Supplemental Table 1. Receptive field properties across V4 and vlPFC sites, including locations of mapped RF centers and estimated widths of fitted RFs.

	Distribution of RF centers in visual degrees (°), reported as (horizontal, vertical) at 25th, 50th, and 75th percentiles	Bootstrapped median RF center ± SEM (n = 1,000)	Observed mean RF center ± SEM	Gaussian width (pooled across X & Y, 25th, 50th, 75th percentiles)	Bootstrapped median width ± SEM (pooled, n = 1,000)	Observed mean ± SEM width (pooled)
V4 Monkey C	(9.8, -3.7)º (10.1, -3.2)º (10.4, -2.7)º	(10.1, -3.2)º ± (0.003, 0.005)º	(10.1, -3.3)º ± (0.09, 0.1)º	4.5º 4.7º 5.0º	4.7º ± 0.001º	4.8º ± 0.05º
vlPFC Monkey C	(7.0, 11.0)º (13.7, 11.5)º (17.6, 12.7)º	(12.7, 11.5)º ± (0.1, 0.03)º	(11.2, 11.3)º ± (3.8, 1.0)º	4.9º 7.4º 12.2º	7.7º ± 0.07º	8.9º ± 1.5º
vlPFC Monkey D	(7.8, -12.9)º (10.2, -11.8)º (12.4, -8.0)º	(10.2, -11.3)º ± (0.04, 0.05)º	(10.0, -10.9)º ± (1.2, 0.8)º	4.6º 7.1º 9.4º	7.2º ± 0.02º	7.3º ± 0.6º

Supplemental Table 2. Visual selectivity across the full population of V4, vlPFC, and CPB sites. Statistical significance was tested using a one-way analysis of variance (ANOVA, with stimulus position as the main factor, two-sided test). P-values were adjusted via false discovery rate.

	Total recordings (32 unique sites/array)	**No. with RFs within 4.0° of stimulus center**	**No. selective at P<0.04**	**No. selective at P<0.04**	**No. selective at P<0.04**
			Baseline window r/t image onset: -149–0 ms	Baseline window r/t image onset: none	Baseline window r/t image onset: 0–50 ms
Monkey C					
V4	**1568**	**208**	88.9%	90.4%	89.4%
vlPFC	**1425**	**79**	12.7%	21.5%	13.9%
Monkey D					
CPB	**1198**	**32**	0.0%	0.0%	0.0%
vlPFC	**1252**	**139**	23.0%	35.3%	21.6%

Supplemental Table 3. Prototype synthesis success rates across V4 and vlPFC sites (significance obtained through Wilcoxon rank sum test), as well as median change in firing rate and number of generations to reach half of a given site's peak mean firing rate.

XDream prototype synthesis across monkeys

	Percent successful experiments (where final synthetic images evoked higher responses than early textures)	Percent experiments where final synthesized images evoked higher responses than the reference images	Median change in firing rate (bootstrapped median ± SEM, n = 1,000 iterations), across **all XDream experiments** adaptive synthetic images, fixed natural images (number of experiments)	Median change in firing rate (bootstrapped median ± SEM, n = 1,000 iterations), across only **successful experiments** adaptive synthetic images, fixed natural images (number of experiments)	Number of generations to reach half-maximum mean response all experiments, only successful experiments Mean number of generations for experiment (mean ± SEM)
V4 Monkey C	79.6% (43 experiments)	75.9% (41 experiments)	20.21 ± 0.10 spikes/s, -8.58 ± 0.06 spikes/s (54 experiments)	25.68 ± 0.10 spikes/s, -6.86 ± 0.06 spikes/s (43 experiments)	13.8 ± 1.3, 13.6 ± 1.3 41.5 ± 2.0
vlPFC Monkey C	10.8% (7 experiments)	6.2% (4 experiments)	-0.27 ± 0.03 spikes/s, -1.15 ± 0.04 spikes/s (65 experiments)	16.56 ± 0.08 spikes/s, -7.99 ± 0.08 spikes/s (7 experiments)	19.1 ± 2.3, 10.4 ± 1.7 36 ± 1.4
vlPFC Monkey D	33.8% (27 experiments)	31.3% (25 experiments)	1.76 ± 0.03 spikes/s, -6.64 ± 0.05 spikes/s (80 experiments)	12.55 ± 0.07 spikes/s -5.45 ± 0.07 spikes/s (27 experiments)	21.3 ± 2.6, 17.1 ± 2.4 42.4 ± 1.6

