HHS Public Access

Author manuscript

Nat Neurosci. Author manuscript; available in PMC 2015 June 01.

Published in final edited form as:

Nat Neurosci. 2014 December ; 17(12): 1784–1792. doi:10.1038/nn.3865.

A category-free neural population supports evolving demands during decision-making

David Raposo[#1,2], **Matthew T. Kaufman**[#1], and **Anne K. Churchland**[1]

[1]Cold Spring Harbor Laboratory, Cold Spring Harbor, NY 11724, USA

[2]Champalimaud Neuroscience Programme, Lisboa, Portugal

[#] These authors contributed equally to this work.

Abstract

The posterior parietal cortex (PPC) receives diverse inputs and is involved in a dizzying array of behaviors. These multiple behaviors could rely on distinct categories of neurons specialized to represent particular variables or could rely on a single population of PPC neurons that is leveraged in different ways. To distinguish these possibilities, we evaluated rat PPC neurons recorded during multisensory decisions. Novel tests revealed that task parameters and temporal response features were distributed randomly across neurons, without evidence of categories. This suggests that PPC neurons constitute a dynamic network that is decoded according to the animal's current needs. To test for an additional signature of a dynamic network, we compared moments when behavioral demands differ: decision and movement. Our novel state-space analysis revealed that the network explored different dimensions during decision and movement. These observations suggest that a single network of neurons can support the evolving behavioral demands of decision-making.

Introduction

Individual neurons are often seen as members of highly specialized categories with response properties making them suitable for particular classes of computations[1,2]. This view has been fruitful for understanding early sensory areas where single neurons can be strongly tuned for task parameters, such as direction of motion[3] or disparity[4].

The assumption of neural categories is reflected in many experimental designs and analysis methods, even those focusing on neural structures far downstream of early sensory areas. This assumption can be evident in the way neurons are sampled: sometimes, neurons must meet certain response criteria to be included for study, such as responsiveness to certain stimuli or activity during a delay period[5-8]. Implicit in this approach is the idea that the

Correspondence and requests for materials should be addressed to Prof. Anne Churchland, churchland@cshl.edu.

Competing financial interests: The authors declare no conflicts of interest.

Author contributions: D.R. and A.C. designed the experiments. D.R. performed the electrophysiology and inactivations. M.K. developed the PAIRS, Variance Alignment, and decode analyses. D.R. and M.K. analyzed the data. A.C. wrote the paper. All authors discussed the results and implications and commented on the manuscript at all stages.

cell's response during one stimulus identifies it as a member of the category being examined. The assumption of categories can also be evident during analysis: pie charts, a common way of summarizing population data[9-11], explicitly assign neurons to categories. Another way of summarizing a population response, averaging over many neurons, likewise reflects the assumption that each neuron is an exemplar of a category, different from other category members mainly because of noise.

An alternative hypothesis is that neurons reflect random combinations of parameters, leading to neural populations in which neurons' responses defy categorization. Theoretical work suggests a major advantage for category free populations: when parameters are distributed randomly across neurons, an arbitrary group of them can be linearly combined to estimate the parameter needed at a given moment[12-14]. This obviates the need for precisely pre-patterned connections between neurons and their downstream targets, and also means that all information is transmitted. This latter property could allow the same network to participate in multiple behaviors simply by using different readouts of the neurons. Experimental work has not tested directly whether neural populations are category-free, but many observations are broadly consistent with this possibility. Specifically, recent studies have demonstrated that neurons in parietal[15-17] and frontal[18] areas have "mixed selectivity": individual neurons are modulated by multiple task parameters. Mixed selectivity would be expected if neurons reflect random mixtures of parameters, but also might exist under other assumptions. Other experimental work has probed for the existence of neural categories defined by the timing of a neuron's response[19]. That work argued against categories, but only tested for categories defined by response sequence. A more general test is thus required. Further, because neurons in that study responded sparsely, it was not possible to test whether the same neurons participated statically or dynamically in the network as the behavioral demands evolved from decision to movement.

Here, we developed a multisensory decision task rich enough to expose the functional organization of a neural population, both at a single moment and over the course of a complex choice with evolving behavioral demands. Our data suggest that in the posterior parietal cortex (PPC), the population is category-free: response features are randomly distributed across neurons. A possible explanation for this configuration is that it confers flexibility, allowing the brain to use the same neurons in different ways, depending on the current needs of the animal. In keeping with this explanation, we found that the population can be decoded instantaneously to estimate multiple task parameters, and that population activity explored different dimensions as the animal's needs evolved from decision formation to movement.

Multisensory decision-making behavior

We trained rats on an established decision-making task[20,21] in which animals reported a judgment about a 1s series of auditory clicks and/or full-field visual flashes (Fig. 1a-c). We refer to this 1s period as "decision formation" because previous studies have demonstrated that stimuli presented throughout the 1 s period influence the animals' decisions[20,21]. Once the stimulus terminated, animals reported whether the event rate of the stimulus was above or below an experimenter-imposed category boundary. They reported decisions via

movement to one of two choice ports. Rats were mostly stationary during stimulus presentation (Supp. Fig. 1a) and did not typically move towards or away from the direction of the port they ultimately chose (Supp. Fig. 1b). Rats mastered the ability to categorize the stimulus and report the decision regardless of whether stimuli were unisensory (visual-only or auditory-only) or multisensory (Fig. 1c). Similarly to other studies[20-22], when auditory and visual stimuli were presented together (multisensory trials), performance was enhanced (Fig. 1c, orange line steeper than green/blue lines).

PPC inactivation reduces visual performance

We first evaluated whether PPC inactivation affected decisions in any of the modalities tested. Even if the area is not causally involved in *all* modalities, having multiple task parameters that modulate neurons can greatly aid our understanding: it allows a broad search for categories that could be defined by a number of features. In addition, modulation from a second, non-causal modality could still be of interest to the animal, since these inputs might be required for behaviors beyond those studied here.

We suppressed spiking activity of PPC neurons using two complementary strategies. First, we made double bilateral infusions of muscimol, a $GABA_A$ agonist, into PPC (2 rats). We compared performance for inactivation versus control days and observed shallower psychometric functions (more errors) on inactivation days (Fig. 1d, $\sigma_{saline} = 2.64 \pm 0.39$; $\sigma_{mucimol} = 4.67 \pm 0.67$; standard errors computed by bootstrapping, see Online Methods). This example was typical: PPC inactivation reliably impaired visual decisions (Fig. 1e, *middle*, Supp. Fig. 2a,b; Mann-Whitney *U* test; $p < 0.001$, pooled across animals). Animals retained some ability to make visual judgments despite the inactivation, suggesting either that inactivation was incomplete, or that structures in addition to PPC support the task, consistent with other studies of parietal inactivation[6,23].

Impairment was specific to visual trials and had no consistent effect on auditory decisions (Fig. 1e, *left*; Supp. Fig. 2a,b; $p = 0.41$, pooled across animals). The sparing of auditory decisions reveals that inactivation did not reduce the animal's motivation, or introduce confusion about the stimulus-response contingency. Further, we observed sparing of the multisensory enhancement, the improved sensitivity for multisensory relative to unisensory decisions (Supp. Fig. 2a,b, orange bars smaller than green/blue bars). This spared enhancement implies that PPC likely does not drive multisensory enhancement and leaves open the possibility that PPC may process visual inputs before they are integrated with other modalities. Under this scenario, we would predict only a small change in multisensory performance even when visual inputs are impaired. This is because multisensory mechanisms can still use the remaining, weak visual signal to improve their estimation alongside the spared auditory signal. The framework for multisensory integration therefore predicts a very minor change in enhancement during inactivation of one modality, even if that modality is clearly impaired (see simulation, Supp. Fig. 2e-g).

One possible explanation for impaired visual decision-making is that the muscimol might have spread to portions of neighboring visual areas, posterior to PPC. This seems unlikely to have driven the effect, because the retinotopic organization of visual areas means that

restricted spread of muscimol would only have affected a portion of the visual field. Since our stimulus was full field, the unaffected parts of the visual field could likely have supported the behavior[24]. Nevertheless, we wished to determine the spread of the inactivation.

To achieve this, we used a second inactivation strategy: DREADD (Designer Receptor Exclusively Activated by Designer Drug), a pharmacogenetic inactivation method that permits visualization of the agent to determine its spread[25]. These effects were similar to the effects of muscimol inactivation in a second set of 2 rats: impairment of visual decisions (Fig. 1f, $\sigma_{saline} = 3.89 \pm 0.64$; $\sigma_{mucimol} = 5.29 \pm 0.71$) and sparing of auditory decisions and multisensory integration (Fig. 1g, Supp. Fig. 2c,d; Visual trials impaired: $p = 0.011$; Auditory trials spared: $p = 0.91$, pooled across animals). Histological examination (Supp. Fig. 3a-h) revealed that DREADD expression was minimal beyond the posterior border of PPC, defined as 5.0 mm posterior to Bregma[26]. For one rat, expression was less than 0.12% of maximal expression; for the second rat, expression was 20.3% of maximal expression (Supp. Fig. 3a-h; Online Methods). The more posterior expression was apparently not the source of the impairment because the rat with more expression posterior to PPC had weaker visual impairment compared to the other rat. We did not detect DREADD expression in other areas (Supp. Fig. 3i).

Overall, results were similar for all animals with both muscimol and DREADD inactivation. Effects were individually significant in 3/4 cases for visual trials and 0/4 cases for auditory trials.

Impairments on visual decision making might be driven by a change in the reliability of incoming visual signals, or by a change in the animal's decision-making strategy. Changes in decision-making strategy could include making "snap judgments" that relied only on evidence presented at the beginning of the trial, or "leaking" evidence and using only evidence presented late in the trial. To distinguish changes in stimulus reliability from changes in strategy, we performed an analysis of animals' decisions that took advantage of the ongoing fluctuations in visual and auditory rates that occur throughout the 1s decision formation period (Online Methods)[20,21,27]. The analysis measures whether fluctuations at any given moment influence the eventual choice. The analysis generates a quantity termed "Excess rate" which measures how strongly each moment in the stimulus influences the animal's eventual choice. This analysis revealed no evidence of snap judgments or evidence leak (Fig. 1h, light blue line is stable over the entire trial). Instead, the analysis indicates that inactivation reduced the signal-to-noise of incoming evidence (Fig. 1h, light blue line below dark blue line). No effects were observed on auditory trials (Fig. 1i, Supp. Fig. 2h-m). The reduced excess rate on visual trials and unchanged excess rate on auditory trials also confirms the outcome of the previous analysis using a model-free approach that does not rely on fitted parameters.

The reduced excess rate over the entire course of visual trials suggests that inactivation reduced the reliability or signal-to-noise of visual signals. Behavioral experiments in which we reduced the brightness of visual flashes affected psychometric functions and excess rate nearly identically to these inactivations[21]. Taken together, our inactivation experiments and

analyses suggest that PPC is required for accurate visual decision-making, perhaps by converting incoming visual signals into evidence for a decision. These observations point to PPC as causal for visual decision-making, laying the foundation for subsequent recording experiments that probe the functional organization and dynamics of cortical networks within PPC.

Choice and modality both modulate neural responses

To evaluate whether PPC neurons demonstrate mixed selectivity, we recorded from single, well-isolated neurons in the left PPC of 5 trained rats. Trials were grouped by modality (indicated in figures by color) and by the animal's choice (indicated in figures by solid vs. dashed line). Rare neurons had pure choice selectivity (Fig. 2a) or pure modality selectivity (Fig. 2b). However, the majority of neurons mixed information about modality and choice (Fig. 2c,d; Supp. Fig. 4). For such neurons, the mixing sometimes resulted in identical firing rates for different conditions (e.g., Fig. 2c, arrow).

