# Mapping Macaque to Human Cortex with Natural Scene Responses

Kasper Vinken[1]*, Saloni Sharma[1], Margaret S. Livingstone[1]

[1]Department of Neurobiology, Harvard Medical School, Boston, MA 02115, USA
*Corresponding author: kasper_vinken@hms.harvard.edu

Neuroscience has long relied on macaque studies to infer human brain function, yet identifying functionally corresponding brain regions across species and measurement modalities remains a fundamental challenge. This is especially true for higher-order cortex, where functional interpretations are constrained by narrow hypotheses and anatomical landmarks are often non-homologous. We present a data-driven approach for mapping functional correspondence across species using rich, naturalistic stimuli. By directly comparing macaque electrophysiology with human fMRI responses to 700 natural scenes, we identify fine-grained alignment based on response pattern similarity, without relying on predefined tuning concepts or hand-picked stimuli. As a test case, we examine the ventral face patch system, a well-studied but contested domain in cross-species alignment. Our approach resolves a longstanding ambiguity, yielding a correspondence consistent with full-brain anatomical warping but inconsistent with prior studies limited by narrow functional hypotheses. These findings show that natural image-evoked response patterns provide a robust foundation for cross-species functional alignment, supporting scalable comparisons as large-scale primate recordings become more widespread.

## Introduction

Neuroscience depends on cross-species and cross-measurement comparisons to understand the principles of human brain function. Research in humans is mostly based on non-invasive techniques such as functional magnetic resonance imaging (fMRI), which measures broad, indirect signals throughout the entire brain. This technique has been widely used to localize function to cortex, for instance to identify visual regions with response biases for edges, shapes, faces, bodies, or other categories[1–6]. Disentangling the neural computations in these regions requires more focal recording techniques that are largely restricted to animal models. Therefore, much of our understanding of human visual processing has been inferred indirectly from electrophysiological measures in the macaque visual system. A central challenge for systems neuroscience is to connect those findings by aligning responses across measurement modalities and identifying functionally corresponding brain regions across species. Establishing such

31    correspondence is not only critical for interpreting human brain function based on findings from animal

32    models but also provides insights into the evolutionary organization of the cortex[7].

33    Existing approaches often rely on narrowly defined functional hypotheses or a small number of diagnostic

34    stimuli to infer correspondence. But such features may fail to capture the full complexity of cortical

35    representations, especially in higher-order areas where conceptual distinctions between regions remain

36    poorly understood[8]. As large-scale, brain-wide recordings in non-human primates become more widely

37    available[9], and as neuroscience expands into more abstract cortical territories, these limitations become

38    more acute, highlighting the need for scalable, hypothesis-agnostic methods that can leverage such data.

39    Here, we propose one such framework: a flexible, data-driven approach for identifying functionally

40    corresponding brain regions across species using response selectivity to a large number of naturalistic

41    stimuli. By leveraging the rich, graded response patterns evoked by diverse natural scenes, this method

42    identifies functional alignment without relying on predefined tuning axes, stimulus categories, or

43    conceptual heuristics.

44    To evaluate this approach, we focus on the ventral face patch system, a well-characterized yet still

45    unresolved test case for cross-species alignment. This system comprises a series of patches that respond

46    more to faces than to other objects ("face selectivity"[3,10]). In humans, the ventral face patches form a

47    sequence that extends posteriorly from the occipital face area (OFA) through the fusiform face area's

48    posterior (FFA-1) and anterior parts (FFA-2) and culminates anteriorly at the anterior temporal lobe face

49    patch (ATL). In the macaque inferotemporal cortex (IT), a comparable sequence unfolds from the posterior

50    lateral face patch (PL) through the middle lateral (ML) and anterior lateral face patches (AL), leading to the

51    anterior medial face patch (AM)[11,12]. These patches contain a high proportion of face-selective neurons

52    called "face cells", which have been studied extensively in macaque visual cortex[10,13–30]. While both species

53    exhibit a broadly conserved posterior-to-anterior axis, the precise correspondence of specific patches

54    remains unresolved[12]. The macaque central IT face patch ML and human FFA are often considered

55    homologous[11,31–33]. This notion is supported by predictions based on full-brain anatomical warping, which

56    suggests an ML to FFA and an AL to ATL correspondence[33,34]. Nevertheless, evidence from the surrounding

57    cortical topography suggests a possible PL to OFA, ML to FFA-1, and AL to FFA-2 correspondence[35] or even

58    a PL/ML to OFA and AL to FFA correspondence[36,37].