Supplemental Table 4. All prototype synthesis experiments targeting vlPFC sites in Monkeys C and D. Significance obtained through Wilcoxon rank sum test (two-sided), with threshold of $P < 0.01$. Individual unit types are designated as follows: single unit (SU), distinct multiunit (MU), and whole-channel multiunit hash ("hash").

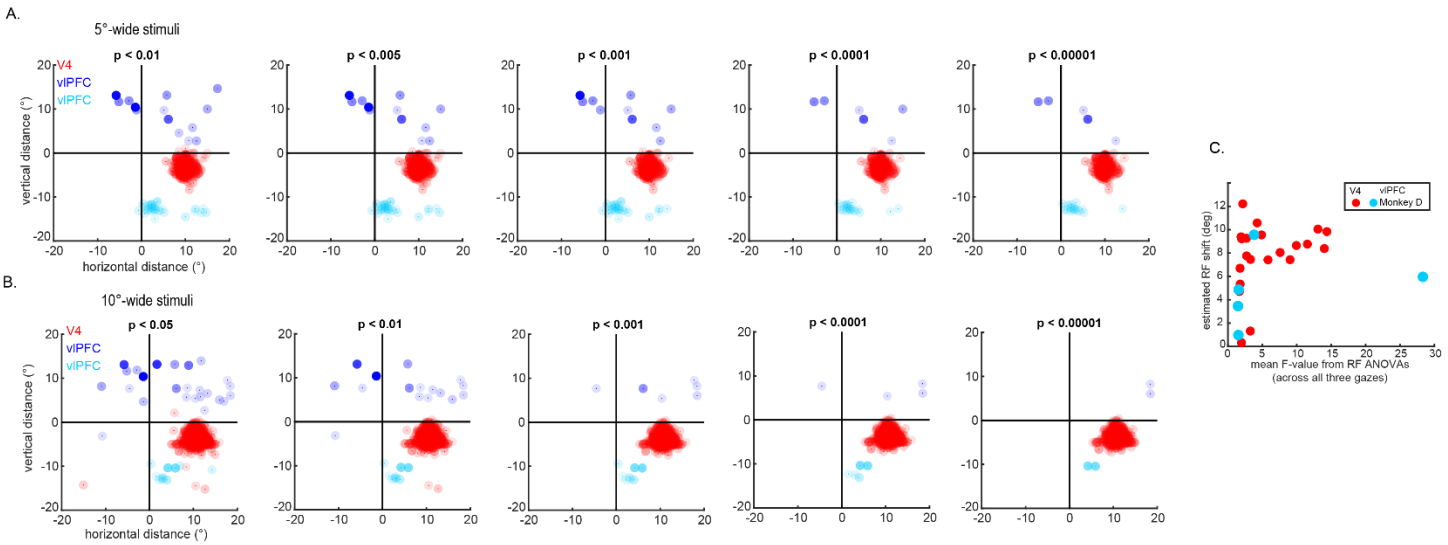
XDream success by site, **p < 0.01. *** = within RF cluster
Monkey C