We assessed the effect of the animal's choice on neural responses during decision formation. We used ROC analysis to generate an index of "choice divergence" that measures how strongly a neuron's firing rate for a correct leftward response diverges from the firing rate for a correct rightward response (e.g, dashed vs. solid traces in Fig. 2a). The divergence could be driven by a number of factors including accumulation of evidence for a decision, a developing motor plan, or a sensory preference for a particular stimulus frequency. Choice divergence became positive about 200 ms after stimulus onset and continued to grow over the course of the decision (Fig. 2e, colored traces). Choice divergence here was computed using "easy" unisensory trials (stimulus rates > 2 events/s from the category boundary) and was similar for auditory and visual stimuli (Fig. 2e, green vs. blue traces; multisensory in Supp. Fig. 5a). Stronger choice divergence was evident on multisensory trials at many points during the trial (Supp. Fig. 5b,c); stronger choice divergence was also evident on easy versus more difficult trials (Supp. Fig. 5d-f).

Responses on multisensory trials were usually well predicted by a linear combination of auditory and visual responses. Simple linear regression revealed that 80.1% of neurons (218 of 272 units) had a multisensory response that was better predicted by the auditory and visual responses than by the multisensory mean (assessed on left-out data). Across all neurons, a linear combination of visual and auditory responses accounted for a median 68.2% of the multisensory variance.

An additional index, "choice preference," captured not only the magnitude of the choice divergence, but also whether it was in favor of a high-rate or low-rate choice (Online Methods). Choice preference was significant in over a third of individual neurons for both auditory and visual trials (Fig. 2f,g, filled bars indicate neurons with significant selectivity 200 ms before movement; 35.5% / 37.3% of neurons were significant on auditory / visual trials; for multisensory, see Supp. Fig. 5g). Strong choice preferences for both ipsi and contra decisions were observed. For both auditory and visual decisions, a slight majority of neurons fired more in advance of ipsilateral, compared to contralateral choices (two-sided sign test, auditory: median choice preference = −0.061, $p = 0.0011$, N = 262 neurons; visual:

median choice preference = −0.038, p = 0.0502, N = 268 neurons). Choice preferences computed during visual and auditory trials were strongly correlated (Supp. Fig. 5h, N = 262 neurons, r = 0.74, $p < 10^{-4}$); preferences during each unimodal stimulus were also correlated with preference during multisensory stimuli (Supp. Fig. 5i, N = 236 neurons; auditory: r = 0.668, p < 0.001; visual: r = 0.807, p < 0.001).

We next assessed the effect of stimulus modality on responses during decision formation. "Modality divergence" measured how strongly a neuron's responses diverged for auditory versus visual trials (e.g., Fig 2b). Compared to choice divergence, modality divergence increased earlier and faster during the stimulus presentation but was weaker overall (Fig. 2e, black trace). "Modality preference" captured not only the magnitude of the modality divergence, but also whether it was in favor of visual vs. auditory stimuli. A third of the neurons (33.8%) had significant modality preference (Fig. 2h). Visual-preferring and auditory-preferring neurons were observed in nearly equal numbers (Fig. 2h, median modality preference was 0.017 and did not differ significantly from 0; p = 0.088; N = 269 neurons). For both choice and modality, similar results were achieved when we assessed selectivity using a rate-based rather than ROC-based analysis (data not shown).

PPC is category-free: neither task parameters nor temporal response patterns define clusters

The data thus far indicate that many individual neurons are strongly modulated by modality or choice. We next investigated how frequently individual neurons had mixed selectivity for modality and choice. If mixed selectivity is common, many neurons should have a non-zero choice preference and a non-zero modality preference. This is exactly what we observed (Fig. 3a). Neural responses were not restricted to pure selectivity (cross-shaped region in Fig. 3a, dashed lines). Instead, most neurons had mixed selectivity for modality and choice. Moreover, a major component of the mixed selectivity was linear: that is, when predicting the neuron's response to a given choice and modality (e.g., a high-rate choice for visual stimuli) linear sensitivity to each task parameter alone was more important than a nonlinear interaction between parameters (Supp. Fig. 6a). The nonlinear component we observed in PPC was smaller and more variable compared to neurons in the prefrontal cortex[18], perhaps suggesting that nonlinear mixed selectivity emerges gradually across cortical areas or depends on the nature of the task.

The existence of individual cells with mixed selectivity would be expected under two scenarios: response features might be randomly distributed across PPC neurons, or particular response features might cluster together, defining categories of neurons that are specialized for particular computations. The scatterplot in Figure 3a hints that choice and modality selectivity are randomly distributed across neurons. For example, choice and modality preferences were uncorrelated (N = 268 neurons, r = 0.074, p = 0.23). This is in keeping with studies from monkey PPC in which selectivity for spatial versus category parameters were likewise unrelated[15].

However, lack of correlation does not conclusively rule out the existence of functional categories: selectivity for task features could still define categories. This could be the case if,

for instance, the points in Figure 3a formed an “X”, or formed clusters that were symmetrically arranged around the origin. We therefore wished to test whether neurons formed categories (broadly construed) or whether, instead, tuning for one feature was independent of tuning for others. To do so, we examined each neuron’s “feature vector”: the pair of values describing how strongly the neuron contributed to decoding choice and modality (see next section; vectors shown in Fig. 3b). Each neuron’s feature vector was compared with its “nearest neighbors” in this feature space. If some neural responses fell into categories, these neurons would tend to have closer neighbors in feature space than if there were not categories[28]. The distribution of nearest-neighbor angles for the population can thus distinguish the presence or absence of such neural categories. We used these nearest-neighbor angles to compute a statistic indicating whether the population had an excess of small nearest-neighbor angles. We term this statistic “PAIRS”: Projection Angle Index of Response Similarity (Online Methods).

The PAIRS test did not indicate categories (Fig. 3c, gray bars). The distributions of nearest-neighbor angles were statistically indistinguishable from a control distribution generated by randomly oriented 2-dimensional vectors (Fig. 3c shows Rat 5, solid black line not different from gray bars, PAIRS index = −0.052; p = 0.632; Online Methods). No evidence for clear categories was present in any animal (PAIRS index for Rats 1-4: −0.135, 0.117, −0.080, −0.142; p-values from Monte Carlo simulations: 0.236, 0.253, 0.399, 0.004; the one significant p-value indicated less clustering than expected by chance). This observation is critical: it argues that neurons with pure selectivity (e.g., Fig. 2a and b) are exceptions and occur about as often as would be expected by chance.

This analysis argues that choice and modality selectivity do not define categories. However, this leaves open the possibility that there are categories defined by other features of the data. More generally, a category might be defined by a shared pattern of firing rates across conditions and time. To test for this, we used Principal Component Analysis (PCA) to identify a set of neural response features that were not imposed by us. This version of the test is thus quite general, because it captures whatever features of the responses were strongest and is sensitive to numerous such features (the 8 principal components from an example animal are shown in Supp. Fig. 6b). The PAIRS test again pointed to a category-free population (Fig. 3d,e; Supp. Fig. 6c). The overall lack of categories was not simply because neural variability caused our analysis to miss structure: when we introduced synthetic categories into the population with noise derived from the real neural data, a strikingly different PAIRS distribution was evident (Fig. 3e, dashed line; Online Methods). Relatively close neuron pairs were occasionally observed, but these differed from chance in only 1 of 5 cases (PAIRS index for Rats 1-5: −0.011, 0.108, −0.038, 0.011, −0.007; p-values from Monte Carlo simulations: 0.621, 0.001, 0.209, 0.491, 0.857; Supp. Fig. 6c). These deviations from the random distribution indicate that a small fraction of neurons do have pairs in feature space. However, such neurons are rare; the vast majority of neurons each reflect a unique combination of response features. Most individual neurons participate in random combinations of response patterns – that is, they randomly mix task parameters and temporal response features.

Decoding choice and modality from a mixed population

Here, we evaluate whether the mixed selectivity of PPC neurons poses any problem for decoding the key task variables. We first tested whether the animal's choice could be decoded from the population response during decision formation. To achieve this, we used a machine learning classifier (Support Vector Machine, SVM)[29,30] as our decoder, trained with single-trial population responses for correct high-rate versus low-rate choices (Online Methods). The decoder successfully identified neural weights so that a weighted sum of the neural population was strongly choice-*de*pendent (Fig. 4a, solid/dashed lines diverge) but mostly modality-*in*dependent (Fig. 4a, blue and green lines nearly overlap).

We trained the decoder using a portion of the stimulus epoch (500-700 ms after stimulus onset), then tested the decode over the entire epoch. Time windows outside the training window test the generality of the decoder and probe the consistency of the population response throughout decision formation. We first examined both correct and incorrect choices for animals with sufficient error trials (2 rats). For both rats (Fig. 4b), PPC activity tracked the animal's choice. As expected, decoder performance grew over the course of the trial. For all 5 rats (Fig. 4c), we examined correct-choice trials at the time point 700–800 ms after stimulus onset (outside the decoder's training epoch). Decoder performance was significantly better than chance for all animals tested (Fig. 4c, red lines; decode performances were 68.9%, 61.2%, 59.4%, 80.2%, 70.1%; 4 rats $p<0.001$, Rat 3: $p = 0.002$). The same decoder did not perform significantly better than chance when estimating stimulus modality (Fig. 4c, blue lines; performance of 48.0%, 51.0%, 54.3%, 48.8%, 48.6%, all rats $p > 0.2$). Our decoding of choice was not perfect, but this is unsurprising: decoding was performed on a sample of tens of neurons out of many thousands, on a task where the animal's performance is likely noise limited. Further, since the decoder indicated the opposite choice for error trials, this implies that correct choices could be distinguished from errors. Moreover, unlike typical monkey experiments[3,5,8], the stimulus was not optimized for each neuron's preference and neurons were not selected based on tuning properties.

The analysis above was restricted to auditory and visual trials. We next tested the ability of the same decoder, trained only on unisensory trials, to classify performance on multisensory trials. Decoder performance was significantly better than chance for all animals tested (Supp. Fig. 7; all $p < 0.001$). This speaks to the generality of the decoder and also highlights that the network is used similarly during decision formation regardless of the modality.

One concern might be that the choice decoder's success resulted from heavily weighting rare neurons with pure selectivity for choice (e.g., Fig. 2a). This was not the case. The neural weights were nonzero for all neurons, indicating that every member of the population contributed (Fig. 4d, black bars). Importantly, the distribution of neural weights in the data did not differ from those that would be expected by chance if neurons reflected random weightings (Fig. 4d, purple lines; for rats 1-5, p-values for kurtosis were 0.696, 0.470, 0.134, 0.198, 0.430; see Online Methods). This is an independent indicator that neurons mix information about task parameters with random weights.

Having examined choice, we tested whether stimulus modality could also be decoded from the population response during decision formation. Again, we trained an SVM classifier with single-trial population responses, this time to distinguish auditory versus visual trials. The decoder successfully identified a weighted sum of neurons such that the population readout was strongly modality-*de*pendent but choice-*in*dependent (Fig. 4e). The decoder was able to estimate modality at a rate better than chance when testing generalization for all 5 animals tested (Fig. 4g, blue lines; decode performances were 59.8%, 56.6%, 60.4%, 83.7%, 70.0%; p-values 0.010, 0.045, 0.002, < 0.001, < 0.001).

Modality was decoded nearly identically whether the animal's choice was correct or incorrect (Fig. 4f); this is expected because errors do not reflect incorrect categorization of modality. As before, the accurate performance seen at the end of the trial was for time points outside the training window. The same decoder did not perform significantly better than chance when estimating choice (Fig. 4g, red lines, performance of 48.0%, 51.6%, 50.4%, 50.2%, 51.5%, all p-values > 0.5). Again, the decoder achieved this performance using all the neurons; the distribution of weights, as with choice, did not differ significantly from the random distribution (Fig. 4h; for rats 1-5, all p-values > 0.4).

The ability of the same population of neurons to reliably and independently represent information about both choice and modality is a direct consequence of mixed selectivity: the joint modulation of neurons by choice and modality allows them to be combined in different ways to give rise to whatever estimate is needed.