59    To further complicate the picture, anatomical evidence may be insufficient on its own to establish the

60    correspondence of face areas, as regions can reorganize, duplicate, segregate, or enlarge through cortical

61    expansion during evolution[38,39,45]. Functional evidence is therefore essential to determine whether two

2

62 regions in macaques and humans serve similar roles in visual processing. Traditionally, functional

63 correspondence is determined based on a conceptual interpretation of the role that a region may serve in

64 brain processing. For macaque face regions, such a conceptual distinction is how neurons respond to heads

65 presented at different orientations. Whereas ML neurons show viewpoint-specific responses, more

66 anterior AL neurons show mirror-symmetric viewpoint invariance[16], a tuning property that can be

67 retrieved from fMRI activation patterns[40]. In humans, one study found evidence for this mirror-symmetric

68 coding in FFA, but not in OFA[41], suggesting an ML to OFA and AL to FFA functional correspondence. An

69 alternative interpretation is that ML may correspond to the more posterior part FFA-1 and AL to the more

70 anterior part FFA-2[12]. Others have suggested that evidence for mirror-symmetric tuning could merely be

71 explained by low-level confounds[42,43]. The lack of consistency amongst studies calls into question the

72 validity of relying on a single, hand-picked tuning property to arbitrate between scenarios of functional

73 correspondence across species. Indeed, recent findings have shown that, even for face cells, the neural

74 tuning should not be reduced to a few human-interpretable features or concepts[29,30].

75 Here, we test whether we can establish the functional correspondence of human and macaque face areas

76 without relying on specific assertions of functional specialization. Rather than focusing on only face images

77 or pre-defined functional properties, we compared neural responses from macaque and human cortex to

78 a shared, diverse set of natural scenes. Specifically, we leveraged an existing large-scale dataset of human

79 fMRI responses to complex natural scenes[44], and recorded neural responses in macaque central and

80 anterior face regions using the same stimuli. This allowed us to evaluate whether response selectivity

81 across a large number of natural images can resolve competing hypotheses about the alignment of face-

82 selective regions across species. We found that a direct comparison of these responses supported a

83 mapping in which macaque ML corresponds to human FFA and AL to ATL, consistent with predictions from

84 full-brain anatomical warping alone[33]. Thus, our approach resolved prior inconsistencies and

85 demonstrated that broad, natural image-evoked response profiles can establish functional alignment

86 across species, without assuming conceptual distinctions between regions.

## Results

88 We presented stimuli from the Natural Scenes Dataset[44] to five macaque monkeys (initials A, OG, P, R, and

89 B1; hereafter referred to as M1-M5), one with an array in primary visual cortex (V1; M1: N=11 reliably

90 responsive multiunit sites, see Methods), two with arrays in CIT at the location of the middle lateral face

91 patch (ML; M2: N=23 units; M3: N=33 units), and two with arrays in AIT targeting the anterior lateral face

92 patch (AL; M4: N=74 units; M5: N=40 units). The stimulus set consisted of a rich set of 700 photographs of

93    a variety of things in a scene context (including animals, humans, sports, food…) captured at varying

94    distances. Most of these stimuli did not prominently feature faces. On average, each of the IT arrays (but

95    not the V1 array) responded most strongly to images with prominent faces (**Fig. 1a**). To evaluate face

96    selectivity of individual units, we computed a face versus no-face d' metric comparing the response to

97    images with prominent faces to those without faces or animals (see Methods). The units recorded from IT

98    arrays had a high level of face selectivity (mean face versus no-face d'; M2, CIT: 1.46, 95% CI[1.24,1.67];

99    M3, CIT: 2.12, 95% CI[1.76,2.48]; M4, AIT: 2.30, 95% CI[2.14,2.47]; M5, AIT: 3.71, 95% CI[3.44,3.98]), unlike

100    the V1 array (M1: 0.15, 95% CI[-0.03,0.34]).

