized firing rates are shown conditioned on the size of the reward associated with the choice port (reward size). For each cluster, we also note the corresponding putative decision variable. **f**, Average response profiles of each of the five identified response clusters in the stimulus epoch. Conventions are the same as in **e**. Two of the clusters did not obviously map on a putative decision variable. **g**, Average response profiles of each of the five identified response clusters in the feedback epoch. Conventions are the same as in panel **e**. Two of the clusters did not map on a putative decision variable. **h**, **i**, Transition probabilities for neurons in a given cluster across subsequent epochs (compare to Fig. 5a). **h**, Transition probability for neurons belonging to a given cluster in the stimulus epoch to belong to a given cluster in the anticipation epoch (normalized per row). **i**, Transition probability between anticipation and feedback epochs (normalized per row). **j**, Neuron-based similarity measures across epochs. Neuronal response profiles are more similar across epochs for paired responses from the same neuron (red) compared to responses of two different neurons (black). Left, comparison of stimulus and anticipation epochs; right, comparison of anticipation and feedback epochs. Two-sample Kolmogorov–Smirnov test. **k**, Cluster-based similarity measures across epochs. Clusters derived from different epochs are more likely to share members if the average response profiles of each cluster are similar. Left, comparison of stimulus and anticipation epochs; right, comparison of anticipation and feedback epochs.



f

Clusters

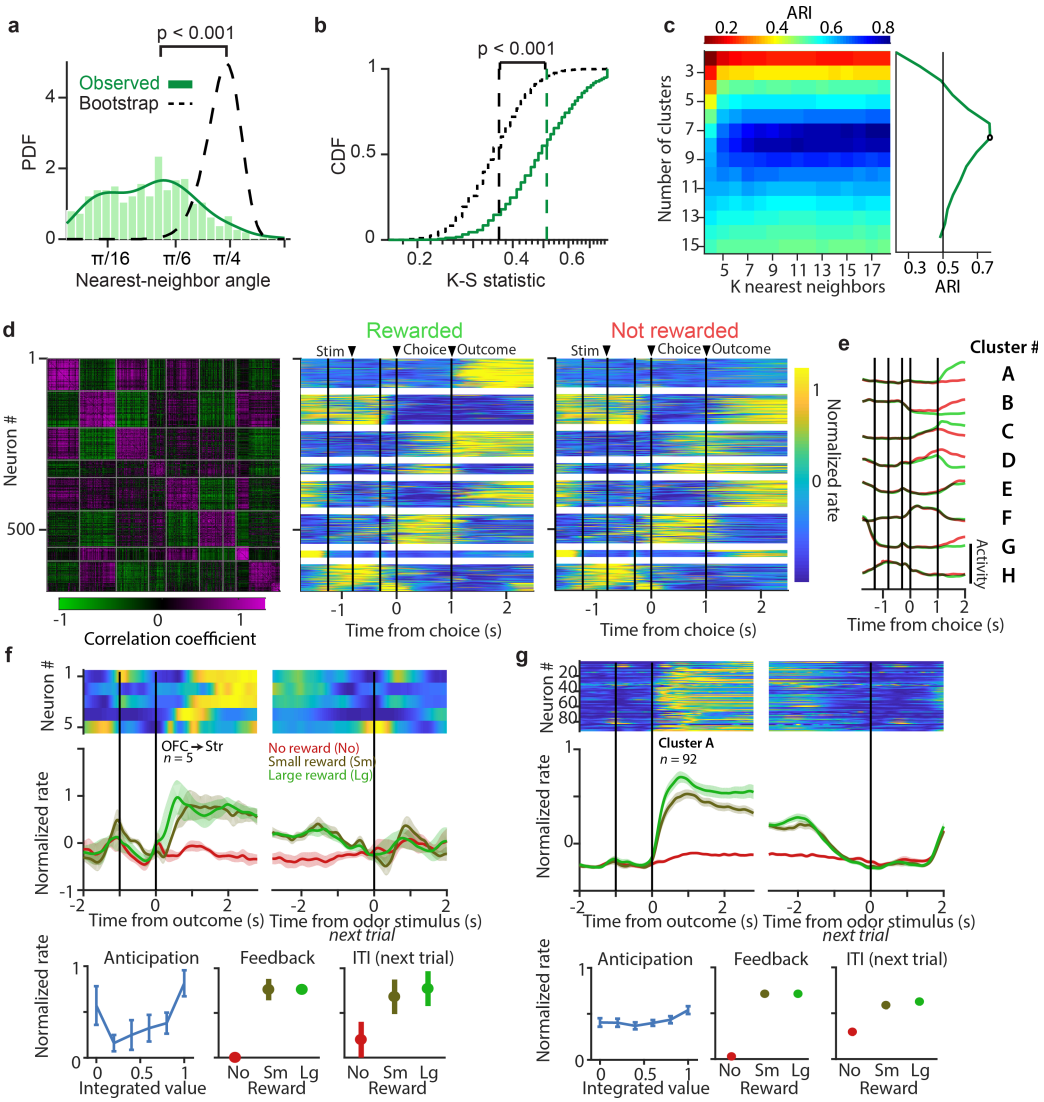
	Canonical Model	Random Rotation	p-value
RMSE	0.451 +/- 0.048	0.455 +/- 0.047	0.9
L1 Penalty	1.119 +/- 0.061	1.347 +/- 0.069	0.004