The above results have two implications. First, the fact that choice and modality can be decoded independently implies that these two representations are nearly orthogonal in the population. That is, presenting a high vs. low rate stimulus evokes one pattern of activity across neurons, presenting a visual vs. auditory stimulus evokes another (different) pattern of activity across neurons, and these two patterns are unrelated. More operationally, a decoder vector summarizes the pattern of activity (across the population) evoked in response to one stimulus versus another. Since choice and modality can be decoded independently, the patterns of activity for different parameters must be uncorrelated, as expected from Figure 3a. This can be verified directly: the average angle between the choice decoder and the modality decoder was 86.5°, only slightly less than a perfectly orthogonal 90°. Second, the consistent decode performance over time implies that the choice representation remains in the same neural dimension (covariance pattern across neurons) in PPC over the course of decision formation, and the modality representation does the same.

The network explores different dimensions during decision and movement

Perhaps the neural state explores yet other dimensions (patterns of neural covariance) when the animal's brain needs to perform a substantially different function[19]. This could permit PPC to control what signals are routed to different areas at different times. As shown recently, exploiting additional dimensions can be particularly useful to control when movement should be produced[31]. We searched for a signature of the neural states either aligning (using the same dimensions at different times) or exploring different dimensions during different epochs. To do so, we compared two moments in the trial where the animal's

behavior differed: during decision formation, when animals remain still to integrate sensory signals (Supp. Fig. 1) and during movement, when they rapidly reorient their bodies to harvest a reward. Examination of PSTHs (two example neurons in Fig. 5a,b) shows that neural activity can differ substantially during decision formation and movement. For example, the neuron in Fig. 5b has an elevated response during trials preceding a leftwards choice, but is then suppressed during the leftwards movement used to report the choice. This "switching" of preference from decision formation to movement was observed frequently (Fig. 5c).

To quantify the alignment of state spaces during decision and movement, we developed a novel analysis that we term Variance Alignment (Online Methods). The intuition behind this analysis is that if neurons' firing rates co-vary in similar ways during decision and movement, then the dimensions that best capture the variance for one epoch will also capture much of the variance for the other epoch. The alternative is that the dimensions that account for a lot of variance in one epoch will account for little variance in the other epoch – that is, that neurons will co-vary in completely different patterns from one epoch to the other. In this latter scenario, the neural state spaces for the two epochs can be described as misaligned.

We tested for alignment by computing an index that describes whether the dimensions that capture most of the variance during movement likewise capture the variance during decision formation. Crucially, this measure describes whether the neural state moves through the same dimensions (i.e., maintains the same patterns of covariance) – not whether the trajectories are similar within those dimensions. Four two-dimensional projections of decision epoch data are shown in Figure 5d-g. Each panel contains four traces, showing the responses to high and low rate stimuli for visual and auditory conditions. A projection onto the first two principal components is shown in Figure 5d, with the magenta ellipse marking one standard deviation of the data in that projection. For comparison, we can view different two-dimensional projections of the exact same data, with the projection chosen based on the movement-epoch activity (Fig. 5e), a random projection (Fig. 5f), or the smallest two principal components (Fig. 5g). The key element of these plots is the size of the magenta ellipse: if a space captures the decision-epoch variance well, then the ellipse should be nearly as large as in Figure 5d. Surprisingly, we found that dimensions that captured considerable variance during the movement captured far less variance during decision formation (Fig. 5e has a much smaller ellipse than Fig. 5d, similar to Fig. 5f), arguing against aligned state spaces for decision formation and movement. Indeed, the amount of alignment was slightly less than would be expected by chance (Fig. 5h). This absence of alignment (black and red lines similar, index near zero) or even significant misalignment (black line below red, index near −1) was present in all animals tested (Rat 1, index = −0.498, $p = 0.0012$; Rat 2, index = −0.145, $p = 0.51$; Rat 3, index = −0.559, $p = 0.003$; Rat 4, index = −0.230, $p = 0.27$; Rat 5, index = −0.219, $p = 0.20$).

For comparison as a positive control, we repeated the Variance Alignment analysis on neural responses during decision formation, comparing two different stimulus conditions: visual and multisensory. This analysis, by contrast, revealed strong alignment (Fig. 5i, index near 1 for rat 4; index = 0.840, $p < 10^{-4}$),present in all rats tested (indices were 0.644, 0.883,

and 0.812 for rats 1, 2, and 3, $p < 10^{-4}$ in all cases). This strong alignment indicates that the inherent noisiness of neural responses does not cause neural states to falsely appear misaligned, and therefore provides reassurance that the misalignment of states during decision and movement indicates a real difference in neural covariance. Thus, PPC employs highly similar patterns of population activity during visual and multisensory stimuli, but explores quite different patterns of population activity during stimulus than during movement.

Finally, we asked whether the misalignment of stimulus and movement activity might be due to simply having different neurons active during these two epochs. This hypothesis predicts that neurons that are strongly modulated during the stimulus should be less modulated during movement, and vice versa. This was not the case: instead, we observed a strong correlation between modulation during stimulus and modulation during movement (Fig. 5j, correlations of log(stimulus) to log(movement) for Rats 1-5 were: $r = 0.709$, $r = 0.655$, $r = 0.577$, $r = 0.666$, $r = 0.763$; all $p < 10^{-4}$). This implies that the misalignment of stimulus and movement activity is a population-level phenomenon, not due to having separate groups of neurons.

Discussion

We used a multisensory decision task to understand the organization and dynamics of PPC, an area that we demonstrate to be causal for visual decisions. We found that PPC neurons have mixed selectivity for two task parameters: the animal's developing choice and the modality of the stimulus. We used a novel test, PAIRS, to demonstrate that task parameters and time-varying response features are distributed randomly across neurons. This configuration does not pose a problem for decoding: a linear SVM could accurately estimate the modality of the stimulus and the rat's choice based on single-trial responses. A final analysis further demonstrated the flexibility of the population: our Variance Alignment test revealed that the network explores different dimensions during decision and movement. This may allow PPC to translate the decision about rate into an abstractly related action. Taken together, these results point to PPC neurons as a category-free population that is combined dynamically as the behavioral demands of a complex decision evolve.

Theoretical motivations for functionally specialized neurons, and their existence in early visual areas[1], has driven a widespread assumption of categories that has influenced both experimental design[5-8] and analysis[9-11,32]. Our finding of a category-free neural population challenges these assumptions about the organization of cortical structures. As a caveat, we note that neural categories defined by other properties, such as cell type or connectivity, might reveal specialization. Indeed, a few studies have found projection-based categories that are functionally specific[33-35], although many other studies report that connectivity-defined categories are functionally diverse[36-38]. By demonstrating here that cortical areas can lack categories defined by selectivity to task parameters (Fig. 3a-c) or by time dependent response features (Fig. 3d,e), our findings invite a new approach to interpreting population data. Specifically, future studies can test directly for the existence of categories, and design appropriate analyses if neurons are shown to reflect random combinations of task parameters, as they are here.

Although individual neurons reflected random combinations of task components (as predicted by theory[13,14,39]), the observed responses were nonetheless structured. Specifically, we observed that the majority of neurons that were driven by choice had "tolerance" for modality: they retained their choice preference whether the stimulus was auditory, visual or multisensory (Supp. Fig. 5h,i). Neurons in monkey inferotemporal cortex (IT) are likewise tolerant: many neurons have a preferred stimulus identity that is stable, though modulated, across many retinal positions[29]. Indeed, our task configuration is reminiscent of that used to study object recognition – just as a given object can be viewed from two different angles, a "low rate" decision here can be informed by two different modalities. In IT, the possibility of using the same linear decoder under many conditions, indicating tolerance, is taken as evidence that the neural data has been reformatted from an original "tangled" representation in earlier sensory areas[29]. In PPC, analogously, a linear decoder was capable of reading out the animal's choice independent of modality. Combined with evidence that PPC responses are nearly linear functions of choice and modality, this suggests that PPC may likewise be at an advanced stage of processing where representations have been untangled to guide decisions. To obtain such a representation, multiple stages of reformatting may be required[40]; this may explain the surprising prevalence of multisensory neurons in early sensory areas[41].

In our study, as in primate work, neural responses in PPC seem likely to reflect a process of transforming ambiguous sensory information into action. As in primate vision studies[5,8], responses gradually diverged according to the eventual decision outcome (Fig. 2a-e); the response divergence had a long latency, but was evident many hundreds of milliseconds before the animal reported the choice. This was true for both auditory and visual decisions. Primate PPC neurons are active in advance of movements driven by auditory stimuli as well[42,43]. However, it was not known whether PPC neurons were causally involved. Our inactivation results were surprising in that auditory decisions were spared despite a clear signature of the developing choice in PPC neurons. Auditory responses in PPC, though apparently not necessary for this task, may be invoked by other decision tasks, such as those that require the animal to decide when to stop accumulating evidence[8], those that require a report of confidence[44], or those that require temporally precise multisensory information[45].

Two methodological differences between our study and primate decision-making studies are notable. First, our stimuli were full-field rather than spatially restricted, and were related abstractly to their required movement (e.g., low rate, move left). These features may explain why we found no bias for contralateral movements (Fig. 2f,g). Second, we recorded from all encountered neurons and used identical stimuli for each. The more traditional approach of using neuron selection criteria and customized stimuli is successful in identifying neurons with strong choice signals, but may leave unexamined neurons with subtler choice signals that nonetheless shape the evolving decision. This point is underscored by our observation that the majority of our neurons contributed to the choice and modality decode, including neurons that were modulated only weakly by those parameters (Fig. 4d,h).

PPC therefore represents multiple behaviorally relevant variables in the same population of neurons, with these representations structured in a way that could allow easy decoding by subsequent cortical areas – perhaps especially those that inform movements. These patterns

of activity are dynamic and task-dependent, and are determined by more than connectivity alone. This use of different patterns of activity could confer flexibility to PPC in converting stimuli into action, and highlights the importance of understanding the population activity over the course of decision formation.

Online Methods

Behavior

Animal training and behavioral task—Adult, male Long-Evans rats (approx. 250g) were trained following previously established methods [20,21,46]. Briefly, rats were trained to wait in the center port while stimuli were presented, and to associate stimuli with reward ports. Stimuli for each trial consisted of a series of "events": auditory clicks, full-field visual flashes, or both together. Stimulus events were separated by either long (100 ms) or short (50 ms) intervals. For the easiest trials, all inter-event intervals were identical, generating rates that were 9 events/s (all long intervals) or 16 events/s (all short intervals). More difficult trials included a mixture of long and short intervals, generating stimulus rates that were intermediate between the two extremes and therefore more difficult for the animal to judge. Rats 1-4 were rewarded with a drop of water for moving to the left reward port following low-rate trials and to the right reward port following high rate trials. Rat 5 was rewarded according to the reverse contingency.

For rat 1, the stimulus began immediately when the rat's snout broke the infrared beam in the center port. For rats 2, 3, 4 and 5, a variable delay separated the time of entrance to the port and the start of the stimulus. The length of this delay was selected from a truncated exponential distribution (rats 2-4: λ = 30 ms, min = 10 ms, max = 200 ms; rat 5: λ = 33 ms, min = 0, max = 1000 ms) to generate an approximately flat hazard function. The total time of the stimulus was usually 1000 ms. For rat 1, the stimulus ranged in duration from 500-1000 ms.

Trials of all modalities and stimulus strengths were interleaved. For multisensory trials, the same number of auditory and visual events were presented. Our previous work has demonstrated that rats make nearly identical decisions regardless of whether stimulus events are presented synchronously or independently [20]. The majority of neurons here were recorded using synchronous stimuli with the exception of a few sessions for Rat 1. No obvious effect of synchronous vs. independent stimuli on the neurons was apparent.

Animals typically completed between 500 and 1200 trials/day. Most experiments had 18 conditions (3 modalities × 6 stimulus strengths) leading to ~ 27-67 trials/condition/day.