101    Similarly, for the human fMRI data, all four localized face area regions of interest (ROIs) responded most

102    strongly to images with prominent faces (**Fig. 1b**). Note that we excluded the mid temporal lobe face ROI,

103    since it was localized in only two subjects. We computed a face versus no-face d' per ROI of each individual

104    subject, confirming that these regions have strongly face selective responses (mean face versus no-face

105    d'; OFA: 1.70, 95% CI[1.09,2.31]; FFA-1: 2.78, 95% CI[2.15,3.40]; FFA-2: 2.72, 95% CI[1.92,3.52]; ATL: 1.93,

106    95% CI[1.38,2.47]). Thus, the CIT and AIT arrays were highly face selective, like human face areas,

107    confirming that we successfully targeted the middle and anterior face-selective regions in IT cortex.

108    To determine whether response patterns across natural scenes can resolve the correspondence between

109    macaque and human face-selective regions, we evaluated three hypothesized alignments (**Fig. 1c**): (1) ML

110    to OFA and AL to FFA; (2) ML to FFA-1 and AL to FFA-2; (3) ML to FFA and AL to ATL. That is, does macaque

111    ML correspond best to human OFA (scenario 1) or to (posterior) human FFA (scenario 2 or 3), and does

112    macaque AL correspond best to human (anterior) FFA (scenario 1 or 2) or to human ATL (scenario 3)? We

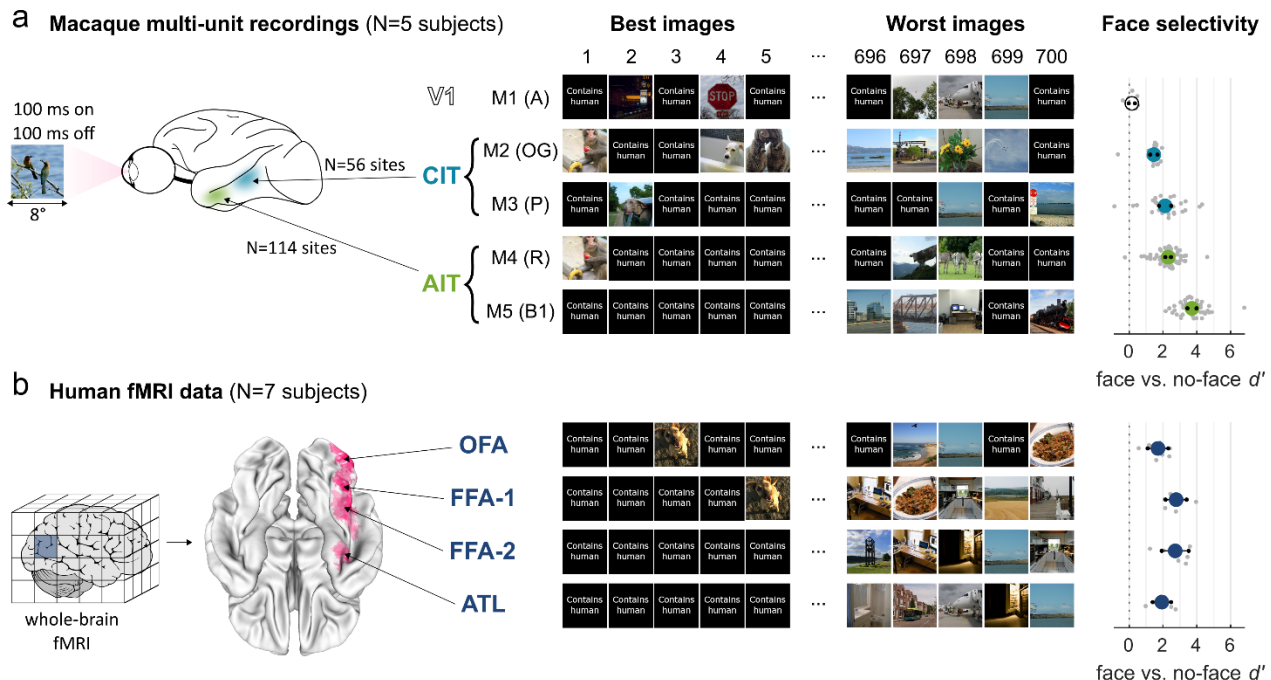113    used array-to-fMRI response pattern similarity to arbitrate between these scenarios.