Ch. number	nExp attempted	nExp successful	Success rate	Successful units	All units tested
Ch. 33*	2	0			1 MU, 1 hash
Ch. 34*	3	0			3 hash
Ch. 35*	7	3	42.9%	2 MU, 1 hash	2 distinct MU, 5 hash
Ch. 42*	5	2	40%	2 hash	5 hash
Ch. 43*	6	1	16.7%	1 hash	2 MU, 4 hash
Ch. 50*	2	0			2 well-isolated SU
Ch. 52	1	0			1 hash
Ch. 57*	10	1	10%	1 MU	5 MU, 5 hash
Ch. 58*	30	20	66.7%	20 MU	23 distinct MU, 7 hash
Ch. 59*	**6**	**0**			5 MU, 1 hash
Ch. 60*	**8**	**0**			6 MU, 2 hash

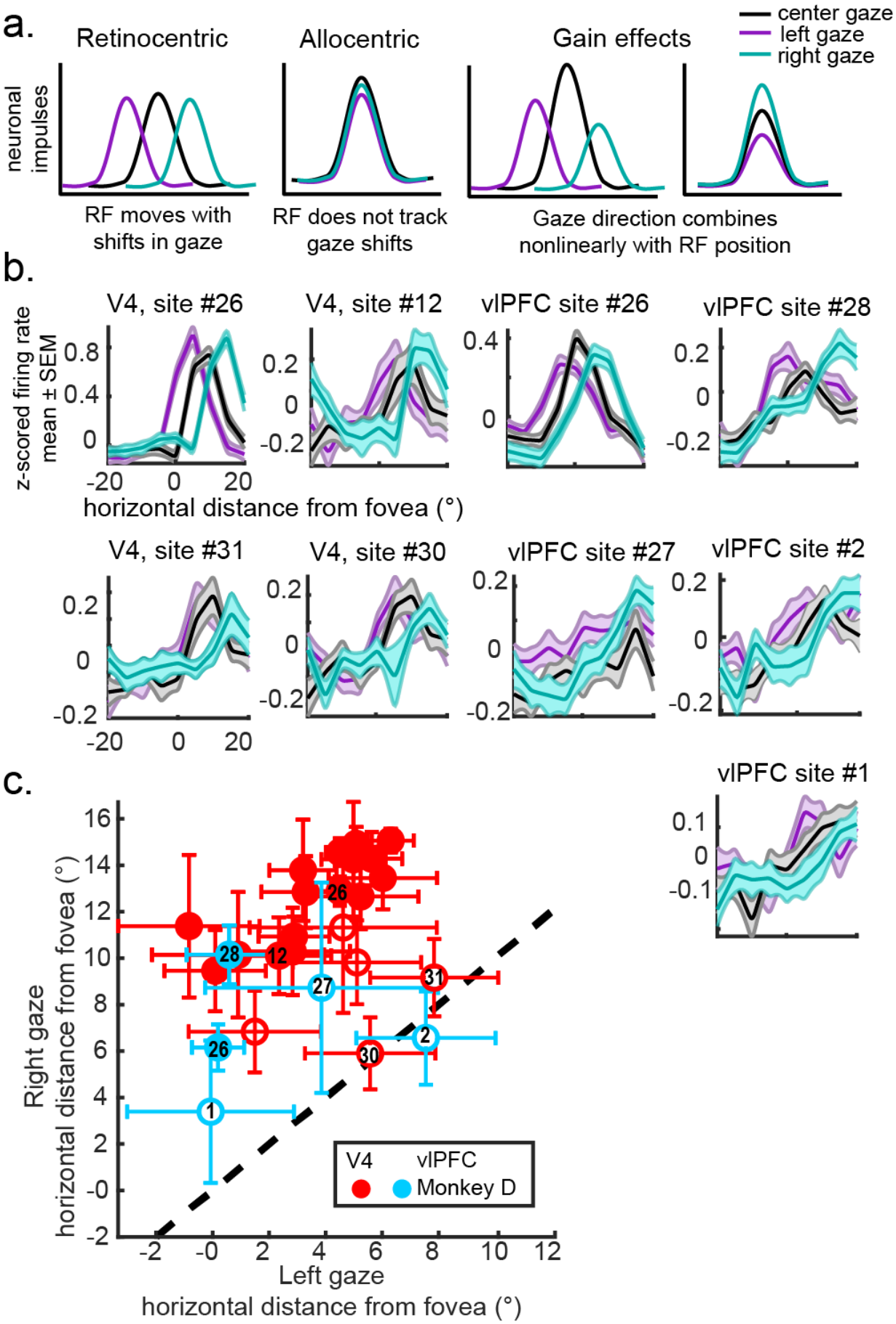
Monkey D

Ch. number	nExp attempted	nExp successful	Success rate	Successful units	All units tested
Ch. 33*	2	0			1 MU, 1 hash
Ch. 34*	3	0			3 hash
Ch. 35*	7	3	42.9%	2 MU, 1 hash	2 distinct MU, 5 hash
Ch. 42*	5	2	40%	2 hash	5 hash
Ch. 43*	6	1	16.7%	1 hash	2 MU, 4 hash
Ch. 50*	2	0			2 well-isolated SU
Ch. 52	1	0			1 hash
Ch. 57*	10	1	10%	1 MU	5 MU, 5 hash
Ch. 58*	30	20	66.7%	20 MU	23 distinct MU, 7 hash

Ch. 59*	**6**	**0**	5 MU, 1 hash
Ch. 60*	**8**	**0**	6 MU, 2 hash



Supplemental Figure 1. Receptive field per different threshold values. Plots show the location of putative RFs (red = V4, blue = vlPFC) as a function of *p*-value thresholds (obtained via a one-way ANOVA, two-tailed, with position as the sole factor), when mapping stimulus widths were 5° (**A**) or 10° **(B)**. **C.** Mean F-value for RF strength as a function of shift in RF centers across different gaze conditions. Each dot shows a given site's measured change in RF center across gaze conditions, plotted against the F-value of their ANOVA test (one-way, stimulus position as factor). This F-value was used as a measure of signal quality. The fact that shifts were generally observed for sites with higher F-values suggest a relationship with signal quality.
All are based on N = 32 channels per array (N = 128 channels total, V4, vlPFC, and CPB).



Supplemental Figure 2. RF spatial coding scheme. We tested if neurons in vlPFC changed their receptive field (RF) position based on gaze direction. (*a*) Schematic of hypothesized RF coding