Analysis of behavioral data—Four-parameter psychometric functions were fit to choice data using the psignifit version 3 toolbox for MATLAB (http://psignifit.sourceforge.net), following the maximum likelihood methods described by Wichmann & Hill [47]. Psychometric functions were parameterized as

$$f(r,\mu,\sigma,\gamma,\lambda) = \gamma + (1-\gamma-\lambda)\left[1+erf\left(\frac{r-\mu}{\sigma\sqrt{2}}\right)\right]$$

where r is the trial event rate, μ and σ are the first and second moments of a cumulative Gaussian function, γ and λ are the guessing and lapse rates (constrained so that $0 \leq \gamma, \lambda \leq 0.1$), and *erf* is the error function. σ is referred to as the psychophysical threshold; smaller σ results in a steeper psychometric function. Standard errors for σ were computed via bootstrap analysis of the choice data (2000 resamples). To assess the magnitude of inactivation effects, we took the ratio of σ values computed for pairs of inactivation vs. control days. This ratio is termed "Impairment ratio" (Fig. 1e,g). Values > 1 indicate impaired performance on inactivation days relative to control days.

The excess rate analysis, described in detail elsewhere[20,21], complements the psychometric function as a means of quantifying the animal's decision-making behavior. The idea is to relate momentary fluctuations in the instantaneous rate with the animal's choice by computing a quantity termed "excess rate" in sliding 200 ms windows in the trial. Consider an example window from 0-200 ms after stimulus onset. Three steps are required. First, we select all trials in which the rate outside the window (e.g., 200-1000 ms) is neutral. The resulting group of trials differ only in the stimulus rate presented from 0-200 ms. Next, we separate trials into groups where the rat made a left vs. right choice. Finally, we average the rate for each group and take the difference in rate between trials preceding right vs. left choices. If the difference is zero, this indicates that trials preceding left and right choices were identical and that the time window under study did not influence the choice. Stimulus rates in excess of zero indicate that the window under study did influence the choice. This process is repeated for sliding 200 ms windows, generating excess rates for every moment in time. Excess rate for data is compared to a shuffle (thin black line on Figs. 1h,i and Supp. Fig. 2h-m) in which trials are randomly assigned to a "left" and "right" pool.

Implants for electrophysiology

Custom implants were prepared in-house. Each assembly contained up to 8 independently moveable tetrodes (Nickel/chrome alloy wire, 12.7 μm, Sandvik–Kanthal). Tetrodes were connected to an EIB-36 narrow connector board (Neuralynx) mounted on the assembly. The assembly was secured within a plastic enclosure prior to implanting. Tetrodes were gold-plated to 300-700 kΩ at 1 kHz; one additional tetrode was used as an internal reference for electrophysiological recordings and plated to ~100 kΩ.

General surgical procedures

All rats subject to surgery were anesthetized with isoflurane and administered 5 mg / kg ketoprofen before surgery for analgesia. Isoflurane anesthesia was maintained by monitoring respiration and foot pinch responses throughout the surgical procedure. Ophthalmic ointment was applied to keep the eyes moistened throughout surgery. Lidocaine solution (~0.1 mL) was injected below the scalp to provide local analgesia prior to performing scalp incisions. 0.05 mg / kg buprenorphine was administered daily for post-surgery analgesia (usually 2-3 days). Animals that received implants or cannulae were trained prior to surgery and then recovered to normal performance. All surgical and behavioral procedures conformed to the guidelines established by the National Institutes of Health and were approved by the Institutional Animal Care and Use Committee of Cold Spring Harbor Laboratory.

Injection surgery procedures—Two rats, 3-5 weeks of age, were anesthetized and placed in a stereotaxic apparatus (Kopf Instruments). Small craniotomies were made over PPC (3.8 mm posterior to Bregma; 2.2, 3.2, and 4.2 mm left / right of midline). One of the rats was subject to unilateral injections (left hemisphere), and the second rat was subject to bilateral injections. Small durotomies were performed at each craniotomy and virus was pressure injected at depths of 400, 600, and 800 μm below the pia (140 nL / depth) using calibrated pipettes and a syringe (rate of ~1 nL / second). 2-3 minutes were allowed following injection at each depth to allow for diffusion of virus. Adeno-associated virus expressing muscarinic receptor hM4D-mCitrine under an hSyn promoter (AAV5-hSyn-HA-hM4D-IRES-mCitrine; construct provided by Bryan Roth, UNC; virus produced by UNC Gene Therapy Center) was used.

Cannulae implant surgery—Rats were anesthetized and placed in the stereotax. Two craniotomies were made on each side of the brain; these were positioned to cover medial and lateral PPC on each side (4.0 mm posterior to Bregma and extending from 2.0 to 3.6 mm left / right of midline). Durotomies were performed and a double guide cannula (PlasticsOne, C235G-1.2) was placed in the brain 100-200 μm below the pia at each craniotomy. The exposed brain was covered with 2% agarose solution and both cannulae were anchored to the skull with dental acrylic (Lang Dental).

Implant surgery procedure—After scalp shaving and incision, the skull was cleaned, and anchoring screws were drilled into 6 locations on the skull. Dental cement (Parkell, Inc.) was applied to the skull surface and a craniotomy was made above left PPC (4 mm posterior to Bregma; 2.5 mm left of midline; ~2.4 mm anteroposterior × ~3.4 mm mediolateral in size). A durotomy was performed and the implant assembly was lowered until the tetrodes just penetrated the pial surface. 2% agarose solution was applied to cover the tetrodes and craniotomy, and dental acrylic (Lang Dental) was applied to secure the implant to the skull. The incision was closed around the base of the implant using Vetbond (3M). Following surgery, tetrodes were advanced in increments of 40-80 μm until action potentials were encountered.

Inactivation

Muscimol inactivation sessions—Muscimol was infused into PPC with a concentration of 0.5-1.0 mg / mL and a volume of 0.3 μL per site. A double-internal cannula (PlasticsOne, C235I/SP), connected to 2 microliter syringes (Hamilton microliter syringe, 7000 series), was inserted into each previously implanted guide cannula. Internal cannulae extended 0.5 mm below the guide (estimated 600-700 μm below the pia). Muscimol was delivered using an infusion pump (Harvard PHD 22/2000) at a rate of 0.1 μL / min. Internal cannulae were kept in the brain for 3-5 additional minutes to allow for diffusion of muscimol. Rats were removed from anesthesia and returned to cages for 30 minutes before beginning behavioral sessions. The same procedure was used in control sessions, where muscimol was replaced with sterile saline.

DREADD inactivation sessions—Prior to DREADD inactivation and control sessions, clozapine N-oxide (CNO, 1 mg / kg) or sterile saline was injected intraperitoneally into rats

expressing the muscarinic receptor hM4D. Animals were returned to their cages for 30 minutes post-injection before beginning behavioral sessions.

Histology

At the conclusion of physiological experiments, animals were deeply anesthetized with ketamine and medetomidine. To indicate the final positions of electrodes, electrolytic lesions were made at the tetrode tips by passing 30 μA current through each electrode for ~10-15 seconds. After lesioning, animals were perfused transcardially with 4% paraformaldehyde. Brains were extracted and post-fixed in 4% paraformaldehyde for 24-48 hours. After post-fixing, 100μm coronal sections were cut from one of the brains on a vibratome (Leica).

At the conclusion of inactivation experiments, both animals that had been injected with DREADD were perfused transcardially with 4% paraformaldehye. One of the brains was extracted, post-fixed and sectioned following the protocol described above. The second brain was post-fixed, then kept in 30% sucrose solution for 48 hours, then frozen at −80°C. 20 μm coronal sections were cut from this brain using a cryostat (Leica CM1850).

In both cases, brain sections were mounted on slides with Vectashield mounting medium.

Quantification of DREADD expression—Brain sections were imaged using an epifluorescence microscope. The resulting images were analyzed with MATLAB software. A region of interest (ROI) was manually defined for each brain section that was analyzed. The ROI extended from ~1.5 to 5.0 mm lateral to the midline and ~0.2 to 1 mm below the pia. A second, smaller ROI was defined near the first one, in a region that was not infected by the virus (which thus should have had no expression). This region was used as a measure for baseline pixel intensity. Average pixel intensity across columns of pixels was calculated inside the first ROI, then normalized by the average pixel intensity inside the second ROI (baseline). To quantify expression levels for a particular brain section, we calculated the area below the average pixel intensity curve for that section and above baseline (a flat line at unity). We used this measurement to compare the expression levels in two places: at the border between PPC and secondary visual cortex, and near the injection site in PPC (Supp. Fig. 3a-h).

Electrophysiological methods

Electrophysiological data was collected daily while animals were engaged in the task. Spike-triggered waveforms were recorded from each tetrode using Digital Lynx SX hardware and Cheetah data acquisition software (Neuralynx, Bozeman, MT). Data were acquired with a sampling rate of 32 kHz, and spike waveforms were bandpass filtered at frequency ranges of 600 – 6,000 Hz. Tetrodes were moved 40-80 μm after each recording session to ensure that independent populations of neurons were sampled across sessions.

Monitoring of head/body orientation during recordings—We used two methods to monitor the animal's orientation during electrophysiology sessions. First, we connected red and green LEDs to the animal's implant and tracked head orientation throughout the behavioral session using Cheetah data acquisition software (Neuralynx, Bozeman, MT).

LED positions were sampled at 30 Hz. Head angles were computed at each sample time and then smoothed with a Gaussian. For the second method, we used an open-source software package (Bonsai; G. Lopes, https://bitbucket.org/horizongir/bonsai) to track the animal's whole body orientation. Body angle was sampled at 100 Hz. The estimates produced using the implant LEDs and body tracking were generally in good agreement, although there tended to be more variability in body angle than head angle (e.g., the rat's head could remain stationary in the central port despite small body movements).

Analysis of electrophysiological data—Raw spike-triggered waveforms were manually clustered using MClust software (A.D. Redish) for MATLAB (Mathworks). Only isolated clusters corresponding to single neurons were included for analysis. Neural recordings were also trimmed or excluded if a portion of the recording had a strongly non-stationary mean firing rate over time, based on automated criteria. In addition, neurons had to satisfy a signal-to-noise criterion. Specifically, the firing-rate range (over conditions and times) divided by the maximal s.e.m. (for all conditions and times) had to be greater than 3.3.

Peri-stimulus time histograms (PSTHs) were computed for two epochs in the trial: a "decision formation epoch" (the time during stimulus presentation and enforced central fixation) and a "movement epoch", for which the spike trains were aligned to the stimulus onset or to the movement onset, respectively. Firing rates were averaged across like trials and smoothed over time with a Gaussian kernel (σ = 50 ms).

Data from 5 rats were analyzed. One animal in the cohort had stimulating fibers implanted alongside tetrodes. This animal was used as a control for a separate optogenetic study. For 9 of 18 electrophysiology sessions in this animal, laser stimulation (473 nm) was introduced through the fibers on 50% of trials. The animal expressed no light-activated ion channels in its brain, however, and laser stimulation had no effect on neural activity or behavior.

Choice selectivity and modality selectivity—PSTHs were constructed from spike trains by averaging firing rates within 10-ms bins and smoothing with a Gaussian kernel (σ = 50 ms). Correct trials were grouped according to two different aspects of the trials. The first way was based on the animal's response: trials ending in a contralateral choice versus trials ending in an ipsilateral choice. The second way was based on the stimulus modality: visual trials versus auditory trials. We used ROC analysis to calculate the ability of an ideal observer to correctly classify the animal's choice or the stimulus modality. This was done on each trial from the smoothed spike trains, at intervals throughout the trial. Choice and modality preference were derived from the area under the ROC curve (AUC) and defined for each time point as 2 × (AUC - 0.5); this value ranged from −1 to 1 [48]. A choice preference of −1 indicates that a cell always fired more during trials ending in an ipsilateral choice; a value of 1 means that the cell always fired more during trials ending in a contralateral choice. Modality preference was computed separately for rightward and leftward trials and averaged. A modality preference of −1 indicates that a cell always fired more during auditory trials; a modality preference of 1 means that the cell always fired more during visual trials.

Choice divergence was computed the same way as choice preference except that each neuron was assigned a "preferred" choice or modality based on its responses at the end of the trial (100-200 ms before movement onset). Choice divergence at other time points was computed based on this preference. This is a closely related measure to the absolute value of choice preference, but this way of computing the index has the advantage that it prevents small fluctuations in selectivity due to noise (either positive or negative) at the beginning of the trial from giving the incorrect impression that the neuron is selective before stimulus onset [49]. For choice and modality preference, significance ($p < 0.01$, one-sided, Fig. 2f-h) was assessed via bootstrapping (1000 iterations). A neuron was considered to have significant modality preference if this value was significant for either rightward or leftward trials.