**Figure 1. Data and hypotheses. a.** Central and anterior IT array locations show face-selective responses. Middle: images ranked according to the array response averaged across units. Right: average face selectivity. Grey markers represent individual units. Black markers indicate 95% CI. **b.** Human face-selective ROIs from the functional localizer experiment of [44]. Middle: images ranked according to the ROI response averaged across voxels and subjects. Right: average ROI face selectivity. Grey markers represent individual subjects. Black markers indicate 95% CI. **c.** Three scenarios of posterior-anterior functional alignment of face-selective cascade in the macaque and human ventral streams.

## Macaque IT arrays correlate with human fMRI face areas and beyond.

Having established the face selectivity of the monkey arrays, where in the human brain do we find similar tuning? Rather than characterizing and interpreting the tuning of neural or fMRI responses as a function of specific visual features or categories, we take an agnostic approach where we compare the full selectivity profile across all images. This approach is motivated by recent work from our lab, where we showed that the tuning of face cells is more complex than what can be characterized with faces alone, applying to all kinds of objects in a meaningful way[29,30]. The assumption that we make with this approach is that the stimulus set of 700 images is rich enough to differentiate between distinct complex neural tuning profiles, even if that tuning is too complex to be human interpretable.

We computed the similarity of responses at each specific array location (i.e., V1, CIT, or AIT) to the responses of each vertex of a subject-averaged human brain (**Fig. 2a**; see Methods, Array-to-fMRI

133    similarity). Briefly, for a given vertex in the human brain, we took the trial-then-subject-averaged response

134    vector and calculated the Pearson correlation with each individual unit's trial-averaged response vector.

135    We also calculated a joint reliability for each unit-to-vertex combination, based on the unit's trial-wise

136    split-half reliability and the vertex's subject-wise split half reliability. The macaque array-to-human brain

137    similarity was then computed by averaging for each vertex the Pearson correlations separately across V1,

138    CIT, and AIT units, and normalizing by the noise ceiling given by the unit-averaged joint reliability values.

139    This analysis yielded three macaque array-to-human brain similarity maps, based on the right human

140    hemisphere: one for V1 units (**Fig. 2b**), one for CIT units (**Fig. 2c**), and one for AIT units (**Fig. 2d**). The array-

141    to-brain similarity maps showed a smooth, graded range of negative to positive correlations values,

142    extending well beyond visual areas. This is consistent with a widespread, correlated engagement of the

143    cortex in response to these images. The map based on the macaque V1 array shows similarity biased

144    towards earlier visual areas, although it also shows marked similarity with the higher human face areas.

145    The maps based on macaque CIT and AIT arrays, in contrast, are biased towards higher human visual areas,

146    particularly face regions, as well as regions beyond visual cortex. The perceptual difference between maps

147    based on CIT and AIT was less clear and required a more direct comparison.

148    To better assess these differences, we Fisher Z-transformed the maps (see Methods, Fisher Z-

149    transformation) and subtracted the map based on CIT from the map based on AIT (**Fig. 3a**). We further

150    masked out vertices that were not correlated or were anticorrelated with both AIT and CIT (Pearson's $r$ /

151    reliability < 0.1). In this AIT minus CIT map, red hues indicate vertices in the human brain that responded

152    more like AIT units than like CIT units. Blue hues indicate vertices that responded more like CIT. This

153    visualization more clearly highlights the fine-grained differences between response similarities with AIT

154    and with CIT. As a sanity check, we subtracted a map based on all IT units from a map based on V1 units

155    (**Fig. 3b**). This V1 minus IT map confirmed that vertices in early visual cortex responded more like V1 units

156    (red hues), whereas vertices in higher visual cortex and beyond responded more like IT units (blue hues).