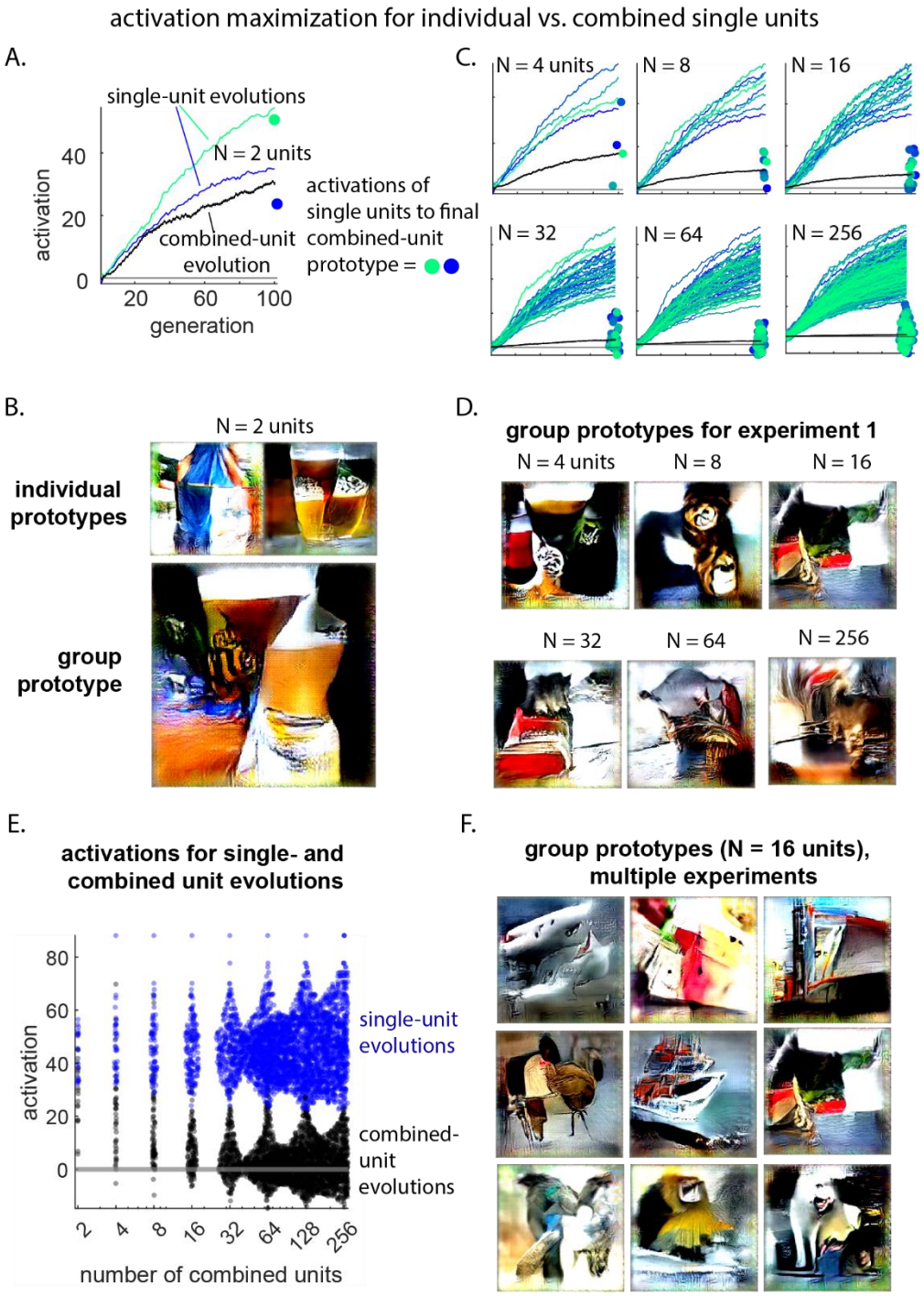
schemes that could be observed, plotted with respect to the display monitor. Responses collected while the monkey fixated at a central fixation point are plotted in black; responses during left- and right-of-center gaze conditions are plotted in purple and teal, respectively. Retinocentric RFs would appear to move with shifts in gaze, while allocentric RFs would consistently signal the same position on the monitor, regardless of stimulating different relative positions on the monkey's retina. For each site, we computed the median shift in RF centers in monitor-centered coordinates. In these experiments, both animals underwent RF mapping (N = 5 experiments), but were required to keep their gaze in one of three locations ("gaze directions"): center (as in the previous RF task), 5° left of center, and 5° right of center (**Figure 1d**). The same absolute positions on the monitor were tested across conditions, so the shift in gaze direction meant that the same position on the monitor tested different relative positions on the retina.
(*b*) Examples of RF coding schemes observed across V4 and vlPFC units, plotted on monitor-centered coordinates. Same color scheme as in (*a*). (*c*) Shift in **estimated RF center for left- and right-gaze conditions,** plotted in monitor-centered coordinates. Retinocentric RFs should fall along the upper left of the diagonal (dashed black line). Red points indicate V4 sites, with blue points indicating vlPFC sites (light blue for Monkey D; no active vlPFC sites in Monkey C during these sessions). Solid markers denote sites that show a significant shift in RF center that follow shifts in gaze, while open markers signal sites without a significant change in RF center. Analyzing responses of sites across days, we applied a two-way ANOVA analysis (two-tailed) for each site (with factors of stimulus position and gaze position). In V4, 38 site-days showed a statistically reliable RF, and of these, 65.8% showed an interaction between RF and gaze V4 ($P < 1x10^{-7}$). In CPB, no sites showed statistically significant RFs. The remaining PFC array showed 10 site-days with robust RFs (also $P < 1x10^{-7}$), and of these 50% showed an interaction with gaze position; these values decreased as the P-value threshold was loosened, down to 30% of PFC sites with $P = 0.06$. Overall, we interpret these results as suggesting the most robust RFs in V4 and PFC were retinocentric, that is, anchored to the position of the fovea. These experiments were performed late in the lifetime of the arrays, so only vlPFC sites from one monkey passed criteria for examination (see Methods), and we could not obtain viable RFs from the second monkey's vlPFC arrays, only from its V4 array. These results are proffered as a proof-of-concept only.