Analysis of response clustering—To test for the presence of neural clusters, we developed a novel analysis, PAIRS: Projection Angle Index of Response Similarity. To calculate the statistic, we first built a matrix of the trial-averaged neural data (the "A" matrix). This matrix had *n* rows by *ct* columns, where *n* is the number of neurons, *c* is the number of conditions (6 conditions: 2 choices times 3 sensory modality conditions), and *t* is the number of time points in the decision epoch (including 300 ms before stimulus onset). We then reduced the dimensionality of this matrix using one of two methods. One method was to perform principal component analysis (PCA) on the A matrix, reducing the dimensionality (number of rows) to 8. This dimensionality was estimated from the data, and the results were not sensitive to the exact dimensionality used. For the alternative "feature based" method, we used the two dimensions specified by our choice decoder and our modality decoder. In this case, the two dimensions were orthogonalized using the Gram-Schmidt algorithm (since they were nearly orthogonal but not perfectly so) to ensure that they captured independent variance. The PCA method is assumed for further description below.

The coefficients matrix resulting from PCA is of size 8 by *n*; that is, each neuron received a single 8-element vector representing its response profile across conditions and over time. If a pair of neurons had similar response profiles, they would receive similar vectors (i.e., the vectors would form a small angle). For each neuron, we found the angle it made with each of its *k* most similar partners (e.g., $k = 3$ partners). For each neuron, these *k* values were then averaged. This produced a distribution of near-neighbor angles, θ_{data}, with one angle per neuron. We took the median of this distribution, denoted $\widetilde{\theta}_{data}$, which will be small for data with strong clustering of neural responses or larger if there is little or no clustering.

For comparison, we generated 10,000 simulated datasets composed of *n* random vectors from a 2- or 8-dimensional Gaussian distribution, as appropriate. For each simulated dataset, we then computed the distribution of angles θ_{random} using exactly the same method of *k* nearest neighbors. This random-derived distribution is plotted as the black line in Fig. 3c,e. Note that the distribution for each simulated dataset depends on the number of neurons in the original dataset (*n*), the dimensionality (here 2 or 8), and *k*, and that each of the 10,000 datasets was independent from the others. We collected the resulting $n \times 10{,}000$ angles, then

computed the median near-neighbor angle $\widetilde{\theta}_{random}$ from this grand distribution. For each rat, we then computed the PAIRS statistic:

$$PAIRS=\frac{\widetilde{\theta}_{random}-\widetilde{\theta}_{data}}{\widetilde{\theta}_{random}}$$

This statistic is 1 if all neurons have at least *k* identical partners, and 0 if clustering is only as strong as expected by chance. Since we could compute $\widetilde{\theta}_{random}$ for each of the 10,000 simulated datasets separately, we used these values to find the distribution of the PAIRS statistic expected by chance. A two-sided p-value was then computed by comparing this distribution to the PAIRS statistic obtained from the neural data.

The value *k* was selected automatically. To do so, we found the smallest value of *k* for which $\widetilde{\theta}_{random}$ exceeded a target value. When using the PCA method, this value was π / 4 (halfway between 0 and orthogonal, π / 2). For our data, this produced values of *k* from 2-4. When using the feature based method, the space was only two dimensional and therefore the feature vectors were packed more tightly. To avoid having overly large values of *k* (which would limit us to finding only large clusters), we chose a smaller value for the target angle: π / 8. This produced values of *k* from 9-24.

We also compared the neural data with synthetic data that had artificial clusters introduced. To do so, we first selected one fifth of our neurons at random. For each neuron selected, we generated a quintet of "partner" neurons related to the original. These partners were produced by resampling trials with replacement, then producing PSTHs as in the real data. In the resulting synthetic data, there were as many faux neurons as in the original data, but each faux neuron was related to four others. These faux neurons had exactly the same amount of noise as their originating neurons. This process was repeated 1,000 times. The distribution of the resulting $\theta_{clustered}$ values is plotted as the dashed line in Figure 3e.

Finally, we note that PAIRS is not only a test for clustering, but also more generally for non-uniformity of the distribution of tuning across neurons. For example, if a strong majority of neurons "preferred" the high-rate stimulus, then the neurons' coefficient vectors would be mostly packed into half the space. The near-neighbor angles would therefore be reduced relative to fully random, and PAIRS would detect "clustering" (properly, non-uniformity). While it is probably not possible to detect all conceivable ways in which the data might cluster, this method is a reasonably general test.

Decoding neural responses

The goal of the decoding analysis was to train a trial-by-trial classifier that could identify left versus right choices but was tolerant of modality, or separated auditory from visual trials but was tolerant of choice. That is, we looked for a weighted sum of neurons such that the result was high for one choice and low for the other regardless of modality, and a second weighted sum of neurons that was high for one modality and low for the other regardless of choice.

For a neuron to be included in this analysis, we required at least 20 successful trials each for all four choice / modality pairs (only visual and auditory modalities were used). Most of our neurons were not recorded simultaneously; we therefore constructed “pseudo-trials” by choosing random trials of the desired condition (defined by choice and stimulus modality) for each neuron. The epoch from 500 to 700 ms after stimulus onset was used for training the classifier. We assembled as many pseudo-trials as possible by randomly sampling trials from each neuron without replacement; because the number of trials from each neuron was balanced across conditions when training the classifiers, the process was limited by the neuron with the fewest trials in any one condition (21-32). We then used a standard machine learning technique, the linear Support Vector Machine (SVM)[30], to train one classifier for choice and a second for modality. Training was repeated 1,000 times with different random samples of pseudo-trials, resulting in 1,000 trained classifiers each for choice and stimulus modality. These were combined using a standard technique called bootstrap aggregation (“bagging”, [50], described below.

Each of these classifiers is a vector consisting of a set of linear weights, with one weight per neuron. We averaged the 1,000 trained classifier vectors to obtain a final classifier orientation for choice and a final classifier orientation for modality. In addition, the classifiers required a threshold. To obtain the thresholds, we randomly sampled additional balanced sets of pseudo-trials, projected them onto our classifier vector, and found the optimal threshold based on Gaussian fits to the two classes (high vs. low rate or visual vs. auditory). Classifier thresholds were averaged across 25 iterations of this process.

To test the performance of the classifier, we randomly sampled additional sets of pseudo-trials. Spike trains were filtered with a 100 ms boxcar, then each time point was classified. This was repeated 1,000 times and performance was averaged. To assess statistical significance, we assessed generalization performance on the epoch from 700 to 800 ms after stimulus onset. Performance of the classifier on 1,000 pseudo-trials was compared with the performance of 10,000 random classifiers on the same number of pseudo-trials. To generate the random classifiers, we first chose a random vector with n elements (n the number of neurons). Since different neurons have unequal variances and high-variance neurons will tend to be used more heavily by a trained classifier, we multiplied each element of the random vector by the standard deviation of the corresponding neuron. P-values are two-sided. Additional cross-validation of the classifier was achieved by using classifiers trained on visual data and tested with multisensory data (Supp. Fig. 7). This analysis also ensures that the success of the classifier outside the training window was not due to temporal correlations in the data because no temporal correlations will exist between data collected on different trials (multisensory vs. visual).

To interpret the weights chosen by the classifier, we examined whether only a subset of neurons might be used heavily, or whether all the neurons were used. To do so, we compared the distribution of weights from the classifier found using the real data with the distribution of weights from the random classifiers described above (shown as purple lines in Fig. 4d,h). To evaluate whether the real classifier was significantly different from the random ones, we computed the kurtosis of the distribution of weights from the data, and the kurtosis for each random classifier. To obtain a p-value, we could then compare the kurtosis

for the real classifier's weights to the distribution of kurtosis expected by chance. If the real classifier's weight kurtosis differed from chance, this would indicate either that fewer neurons were strongly involved in the classifier compared to a random classifier, or that neurons contributed more uniformly to the classifier than expected by chance. Neither of these was observed (all $p > 0.1$, two-sided).

Variance Alignment analysis

We initially reduced the dimensionality of the data as above to *k* dimensions (chosen as 8) using PCA. This step de-noised the data. For this analysis, the A matrix on which we performed PCA contained data from both the decision and movement epochs together; this ensured that the resulting space captured the structure of both epochs. We then determined the shape of the variance ellipsoid for the *movement* epoch alone (−200 to 800 ms from movement onset). That is, we rotated the data in the *k*-dimensional space so that the first dimension captured as much movement-epoch variance as possible, the second the next most, etc. This was accomplished using PCA on the (*k*-dimensional) movement-epoch data alone, retaining all components. The *decision* data was then rotated into this movement-determined orientation. For each dimension *d* (1 to *k*, horizontal axis in Fig. 5h,i), we could then determine how much variance was present in the first *d* dimensions of the decision data. These values were normalized by the maximum possible variance that could be captured in the same number of dimensions: if the rotation were found using PCA on the decision-epoch data itself. Perfect alignment would produce a unity Variance Alignment value, while maximal misalignment defines the lower bound (i.e., if the highest-variance dimension during the movement epoch were the lowest-variance dimension during the decision epoch). To determine the chance Variance Alignment, 10,000 randomly oriented orthogonal bases for the *k*-space were chosen. The confidence intervals shown in Figure 5h,i were derived from these random bases (not corrected for multiple comparisons).

To obtain a summary "Variance Alignment index", we computed the area between the data curve and the chance curve. If the data curve was above the chance curve, the index was taken as positive and was normalized by the area between the perfect alignment curve and the chance curve. If the data curve was below the chance curve, the index was taken as negative and was normalized by the area between the chance curve and the maximally misaligned curve. The index thus ranges from −1 (maximally misaligned) to 1 (perfectly aligned). This index was also computed for each of the 10,000 random orientations. The resulting chance distribution was used to calculate a p-value (two-sided).

For a control comparison, we repeated this analysis on two different modality conditions, visual and multisensory, during the decision epoch. This is a useful comparison because if activity patterns during the visual condition and multisensory condition are aligned, it demonstrates that the finding of chance-level alignment during decision formation and movement truly results from misaligned states and not noise. For this analysis, we determined the ordering of dimensions using data from multisensory trials. Then, data from visual trials was rotated into the multisensory-determined orientation. This analysis was performed on rats 1-4; rat 5 was excluded because this animal had some neurons for which multisensory trials were not collected.

To better interpret the result, we asked whether neurons that had strong stimulus-epoch modulation tended to have strong or weak movement-epoch modulation. To measure the depth of modulation for each neuron, we first created a vector containing the trial-averaged firing rate at each time point for each condition. For each neuron, there was one vector for the stimulus epoch (starting 300 ms before stimulus onset) and one vector for the movement epoch (−200 to 800 ms from movement onset). The variance of each vector was then taken. Since the resulting distributions were approximately log-normal, we took the log of these values before correlating them.

Testing for linear and nonlinear components of neurons' responses

We wished to test how much of neurons' tuning was a linear function of choice or stimulus modality, and how much was a function of nonlinear interaction between choice and stimulus modality. To determine this, we considered one neuron at a time, and analyzed only successful visual and auditory trials (multisensory trials and failures were excluded). We first reduced the neuron's response on each trial to a single number: the spike count in a 200 ms window of time (600 to 800 ms after stimulus onset). This produced a spike count vector y with as many elements as included trials. Our model of the neuron's response was:

$$y = X\beta + \eta$$

where X is a design matrix (see below), β is a vector of regression coefficients, and η is a noise term. The design matrix X was of size r by 4, with r the number of trials. Each row summarized the conditions for the corresponding trial. For each row, the first element was always one, to capture the mean across trials. The second element captured choice preference, set to +1 for rightward trials and −1 for leftward trials. The third element captured modality preference, set to +1 for visual trials and −1 for auditory trials. The last element captured the interaction, set to +1 for visual rightward trials and auditory leftward trials, and −1 otherwise.