g

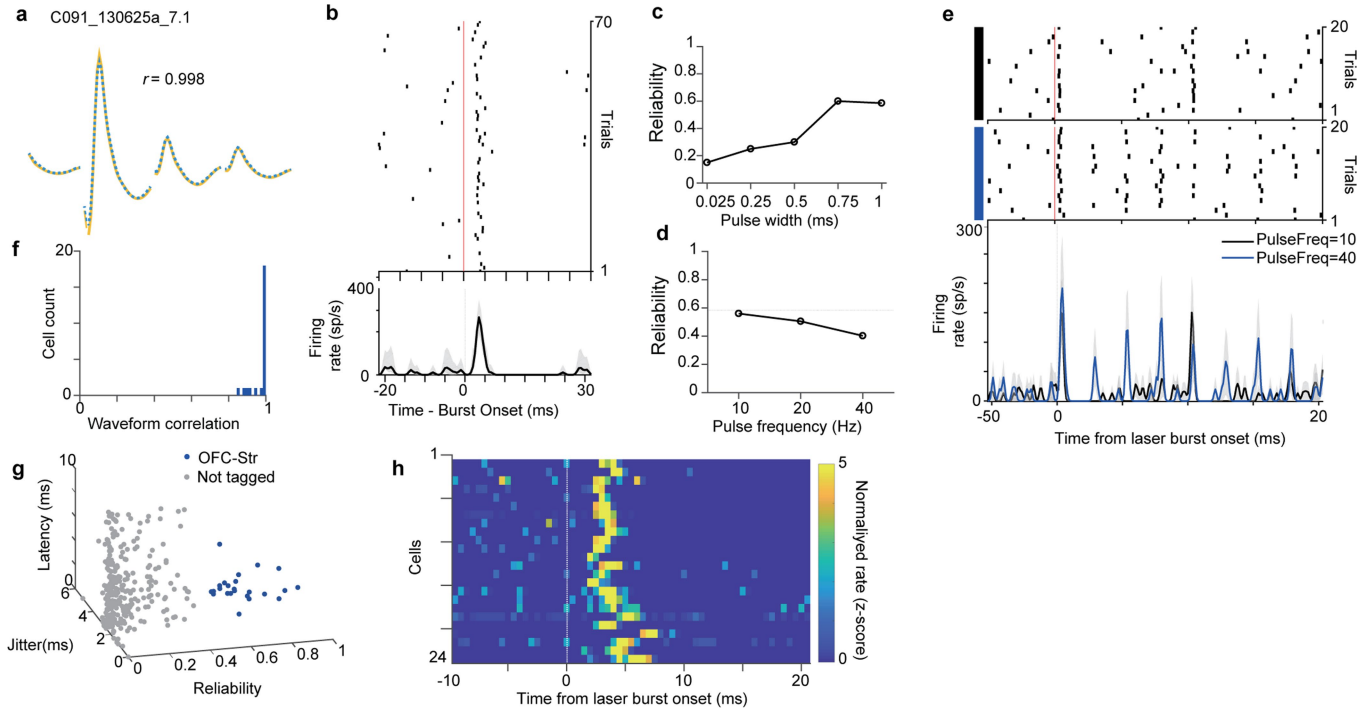
Single neurons

	Canonical Model	Random Rotation	p-value
RMSE	2.451 +/- 0.427	2.477 +/- 0.427	1
L1 Penalty	0.728 +/- 0.120	0.846 +/- 0.123	0.0002

Extended Data Fig. 8 | Negative confidence is quantitatively represented in a cluster of OFC neurons. a, Response profile of neurons in cluster 2 correspond to the decision-variable confidence$^{(+)}$. Panel (i): peri-stimulus time histogram of normalized firing rate, grouped based on the degree of evidence supporting choice. Panel (ii): trial-by-trial fit of each neuron in cluster 2 to choice evidence reveals significant representation of statistical decision confidence variable (R^2 with $P < 0.05$ based on bootstrap). Panels (iii–vi): normalized average tuning curves for neurons in cluster 2. Panel (iii) shows firing rate as a function of stimulus difficulty and choice (vevaiometric curve). Panel (iv) shows choice accuracy as a function of stimulus difficulty and firing rate (conditioned psychometric curve); (v) shows choice accuracy as a function of firing rate (calibration curve); and (vi) shows firing rate as a function of evidence supporting choice. Note that panels (iii) and (v) are replications of Fig. 4a bottom panels. **b**, Response profile of neurons in cluster 1, corresponding to a decision variable representation of confidence$^{(-)}$ (same convention as panel **a**). **c**, Response profile of neurons in cluster 3, corresponding to a decision-variable of integrated value$^{(-)}$. Panels (i–v): the representation of integrated value is analysed similarly to confidence in panel **a**, with the following changes: (ii): trial-by-trial fit of integrated value