157    Visual inspection of the response similarities with the target human face regions indicates that human OFA

158    and FFA-1 responses were more like macaque CIT than AIT (predominantly blue hues), whereas human

159    ATL responses were more like macaque AIT than CIT (predominantly red hues). This was less clear for FFA-

160    2: some vertices corresponded better to CIT, others to AIT. For a quantitative analysis, we averaged unit-

161    to-vertex correlations across vertices for which the fraction of subjects that had a given ROI present

162    exceeded 0.33. This analysis confirmed the posterior-to-anterior trend observed through visual inspection,

163    culminating in the largest AIT minus CIT difference in ATL (**Fig. 3c**).

164    Since FFA vertices did not show a higher similarity to AIT than to CIT units, these results are inconsistent

165    with scenario 1, where the macaque AIT face region AL corresponds to human FFA (and CIT face region

166    ML to human OFA). However, thus far these results remain consistent with both scenario 2 and 3. Next,

167    we used each human subject's individually localized ROIs in both the left and right hemisphere to

168    determine whether macaque AIT units are better aligned with human FFA-2 (scenario 2) or with ATL
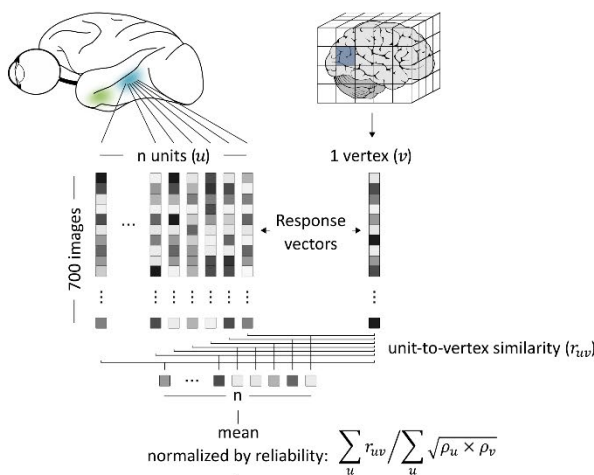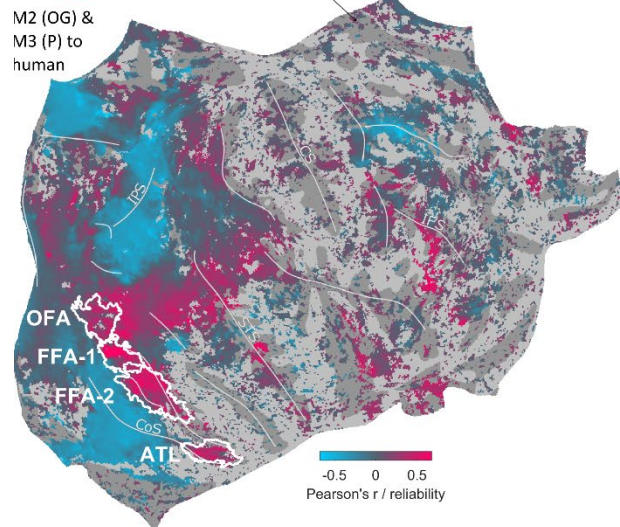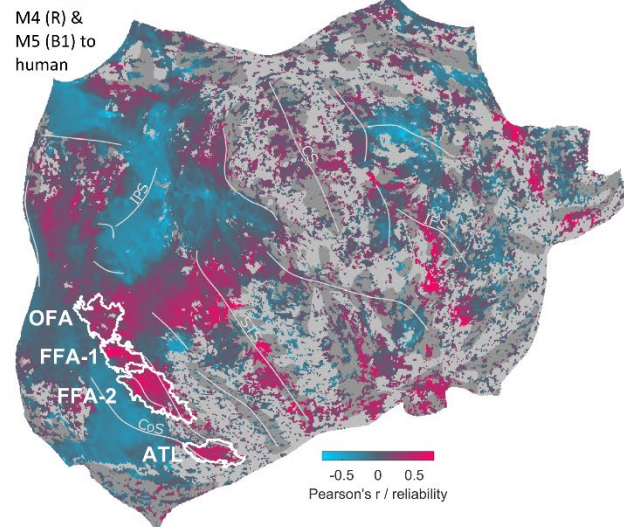
169    (scenario 3).
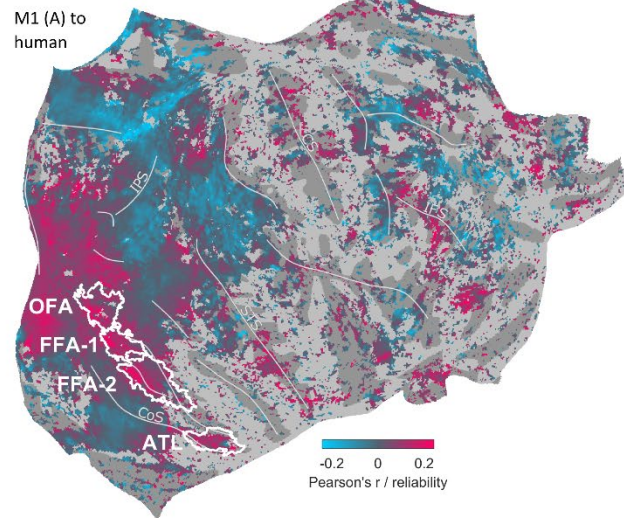