To find β, we performed a Generalized Linear Model regression (GLM regression, [51], with a Poisson noise distribution (since single-trial spike counts are typically assumed to have Poisson noise). In order to have orthogonal columns of X, it was necessary to have equal numbers of trials for each combination of left and right and visual and auditory. We therefore randomly down-sampled trials to make these groups equal size before performing the regression. This was done 100 times for each neuron, and the resulting β's were averaged. For analysis, we examined the two linear terms of β and compared them with the final, interaction term of β.

Statistics

The statistical tests used to evaluate each measure are listed in the corresponding sections above. A brief summary of the three types of tests is outlined here. First, non-parametric tests (Mann-Whitney U) were used to determine significance of the behavioral effects from inactivation. Second, bootstrap tests were used to measure the significance of the multisensory enhancement, the strength of choice and modality tuning, and decoder performance. Finally, Monte Carlo simulations were used to evaluate the PAIRS analysis

and the Variance Alignment analysis. No statistical tests were done to predetermine sample size. The study was not blinded. A supplementary methods checklist is available.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

We thank Petr Znamenskiy, Santiago Jaramillo and Tony Zador for technical advice, Kathy Rockland and Angela Licata for help with histology, John Sheppard and Michael Ryan for help with electrophysiology and neural clustering, and Matteo Carandini and Antonio Rangel for providing input to early versions of the manuscript. Funding for this work was provided by NIH EY019072 and EY022979, the John Merck Fund, the McKnight Foundation, the Marie Robertson Memorial Fund of Cold Spring Harbor Laboratory, and a Swartz Foundation fellowship.

References

1. Barlow HB. Summation and inhibition in the frog's retina. The Journal of physiology. 1953; 119:69–88. [PubMed: 13035718]
2. Kuffler SW. Discharge patterns and functional organization of mammalian retina. J Neurophysiol. 1953; 16:37–68. [PubMed: 13035466]
3. Britten KH, Newsome WT, Shadlen MN, Celebrini S, Movshon JA. A relationship between behavioral choice and the visual responses of neurons in macaque MT. Visual neuroscience. 1996; 13:87–100. [PubMed: 8730992]
4. Nienborg H, Cumming BG. Macaque V2 neurons, but not V1 neurons, show choice-related activity. J Neurosci. 2006; 26:9567–9578. doi:26/37/9567 [pii] 10.1523/JNEUROSCI.2256-06.2006. [PubMed: 16971541]
5. Roitman JD, Shadlen MN. Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. J Neurosci. 2002; 22:9475–9489. doi:22/21/9475 [pii]. [PubMed: 12417672]
6. Balan PF, Gottlieb J. Functional significance of nonspatial information in monkey lateral intraparietal area. J Neurosci. 2009; 29:8166–8176. doi:29/25/8166 [pii] 10.1523/JNEUROSCI. 0243-09.2009. [PubMed: 19553456]
7. Georgopoulos AP, Kalaska JF, Caminiti R, Massey JT. On the relations between the direction of two-dimensional arm movements and cell discharge in primate motor cortex. J Neurosci. 1982; 2:1527–1537. [PubMed: 7143039]
8. Churchland AK, Kiani R, Shadlen MN. Decision-making with multiple alternatives. Nature neuroscience. 2008; 11:693–702. doi:10.1038/nn.2123. [PubMed: 18488024]
9. Arimura N, Nakayama Y, Yamagata T, Tanji J, Hoshi E. Involvement of the globus pallidus in behavioral goal determination and action specification. J Neurosci. 2013; 33:13639–13653. doi: 10.1523/JNEUROSCI.1620-13.2013. [PubMed: 23966686]
10. Viswanathan P, Nieder A. Neuronal correlates of a visual "sense of number" in primate parietal and prefrontal cortices. Proceedings of the National Academy of Sciences of the United States of America. 2013 doi:10.1073/pnas.1308141110.
11. Roth ED, Yu X, Rao G, Knierim JJ. Functional differences in the backward shifts of CA1 and CA3 place fields in novel and familiar environments. PloS one. 2012; 7:e36035. doi:10.1371/ journal.pone.0036035. [PubMed: 22558316]
12. Ganguli S, Sompolinsky H. Compressed sensing, sparsity, and dimensionality in neuronal information processing and data analysis. Annual review of neuroscience. 2012; 35:485–508. doi: 10.1146/annurev-neuro-062111-150410.
13. Salinas E. Context-dependent selection of visuomotor maps. BMC Neurosci. 2004; 5:47. doi: 10.1186/1471-2202-5-47. [PubMed: 15563737]

14. Pouget A, Sejnowski TJ. Spatial transformations in the parietal cortex using basis functions. Journal of Cognitive Neuroscience. 1997; 9:222–237. doi:10.1162/jocn.1997.9.2.222. [PubMed: 23962013]
15. Rishel CA, Huang G, Freedman DJ. Independent category and spatial encoding in parietal cortex. Neuron. 2013; 77:969–979. doi:10.1016/j.neuron.2013.01.007. [PubMed: 23473325]
16. Meister ML, Hennig JA, Huk AC. Signal multiplexing and single-neuron computations in lateral intraparietal area during decision-making. The Journal of neuroscience. 2013; 33:2254–2267. doi: 10.1523/JNEUROSCI.2984-12.2013. [PubMed: 23392657]
17. Freedman DJ, Assad JA. Distinct encoding of spatial and nonspatial visual information in parietal cortex. J Neurosci. 2009; 29:5671–5680. doi:10.1523/JNEUROSCI.2878-08.2009. [PubMed: 19403833]
18. Rigotti M, et al. The importance of mixed selectivity in complex cognitive tasks. Nature. 2013; 497:585–590. doi:10.1038/nature12160. [PubMed: 23685452]
19. Harvey CD, Coen P, Tank DW. Choice-specific sequences in parietal cortex during a virtual-navigation decision task. Nature. 2012; 484:62–68. doi:10.1038/nature10918. [PubMed: 22419153]
20. Raposo D, Sheppard JP, Schrater PR, Churchland AK. Multisensory decision-making in rats and humans. The Journal of neuroscience. 2012; 32:3726–3735. doi:10.1523/JNEUROSCI. 4998-11.2012. [PubMed: 22423093]
21. Sheppard JP, Raposo D, Churchland AK. Dynamic weighting of multisensory stimuli shapes decision-making in rats and humans. Journal of Vision. 2013; 13 doi:10.1167/13.6.4.
22. Angelaki DE, Gu Y, DeAngelis GC. Multisensory integration: psychophysics, neurophysiology, and computation. Curr Opin Neurobiol. 2009; 19:452–458. doi:S0959-4388(09)00072-5 [pii] 10.1016/j.conb.2009.06.008. [PubMed: 19616425]
23. Li CS, Mazzoni P, Andersen RA. Effect of reversible inactivation of macaque lateral intraparietal area on visual and memory saccades. J Neurophysiol. 1999; 81:1827–1838. [PubMed: 10200217]
24. Glickfeld LL, Histed MH, Maunsell JH. Mouse primary visual cortex is used to detect both orientation and contrast changes. J Neurosci. 2013; 33:19416–19422. doi:10.1523/JNEUROSCI. 3560-13.2013. [PubMed: 24336708]
25. Rogan SC, Roth BL. Remote control of neuronal signaling. Pharmacological reviews. 2011; 63:291–315. doi:10.1124/pr.110.003020. [PubMed: 21415127]
26. Reep RL, Chandler HC, King V, Corwin JV. Rat posterior parietal cortex: topography of corticocortical and thalamic connections. Exp Brain Res. 1994; 100:67–84. [PubMed: 7813654]
27. Brunton BW, Botvinick MM, Brody CD. Rats and humans can optimally accumulate evidence for decision-making. Science. 2013; 340:95–98. doi:10.1126/science.1233912. [PubMed: 23559254]
28. Indyk P, Motwani R. Approximate nearest neighbors: towards removing the curse of dimensionality. Proc. Annu. ACM Symp. Theory Comput. 1998; 30th:604–613.
29. Rust NC, Dicarlo JJ. Selectivity and tolerance (“invariance”) both increase as visual information propagates from cortical area V4 to IT. J Neurosci. 2010; 30:12978–12995. doi:10.1523/ JNEUROSCI.0179-10.2010. [PubMed: 20881116]
30. Cortes C, Vapnik V. Support-vector networks. Machine Learning. 1995; 20:273–297.
31. Kaufman MT, Churchland MM, Ryu SI, Shenoy KV. Cortical activity in the null space: permitting preparation without movement. Nature neuroscience. 2014 doi:10.1038/nn.3643.
32. Russo GS, Backus DA, Ye S, Crutcher MD. Neural activity in monkey dorsal and ventral cingulate motor areas: comparison with the supplementary motor area. J Neurophysiol. 2002; 88:2612–2629. doi:10.1152/jn.00306.2002. [PubMed: 12424298]
33. Segraves MA. Activity of monkey frontal eye field neurons projecting to oculomotor regions of the pons. J Neurophysiol. 1992; 68:1967–1985. [PubMed: 1491252]
34. Movshon JA, Newsome WT. Visual response properties of striate cortical neurons projecting to area MT in macaque monkeys. J Neurosci. 1996; 16:7733–7741. [PubMed: 8922429]
35. Chen JL, Carta S, Soldado-Magraner J, Schneider BL, Helmchen F. Behaviour-dependent recruitment of long-range projection neurons in somatosensory cortex. Nature. 2013; 499:336–340. doi:10.1038/nature12236. [PubMed: 23792559]

36. Churchland AK, Lisberger SG. Discharge properties of MST neurons that project to the frontal pursuit area in macaque monkeys. J Neurophysiol. 2005; 94:1084–1090. doi:Doi 10.1152/Jn.00196.2005. [PubMed: 15872067]
37. El-Shamayleh Y, Kumbhani RD, Dhruv NT, Movshon JA. Visual response properties of V1 neurons projecting to V2 in macaque. J Neurosci. 2013; 33:16594–16605. doi:10.1523/JNEUROSCI.2753-13.2013. [PubMed: 24133263]
38. Pare M, Wurtz RH. Progression in neuronal processing for saccadic eye movements from parietal cortex area lip to superior colliculus. J Neurophysiol. 2001; 85:2545–2562. [PubMed: 11387400]
39. Sussillo D, Abbott LF. Generating coherent patterns of activity from chaotic neural networks. Neuron. 2009; 63:544–557. doi:10.1016/j.neuron.2009.07.018. [PubMed: 19709635]
40. Pagan M, Urban LS, Wohl MP, Rust NC. Signals in inferotemporal and perirhinal cortex suggest an untangling of visual target information. Nature neuroscience. 2013; 16:1132–1139. doi:10.1038/nn.3433. [PubMed: 23792943]
41. Ghazanfar AA, Maier JX, Hoffman KL, Logothetis NK. Multisensory integration of dynamic faces and voices in rhesus monkey auditory cortex. J Neurosci. 2005; 25:5004–5012. doi:10.1523/JNEUROSCI.0799-05.2005. [PubMed: 15901781]
42. Linden JF, Grunewald A, Andersen RA. Responses to auditory stimuli in macaque lateral intraparietal area. II. Behavioral modulation. J Neurophysiol. 1999; 82:343–358. [PubMed: 10400963]
43. Mullette-Gillman OA, Cohen YE, Groh JM. Eye-centered, head-centered, and complex coding of visual and auditory targets in the intraparietal sulcus. J Neurophysiol. 2005; 94:2331–2352. [PubMed: 15843485]
44. Kiani R, Shadlen MN. Representation of confidence associated with a decision by neurons in the parietal cortex. Science. 2009; 324:759–764. doi:324/5928/759 [pii] 10.1126/science.1169405. [PubMed: 19423820]
45. Shams L, Kamitani Y, Shimojo S. Visual illusion induced by sound. Brain Res Cogn Brain Res. 2002; 14:147–152. doi:S0926641002000691 [pii]. [PubMed: 12063138]
46. Carandini M, Churchland AK. Probing perceptual decisions in rodents. Nature neuroscience. 2013; 16:824–831. doi:10.1038/nn.3410. [PubMed: 23799475]
47. Wichmann FA, Hill NJ. The psychometric function: II. Bootstrap-based confidence intervals and sampling. Percept Psychophys. 2001; 63:1314–1329. [PubMed: 11800459]
48. Feierstein CE, Quirk MC, Uchida N, Sosulski DL, Mainen ZF. Representation of spatial goals in rat orbitofrontal cortex. Neuron. 2006; 51:495–507. doi:S0896-6273(06)00513-7 [pii] 10.1016/j.neuron.2006.06.032. [PubMed: 16908414]
49. Erlich JC, Bialek M, Brody CD. A cortical substrate for memory-guided orienting in the rat. Neuron. 2011; 72:330–343. doi:S0896-6273(11)00607-6 [pii] 10.1016/j.neuron.2011.07.010. [PubMed: 22017991]
50. Breiman L. Bagging predictors. Machine learning. 1996; 26 doi:10.1023/A:1018054314350.
51. Nelder JA, Wedderburn RW. Generalized linear models. Journal of the Royal Statistical Society, Series A (Royal Statistical Society). 1972; 135:370–384. doi:10.2307/2344614.