Retinotopic coding analysis (site-by-site). In addition to the population result, we performed further analyses on a site-by-site basis. Above, we described that 24/32 V4 sites showed a robust RF at the center-gaze position (0,0)° and performed our analyses using data only from the *left-* and *right-*gaze conditions [(-5,0)°, (5,0)°]. Here, we focused on the sites that showed very robust RFs in every gaze condition. Out of the 24 active V4 sites noted above, eight V4 sites met this criterion. They demonstrated no significant shifts in estimated RF center while plotted in retinocentric coordinates ($P > 0.05$ after correction by FDR) and a significant shift in estimated RF center when plotted in allocentric coordinates ($p \leq 0.05$, $F \geq 2.0$; F-value range of 7.2 to 85.2). An additional five V4 sites met retinocentric criteria when including sites that, in allocentric coordinates, had a reliable effect size ($F \geq 2.0$), despite having a *P*-value that failed to reject the null hypothesis ($P > 0.05$); for these three sites, the ANOVA F-values ranged from 2.0 to 5.6, while the FDR-corrected *P*-values ranged from 0.06 to 0.30.

There were four vlPFC sites in Monkey C that had a significant RF during center gaze, but none met all criteria for retinocentric RFs. Two of the four vlPFC sites in Monkey C had no significant differences from the retinocentric permutation test and an F-value greater than 2.0 from the allocentric ANOVA (*P*-values were 0.06 and 0.29 after FDR; F-values were 5.7 and 2.4, respectively), suggesting a reliable effect size. In Monkey D, two of the 13 active vlPFC sites met the retinocentric criteria ($P > 0.05$ for retinal permutation test, $P \leq 0.05$ for allocentric ANOVA, $F \geq 2.0$; F-value ranged from 15.8 to 19.8). One additional vlPFC site in Monkey D met the retinocentric criteria in the permutation test ($P > 0.05$) and in having a meaningful effect size from the allocentric ANOVA ($F = 2.5$), though this site failed to reject the null hypothesis under allocentric coordinates ($P = 0.51$).

When considering all sites that met the retinocentric criteria and had allocentric F-values larger than 2.0, there was a total of 13 V4 sites in Monkey C and five vlPFC sites across Monkeys C and D that exhibited RFs on retinocentric coordinates. We did not observe any sites in either vlPFC or V4 that demonstrated allocentric RFs, as defined by the inverse of this trend: a significant difference ($P \leq 0.05$) in the retinocentric condition and no significant differences ($P > 0.05$) in the allocentric condition; however, though we did not observe them, we cannot rule out the possibility that non-retinocentric RFs could exist elsewhere in vlPFC.

Gain effect of gaze direction. Sites exhibiting retinocentric RFs were also tested to determine if there appeared to be a gain effect of gaze direction on response amplitude. None of the 13 V4 sites showed a significant difference in response amplitude as a function of gaze direction, although one did demonstrate an *F*-value of 2.0 ($P = 0.98$ after FDR). None of the five vlPFC sites across subjects demonstrated a gain effect of gaze direction ($P > 0.05$, *F*-values ranged from 0.07 to 0.49). We thus do not report the presence of gain effects of gaze direction in vlPFC RFs, consistent with the general lack of this trend in V4 and ventral stream RFs.