reveals significant representation of negative integrated value (R^2 with $P < 0.05$ based on bootstrap). (vi): firing rate as a function of negative integrated value. (v): Choice accuracy as a function of firing rate (*$P < 0.01$, *t*-test). (vi): Firing rate as a function of integrated value. Note that panel (iii) and (v) are replications of Fig. 4b bottom panels. **d**, Single neurons encode coherent combinations of confidence and reward size. Each neuron's response profile was fit to a two-parameter model representing confidence and reward size. For most neurons, regression coefficients (β) for each component share the same sign. Data are shown for all neurons (grey), and neurons with significant beta coefficients for both components are shown in blue ($P < 0.01$ threshold). Polar histogram is significantly different from uniform ($P < 0.01$). **e**, Elementary task variables defined for the regression model. Each task variable was *z*-scored according to the weight of its non-masked conditions, with masked conditions subsequently set to zero. **f**, Detailed results of the LASSO model from Fig. 4g for neuronal clusters, using both the canonical design matrix (corresponding to decision variables) and null models (corresponding to random rotations of the design matrix). Errors are shown as median ± s.e.m. *P* value calculated as paired sign-rank test. **g**, Detailed results of the LASSO model shown in Fig. 4h for single neurons for both the canonical design matrix (corresponding to decision variables) and null models (corresponding to random rotations of the design matrix). Errors are shown as median ± s.e.m. *P* value calculated as a two-sided *t*-test.