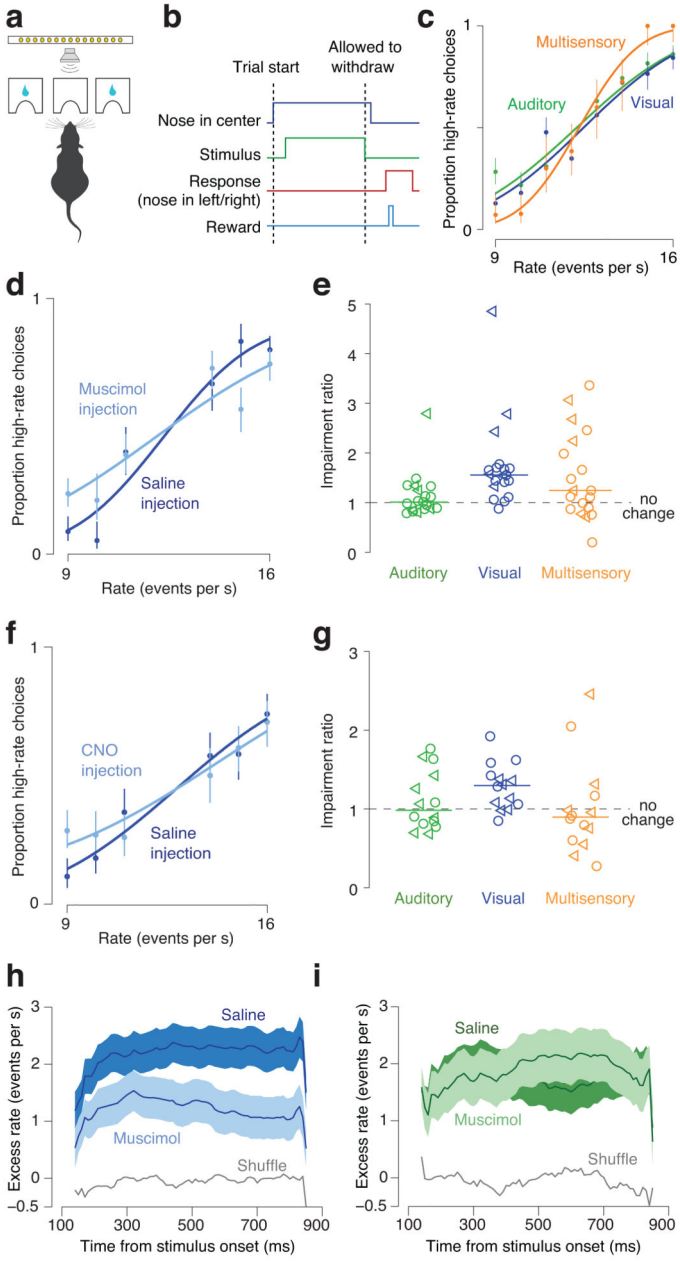

Figure 1. Multisensory decision-making in rodents: behavior and inactivation

a, Schematic drawing of rat in behavioral apparatus. Visual stimuli were presented via a panel of diffused LEDs (top); auditory stimuli were presented via a centrally positioned speaker. **b,** Schematic of task timeline. **c,** Example data (818 trials; 1 session) from a single animal (rat 4). Smooth lines are fits (cumulative Gaussian). Error bars reflect the Wilson binomial confidence interval. **d-e,** Effects of muscimol inactivation. **d,** Example psychometric functions for 1 animal (visual trials only). Dark blue: a single session following saline injection. Light blue: a single session the next day following muscimol injection. **e,** Effects of inactivation on performance for auditory (green), visual (blue) and multisensory (orange) trials. Ordinate: impairment ratio: the ratio of values for σ parameter

from cumulative Gaussian fit to the data (Online Methods). A value of 1 indicates no effect; values > 1 indicate performance was worse on a single inactivation session (muscimol) relative to the previous control session (saline). Symbols: individual animals (N = 2); horizontal lines: median across animals and sessions. **f,g,** Same as d,e but for separate inactivation experiments implemented with DREADD (N = 2). **h-i,** Effect of stimulus on decisions. **h,** Excess rate is higher for a single rat on visual trials with saline (dark blue) vs. muscimol (light blue) sessions. Values on abscissa: centers of sliding windows. Shaded regions: confidence bounds (mean ± s.e.m.). **i,** Excess rate for the same rat on auditory trials with saline (dark green) and muscimol (light green).

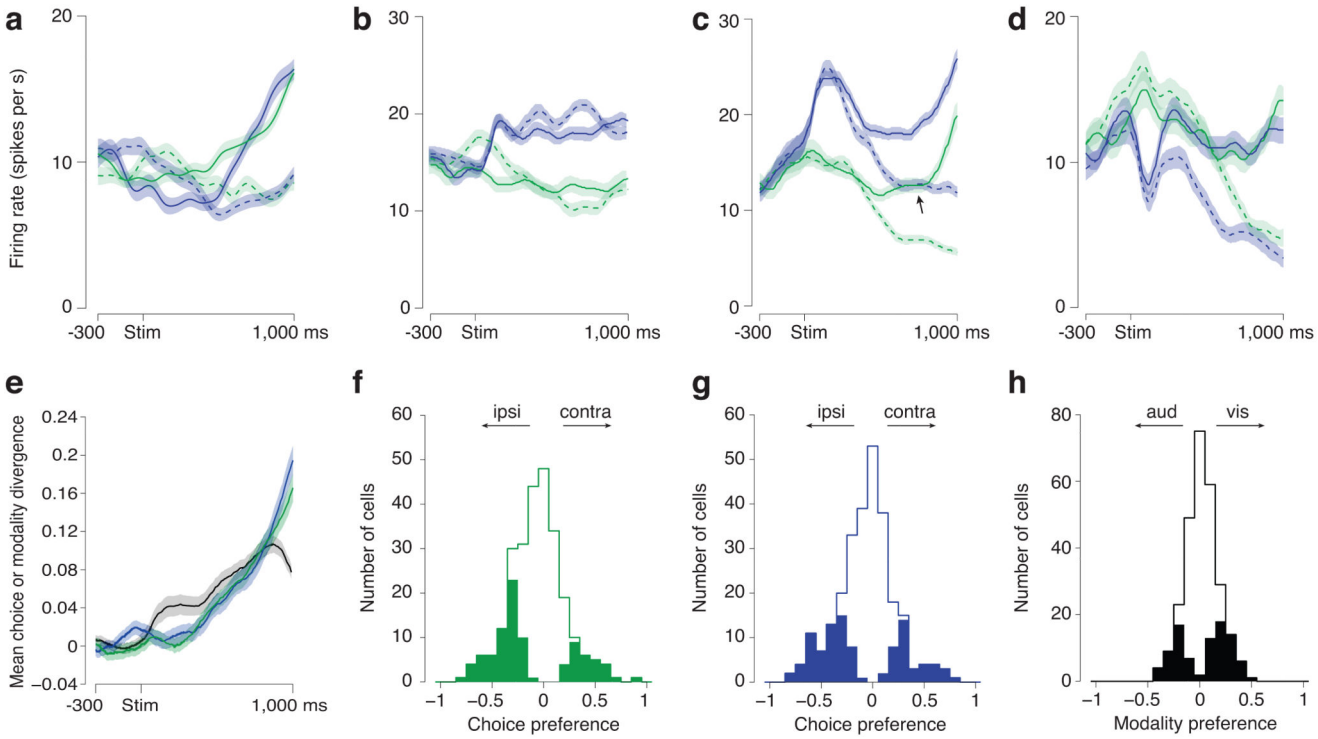

Figure 2. PPC neurons show mixed selectivity for choice and modality

Plots display visual/auditory trials; for multisensory trials, see Supplementary Figure 4-5. **a-d,** Peri-stimulus time histograms for four single neurons. Mean spike counts were computed in 10 ms time windows smoothed with a Gaussian (σ = 50 ms). Error trials were excluded. Trials grouped by stimulus rate. Solid line: low rate stimulus; dashed line: high rate stimulus. Color indicates modality. Transparent fills: s.e.m. Responses aligned to the time the visual or auditory stimulus began ("Stim"). **a,** Neuron reflects mainly categorical choice (392 trials). **b,** Neuron reflects mainly stimulus modality (414 trials). **c,** Neuron mixes categorical choice and modality (586 trials). Arrow highlights ambiguous moment in which high rate visual and low rate auditory stimuli gave rise to the same firing rate. **d,** Neuron mixes categorical choice and modality and displays complex temporal dynamics (440 trials). **e,** Choice divergence (Online Methods) for auditory trials (green; average of 262 neurons) and visual trials (blue; average of 268 neurons), and modality divergence (black; average of 266 neurons). Transparent fills: s.e.m. (bootstrap). **f,** Histogram of choice preference for auditory trials, measured 200 ms before decision end. Filled bars indicate neurons where index was significantly different from 0 ($p < 0.01$, 1000 bootstraps). **g,** Same as (f) but for visual trials. **h,** Same as (f-g), but for modality.

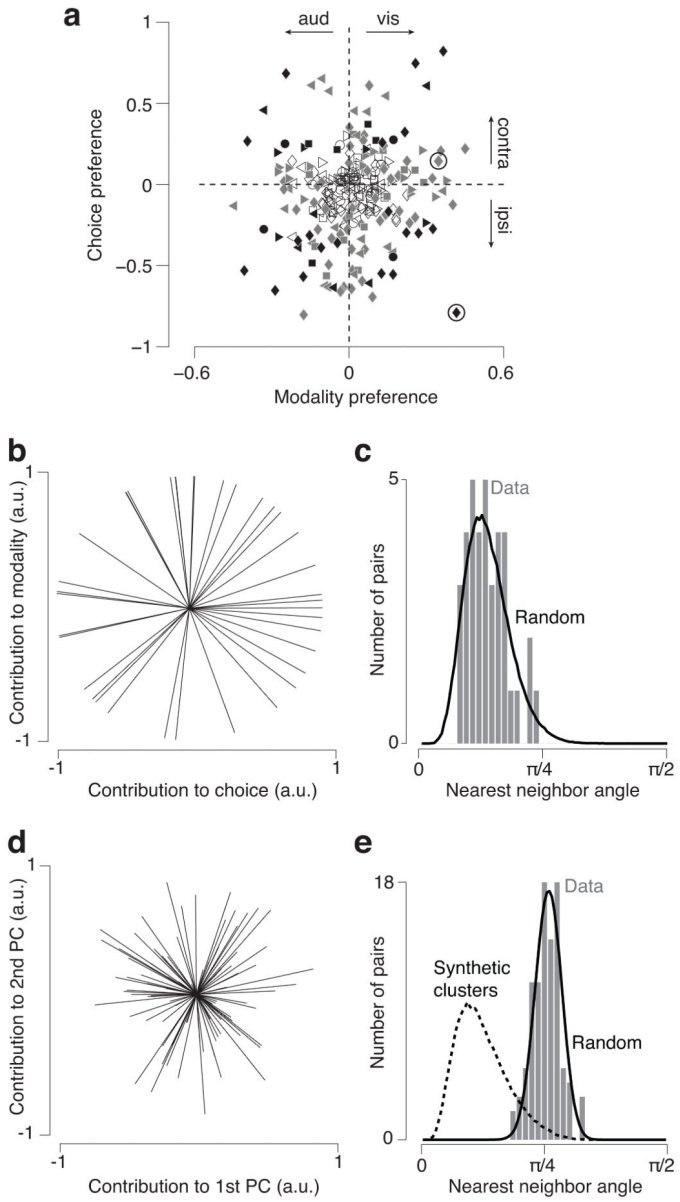

Figure 3. Neural responses defy categorization

a, Choice and modality preferences are unrelated. Each point shows values for a single neuron. Abscissa: average modality preference for high- and low-rate trials. Ordinate: choice preference for visual trials. A nearly identical outcome was achieved when choice preference was computed from auditory trials (data not shown). Shading indicates significance: open: neither choice nor modality preference was significant; gray: choice or modality preference was significant; black: choice and modality preference were significant. Top circled point shows modality-selective neuron in panel (2b); lower circled point shows mixed selectivity neuron in panel (2c). Dashed lines define a region along which neurons would tend to cluster if they had pure selectivity for choice or modality. **b**, Low-dimensional summaries of neurons' responses. Each line is a "feature vector" showing the degree to