170
171    **Figure 2. Macaque array-to-human brain similarity maps. a.** Illustration of how we computed the
172    similarity between array unit responses and human fMRI vertex responses (subject-averaged in fsaverage
173    space). For each vertex $v$, the unit-to-vertex correlations $r_{uv}$ were averaged across units $u$ and then
174    normalized by the Spearman-Brown-corrected joined reliability values $\sqrt{\rho_u \times \rho_v}$, averaged across units $u$.
175    **b-d.** Full right hemisphere human brain maps of similarities with macaque V1 array responses (b), CIT array

176   responses (c), and AIT array responses (d). Vertices for which the noise ceiling (unit-averaged joint unit-
177   vertex reliability) was below 0.2 were masked out. White outlines indicate clusters of vertices for which
178   the fraction of subjects that has a given ROI present exceeds 0.33.

179



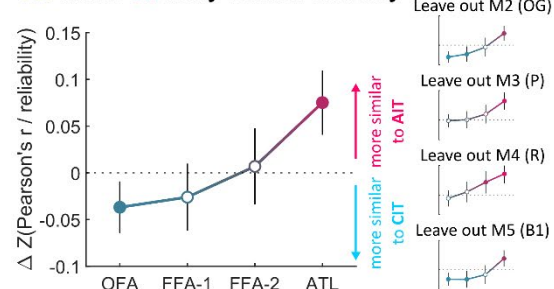**Figure 3. Monkey AIT matched higher human visual (face) areas better than did monkey CIT. a.** Full right
hemisphere human brain map showing the difference in similarity with macaque AIT and CIT array
responses (subject-averaged in fsaverage space). Red hues indicate higher similarity to AIT units than to
CIT units. Blue hues indicate higher similarity to CIT units than to AIT units. **b.** Human brain map showing
the difference in similarity with macaque V1 and IT array responses (red hues: higher similarity to AIT
units; blue hues: higher similarity to V1). **c.** AIT minus CIT array-to-brain similarity values averaged for the
ROI contours indicated in (a). Error bars indicate 95% CIs based on a two-sample *t*-test comparing AIT units
(N=116) to CIT units (N=48). Filled markers denote p<0.05. Insets show the consistency of this trend across
monkeys by repeating the analysis while systematically leaving out one monkey at a time (akin to jackknife
resampling).

## Macaque CIT maps best onto human FFA, and AIT onto ATL

To further arbitrate between scenario 2 and 3, and to account for human subject level variability, we
computed the similarity of monkey array units to each human subject's individually defined ROIs. The
methods to compute monkey array-to-human brain similarity were analogous to **Fig. 2a** but based on the

8

195 fMRI voxel responses (trial-averaged beta values) of each human subject's native space. For each human

196 subject we then averaged array-to-brain similarity values across all voxels within each ROI (see **Fig. 4a**).