Extended Data Fig. 9 | Time course clustering and positively outcome selective OFC-striatum projection neurons. a, **b**, The ePAIRS (**a**) and eRP (**b**) tests reveals significant non-random clustering in the OFC population based on response profiles with temporal but not task-related information (see Methods). For ePAIRS, nearest neighbour angles were smaller than expected, suggestive of clustering (rather than dispersion). **c**, Clustering results and hyperparameter selection for temporal clustering (most stable configuration: $c = 8$ clusters and $k = 13$ nearest neighbours; compare to Fig. 3f; see Methods). **d**, Analysis of temporal response profiles (left); spectral clustering of temporal response profiles without tuning information reveals eight clusters with high within-cluster similarities ($n = 7$ rats combined; Methods). Dynamics of the trial-averaged time course for single neurons in the eight clusters for rewarded trials (middle) and error trials (right) are shown. We separated rewarded and error trials for this analysis as the actions performed during the outcome period are very different (drinking water versus return to centre port). **e**, Average dynamics of the trial-averaged time course for the eight clusters (green, rewarded trials; red, error trials). **f**, Average activity in correct trials (top) and average PSTH grouped by outcome (error, small reward, large reward) of identified OFC–striatum projecting neurons that positively encoded outcome. Lower panels show that neurons are positively tuned to integrated value in the anticipation period and positive tuning to outcome in the feedback epoch and ITI. Conventions are the same as in Fig. 5h. **g**, Average PSTH of neurons in cluster A whose dynamics match those of optogenetically identified neurons encoding positive outcome (excluding optogenetically identified OFC–striatum projection neurons). Note that the coding of integrated value is weaker than for the negative population but still significant ($r(90) = 0.1$, $P = 0.02$). Conventions are the same as in panel **f**.



Extended Data Fig. 10 | Optogenetic identification of OFC–striatum projection neurons. a–e, An example neuron showing reliable light-evoked responses. **a**, For an example neuron, average waveforms of spontaneous (yellow) and light-evoked spikes (blue) across four tetrodes are very similar. **b**, Spike raster (top) and PSTH (bottom) for the light-activated cell in panel **a** aligned to light onset (1 ms duration, first stimulus in a train). **c**, Reliability of the evoked responses to the first stimulus as a function of pulse duration. **d**, Probability of light-evoked spikes as a function of stimulation frequency (1 ms duration, 20 repetitions). **e**, Spike raster (top) and PSTH (bottom) aligned to light onset for stimulation trials at 10 Hz and 40 Hz. **f**, Histogram of Pearson's correlation coefficients between the waveforms of spontaneous and light-evoked for identified OFC–striatum projecting neurons. **g**, Quantification of light-evoked responses, showing latency and jitter of light-evoked spikes for tagged neurons as a function of the reliability of evoking a response to light. Putative OFC–striatum projection neurons are shown (blue points). **h**, *z*-scored PSTH of all identified OFC–striatum projection neurons in response to 1–3 ms blue light stimulation.

nature research

Corresponding author(s): Adam Kepecs

Last updated by author(s): Jun 5, 2019

Reporting Summary

Nature Research 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 Research policies, see Authors & Referees 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	Spike data was collected using NEURALYNX software, with spike sorting using MCLUST.
Data analysis	Code for ePAIRS and eRP analyes are available at https://github.com/agvaughan/EllipticalClustering.

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/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research 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 list of figures that have associated raw data
- A description of any restrictions on data availability

Neuronal data from all experiments is available on reasonable request.

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	Sample sizes are described by number of neurons, and number of unique animals. Although there is no clear guideline for the number of animals or neurons required to validate a given result, we have attempted to maximized the reliability of our result by performing a full replicate of our analysis on a second cohort of animals.
Data exclusions	No animals were excluded from analysis. Neurons were selected according to strict spike-sorting criteria, as described in the manuscript; we note that any failures of spike sorting (i.e. multi-unit recordings) would generally bias our results against the conclusions we obsereport.r
Replication	As discussed in the manuscript, we replicated our main result through a complete re-analysis of data drawn from a new cohort of animals trained separately.
Randomization	Not applicable as there is no group allocation in this study.
Blinding	Not applicable as there is no group allocation in this study. However, criteria for analysis of replicate data were determined before analysis (and matched the original analysis as closely as possible).

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
☐	☒ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

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	Male Long Evans rats (~250 g) were used for all experiments.
Wild animals	N/A
Field-collected samples	N/A
Ethics oversight	All procedures involving animals were approved by the Cold Spring Harbor Laboratory Institutional Animal Care and Use Committee and by the Animal Research Committee of Doshisha University, and were carried out in accordance with National Institutes of Health standards.

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