which each neuron contributes to choice and modality. **c**, Neurons are rarely more similar to one another than expected by chance. Histogram shows the distribution of angles between each neuron and its nearest neighbors (N = 45 neurons, Rat 5). Black line: distribution of nearest-neighbor angles for random 2-D vectors. **d,** Low-dimensional summaries of neurons' responses. Responses are shown projected into 2 dimensions of 8 used total. Each vector shows the contribution of one neuron to the two dimensions. The apparently random distribution of vectors suggests that neurons do not tend to cluster. **e,** Neurons are rarely more similar to one another than expected by chance. Histogram shows the distribution of angles between each neuron and its nearest neighbors (N = 77 neurons; Rat 1). Black line: distribution of nearest-neighbor angles for random 8-D vectors. Dashed: distribution of nearest-neighbor angles is left-shifted when "synthetic clusters" are introduced (Online Methods).

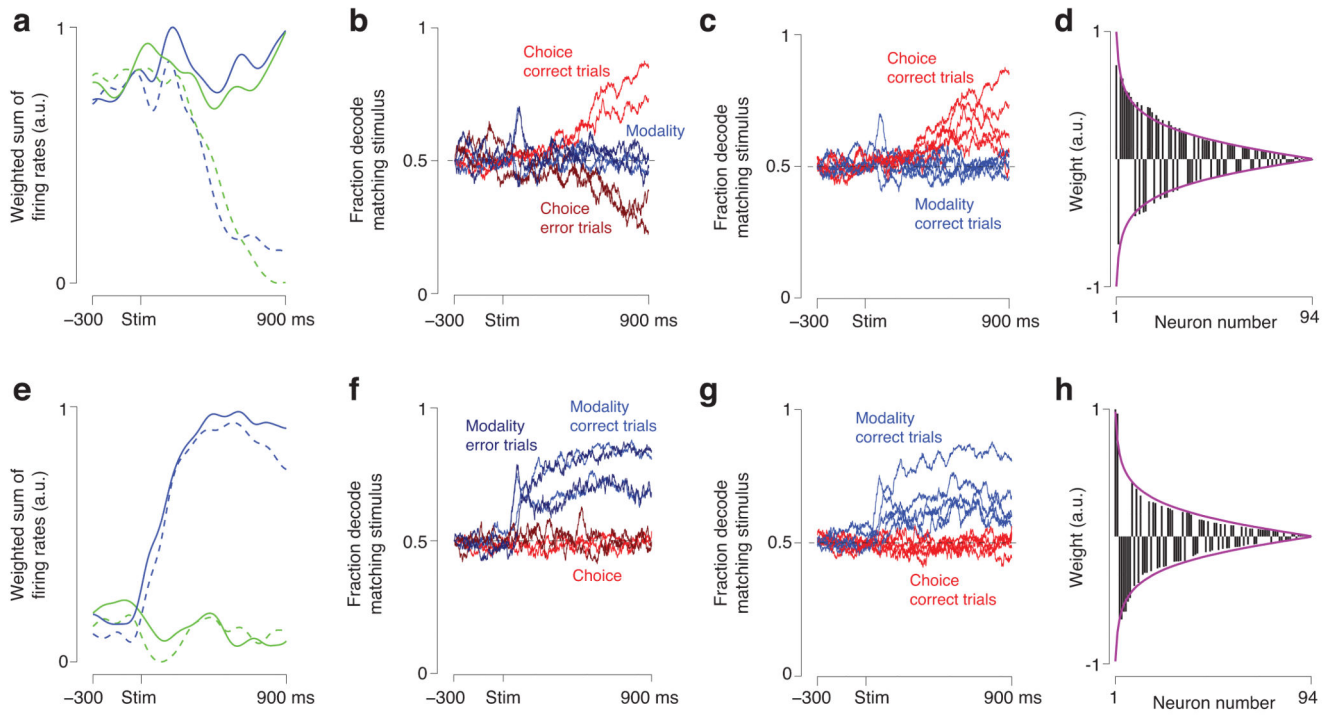

Figure 4. Choice and modality can be decoded from population activity

a, Weighted sums of neural responses; weights were chosen by the classifier. Blue, visual; green, auditory; dashed lines: high-rate trials; solid lines: low-rate trials. Data from Rat 4, N = 94 neurons. **b,** The choice decoder could correctly classify responses as left vs. right on trials where the rat was successful (bright red traces, one per rat), but is at chance for auditory vs. visual (blue traces). On trials where the rat chose the incorrect port, the decode tracked the rat's choice (brown traces). Traces reflect the average of 1000 classifications. **c,** Same as (b) for all 5 rats, correct trials only. Each animal has one trace for modality and one for choice. **d,** Bars: values of the weights used to generate the traces in (a), ordered by magnitude. Purple lines: values of randomly generated 94-dimensional vectors ordered by magnitude. **e-h,** Same as a-d but for the modality decoder. The modality decode was nearly identical whether the rats chose the correct or incorrect port (f, bright vs. dark blue).

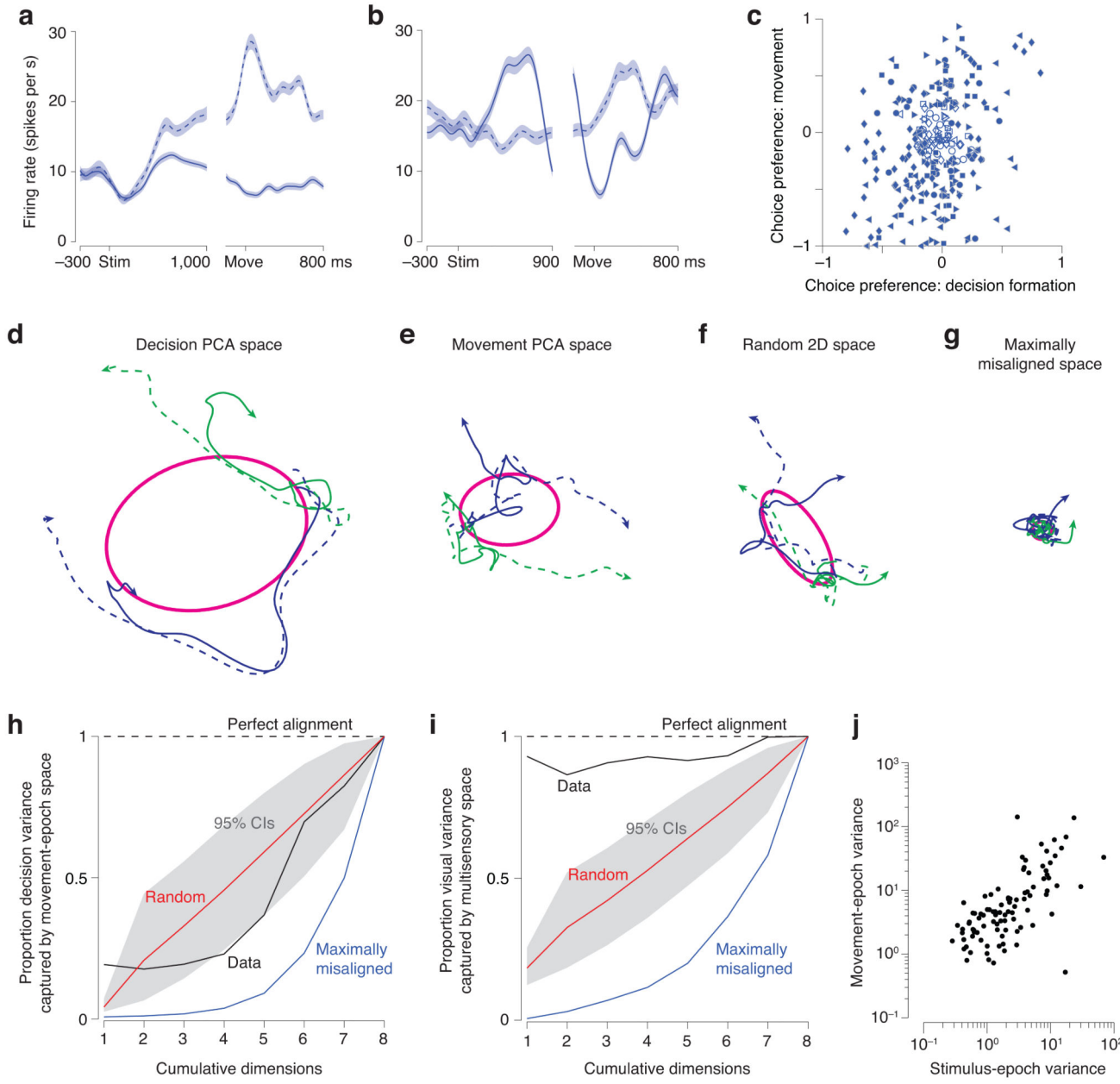

Figure 5. PPC neurons exhibit different covariance patterns during decision formation and movement

a, Example neuron with a sustained preference for high rate stimuli (dashed line above solid line). Responses were aligned to stimulus onset (left) and to movement onset (right). Alignment to these two events was necessary because the time between the stimulus end and the animal's movement varied slightly from trial to trial. Traces reflect averaged responses of all correct visual trials (and s.e.m.) computed as in Figure 2a-d. **b,** Example neuron that switched its preference over the course of the trial. **c,** Choice preference during decision formation (200 ms before decision end, abscissa) and movement (200 ms after animal leaves choice port, ordinate), frequently differed but were nonetheless correlated across all cells; N

= 268 neurons, r = 0.302, p < 0.001. Symbols: individual animals. **d-g,** Two-dimensional projections of decision-epoch data (same data for all panels). Magenta ellipses indicate 1 s.d. of the data projected into the space specified by the panel's title. Color/linestyle are the same as in Figure 2. **d,** Space chosen as first two PCs of decision-epoch data. **e,** Space chosen as first two PCs of movement-epoch data. **f,** Space chosen randomly from top 8 PCs. **g,** Space chosen as PCs 7-8 of decision-epoch data. **h,** Variance Alignment analysis indicates that activity patterns across neurons differ substantially (and are not just, e.g., scaled) for decision formation vs. movement. Ordinate: decision variance captured in the dimensions used during movement, normalized by how much could maximally be captured in the same number of dimensions. Abscissa: cumulative dimensions included. Black trace: data; other traces: alignment values expected under several scenarios for comparison (see labels). **i,** Same as (h) except that ordinate shows amount of decision variance for visual trials captured in the dimensions used on multisensory trials, normalized by how much could maximally be captured in the same number of dimensions. Note that unlike (d), the black "data" trace is close to 1, as expected. **j,** Strength of stimulus modulation for each neuron correlates with strength of movement modulation. Data in (d-j) from Rat 4.