197 For both CIT arrays and AIT arrays, there was no significant difference in similarity to human brain

198 responses in FFA-2 versus FFA-1 (CIT: t(6)=0.61, p=0.5639; AIT: t(6)=1.11, p=0.3094). There was also no

199 statistically significant interaction between AIT versus CIT arrays and the difference in similarity to human

200 brain responses in FFA-2 versus FFA-1 (t(6)=2.35, p=0.0571; see **Fig. 4b**). These results are inconsistent

201 with the notion that macaque ML and AL are best distinguished in how well they match human FFA-1 and

202 FFA-2 (scenario 2).

203 In contrast, there was a significant interaction between AIT versus CIT arrays and the difference in similarity

204 to human brain responses in ATL versus FFA-2 (t(6)=4.17, p=0.0059). Indeed, when pooling similarity to

205 FFA-1 and FFA-2, we found for AIT arrays that similarity to ATL was significantly higher than similarity to

206 FFA (t(6)=2.71, p=0.0353) but not for CIT arrays (t(6)=0.55, p=0.6010). Similarity to OFA was significantly

207 lower than similarity to FFA for both AIT arrays (t(6)=3.81, p=0.0089) and CIT arrays (t(6)=2.66, p=0.0374;

208 confirming the rejection of scenario 1). Thus, overall, these results support the notion that macaque AL is

209 functionally best aligned to human ATL, and macaque ML to human FFA (scenario 3).



210
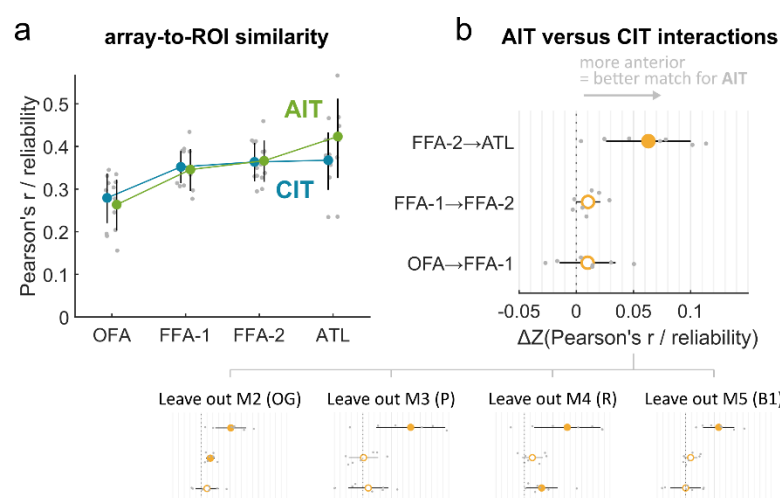
**Figure 4. Monkey AIT matched ATL better than it matched FFA. a.** Average array-to-ROI similarity. Grey
markers represent individual human subjects. Error bars indicate 95% CIs based on a one-sample *t*-test
(N=7 human subjects). **b.** Regression interaction terms between monkey array (AIT vs. CIT) and a one-step
transition in the human ROI cascade (ATL vs. FFA-2, FFA-2 vs. FFA-1, or FFA-1 vs. OFA). For example, for the
interaction with ATL vs. FFA-2 (indicated as FFA-2 -> ATL in the figure), the positive value means that the
increase in array-to-brain similarity from human FFA-2 to ATL is larger for macaque AIT than for CIT. Error
bars indicate 95% CIs based on a paired-sample *t*-test (N=7 human subjects). Filled markers denote p<0.05.
Insets show the consistency of this trend across monkeys by repeating the analysis while systematically
leaving out one monkey at a time (akin to jackknife resampling).

## Discussion

We introduce a general framework for identifying fine-grained functional correspondence across species by leveraging population-level neural responses to naturalistic stimuli. Applying this approach to the ventral face patch system, we demonstrate that rich, graded response patterns provide a robust and scalable basis for mapping functional correspondence across species and measurement modalities, without requiring predefined hypotheses or hand-picked stimuli. By directly comparing macaque electrophysiology with human fMRI responses to a shared set of natural scenes, we identified a functional alignment of face-selective areas in which macaque ML corresponds to human FFA and AL to ATL. Despite being derived purely from neural responses, this mapping was consistent with predictions from anatomical topology, suggesting that large-scale cortical layout remains a strong predictor of functional organization across primates. These findings illustrate that neural responses to natural scenes can serve as a reliable basis for cross-species functional inference, offering a flexible, hypothesis-agnostic alternative to traditional methods for comparative neuroscience.