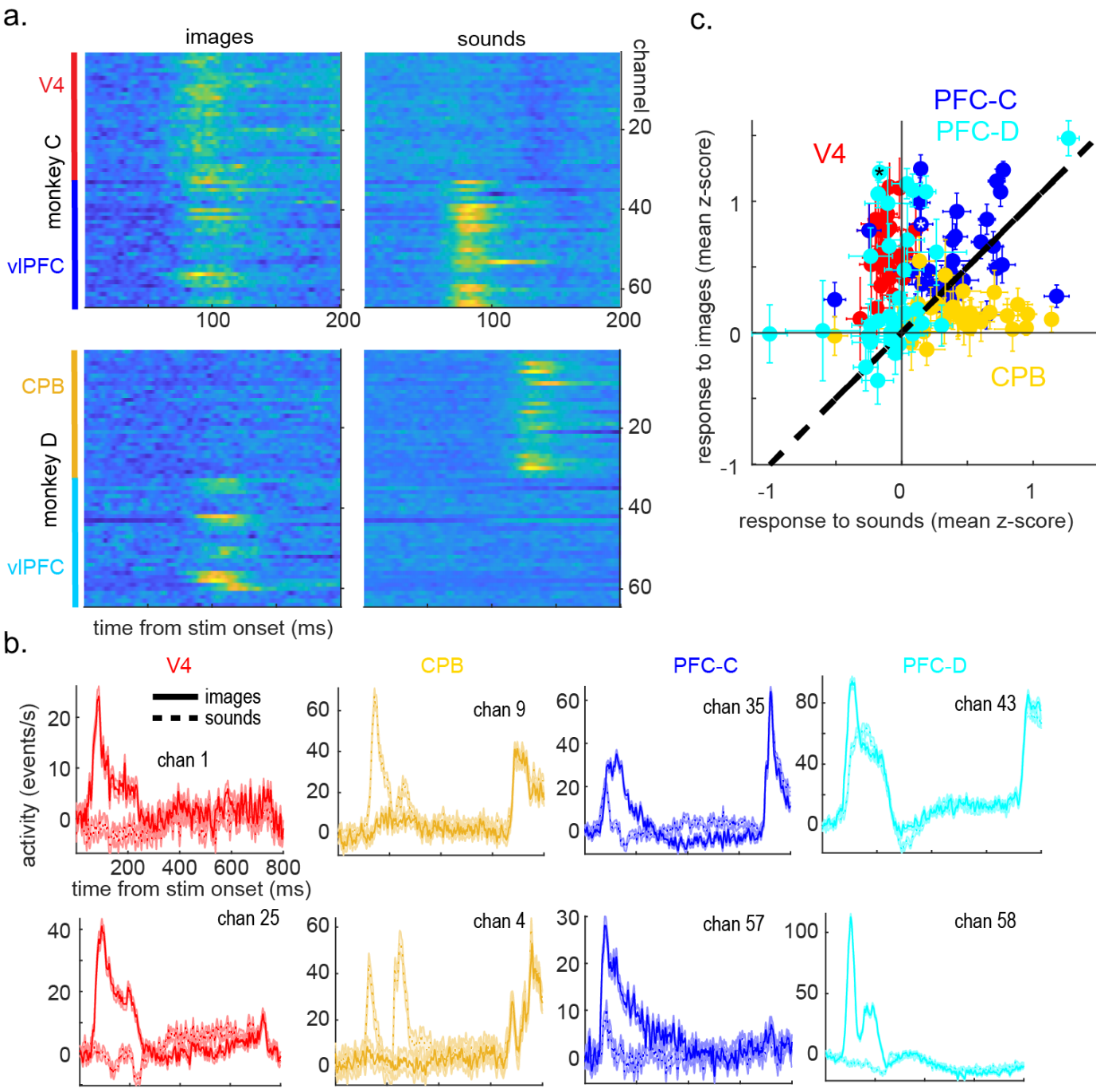


Supplemental Figure 3. Combining single units during the activation maximization process. A. Activity of two units, as images are optimized just for each (color traces) or as images are optimized for both units concurrently (black line), over 100 iterations (generations). Dots show the final activation of the unit to the *combined-unit* prototype. **B.** The singly optimized images (*individual prototypes*) and the combined-unit (*group*) prototype. **C.** As in A but combining more units. **D.** Combined-unit prototypes for larger groups. **E.** Results aggregating multiple experiments, showing the activity (y-axis) of 256 units when images are optimized singly (blue dots), or when they are combined in increasing numbers (black dots, x-axis). **F.** Group prototypes across experiments.

We considered the possibility that the conglomerated activity of many neurons could lead to a more general/amorphous prototype, particularly if the individual neurons comprising the signal had specific and different tuning profiles. Our working hypothesis is that evolutions based on multiunit

activity work largely (perhaps only) when the subjacent single units are similarly tuned. We have made relevant observations over several years. One of our previous studies showed that in evolutions involving a single electrode, single- and multiunit signals were strongly correlated in their tuning (Wang and Ponce, 2022, Tuning landscapes of the ventral stream). We have also tried to maximize the activity of different channels in monkeys B and A, with no interesting results. However, to determine if our closed-loop activation-maximization paradigm works when units are tuned for different images, we conducted a simulation using artificial neural networks (ANNs). We used AlexNet layer fc8 (N = 1000); we defined a range of single (hidden) unit group sizes (G = powers of 2 from 2 to 256). For each experiment Ei where i = 1 to 10, we selected G units (starting with G = 2), then started the evolution process, creating synthetic stimuli that maximize the averaged activation of the selected unit group. We then repeated for G=2, G=4, …, G=256. For each evolution we collected the mean mixed-population response; after each evolution converged to form a mixed-group prototype Pn, we measured the activation of individual units to one of these final generation (no. generations = 1000) mixed-group prototype examples. We then examined the group prototypes and how the individual units responded in each group size.

We found that CNN units acted in a way that resembled our biological observations. First, we found that it was possible to maximize the activity of combined units at the same time (Fig. S3a, right), however, the mixed activation decreased as a function of the number of units combined. For example, on average, when fc8 units were activated separately, the mean activation value was 45.33±0.14; when pairs of units were combined, their mean activation was 29.9±2.8 or down to 66%; combining four units dropped activity to 43% (19.5±2.5), eight units, 31% (14.1±1.6); 16 units, 19% (8.6±0.9) and by 256 units, to 3% (1.4±0.1, Fig. S3a, left). This was in a noiseless simulation: if we consider the response variability of real neurons, it is evident that evolutions would fail because the activation rise would be immediately swamped by noise. When concurrent activation maximization was successful, prototypes did not degrade as much as combined: for example, the mixed prototype of the blue sarong fc8 output unit (776) and the beer glass unit (442) showed a beer glass with a touch of blue (Fig. S3B). As more units were added, the group prototype became larger, covering the whole image as a texture, rather than becoming compressed into a black round object (Fig. S3c). Across all experiments, the evolutions began to “fail” when N = 16 units were combined (Fig. S3E), and their group prototypes were all extended and comprised multiple parts. We conclude that (a) successful evolutions combining units with different selectivity lead to extended prototypes that combine their individual preferred images, and that (b) mixing too many units results in failed evolutions. Because we were able to evolve from some PFC sites, both single- and multiunits, we believe that the MU prototype shapes are not the result of mixing differently tuned units, but from mixing the activity of similarly tuned neurons.



Supplemental Figure 4. Neuronal responses to images and sounds across V4, vlPFC, and CPB sites. (**a**) Responses of individual channel sites to images (left column) and sounds (right column); channels reported activity from sites in V4 (red, monkey C), caudal parabelt (gold, monkey D), and vlPFC (dark and light blue, monkeys C and D). (**b**) Examples from individual sites, showing responses to images (solid lines) and sounds (broken lines). (**c**) Mean z-scored response per site (individual dots, ± SEM) to images and sounds. Reward delivery occurs around 600 ms. All experiments tested N = 32 sites per array, 128 sites total).