These results help resolve a long-standing ambiguity in the posterior-to-anterior cross-species alignment of face-selective regions. Previous accounts have proposed three competing scenarios: (1) ML aligns with OFA and AL with FFA; (2) ML aligns with FFA-1 and AL with FFA-2; (3) ML aligns with FFA and AL with ATL. Using response patterns across a large set of natural images, we found that recordings from macaque CIT (targeting ML) correlated best with human FFA (both posterior and anterior), with no additional increase for ATL, thereby rejecting scenario 1. In contrast, recordings from macaque AIT (targeting AL) correlated most strongly with human ATL, ruling out scenario 2. Finally, our analyses showed that ML and AL are best distinguished by their correspondence to anterior FFA versus ATL. Together, these findings support scenario 3: a functional alignment of ML with FFA and AL with ATL. This alignment suggests a more anterior human homologue of AL than is often assumed, underscoring the importance of considering broad representational properties in higher-order cortex.

One potential explanation for previous discrepancies based on cortical topography[35–37] is that evolutionary changes may reshape cortical organization[38,39,45]. Such changes can undermine the reliability of using local topography to infer functional correspondence. On the other hand, our conclusions do converge with a posterior-to-anterior mapping derived from full-brain anatomical warping[33], suggesting that more holistic anatomical approaches may better capture conserved organizational principles. Previous studies have also compared a range of functional properties, such as mirror-symmetric tuning, the face inversion effect, face familiarity, and face selectivity, often using small sets of diagnostic stimuli[11,12,37,40,46,47]. While these

10

251 approaches are attractive for their simplicity and interpretability, they are constrained by their reliance on

252 small and highly curated stimulus sets and may miss broader feature tuning. Moreover, recent work has

253 suggested that observations of mirror-symmetric tuning in the human face selective areas may originate

254 from low-level confounds and analysis choices[42,43]. Importantly, these observations highlight a common

255 theme: broader approaches, whether leveraging full-brain anatomical alignment or a comprehensive

256 range of natural stimuli, converge on a consistent cross-species mapping, whereas narrower methods

257 based on local topography or curated stimuli can yield divergent conclusions[48].

258 More broadly, our findings illustrate how data-driven approaches using naturalistic stimuli offer a powerful

259 tool for fine-grained functional alignment well beyond the domain of face-selective regions. Traditional

260 methods that rely on small, highly curated stimulus sets risk drawing inaccurate conclusions when the

261 chosen stimuli fail to capture the richness of underlying neural representations. Under stimulus-poor

262 conditions, the inclusion or exclusion of a single image can drastically impact observed response profiles.

263 For example, while face-selective regions in both species responded more strongly to faces on average,

264 the specific images that elicited the strongest responses (Fig. 1, "best" images) differed across species—

265 highlighting nuanced differences in feature tuning that could be either amplified or overlooked with only

266 a few stimuli. These challenges are exacerbated when the stimuli are overly constrained by a conceptual

267 notion such as discrete stimulus categories[49]. Growing evidence suggests that high-level visual cortex,

268 including face-selective cells, is best characterized by an integrated feature space, where category selective

269 responses are carried by domain-general features[29,30,50–53]. By sampling broadly from the natural image

270 space, our approach captures these representational subtleties without imposing categorical assumptions.

271 This strategy enables a more flexible, scalable, and hypothesis-agnostic framework for assessing functional

272 alignment, particularly in higher-order areas where feature tuning is complex or poorly understood.

273 Establishing cross-species functional correspondence in primates is a long-standing challenge, complicated

274 by issues such as differences in measurement techniques[7]. The advent of fMRI, usable in both monkeys

275 and humans, has enabled an explosion of comparative studies inferring correspondence by comparing

276 fMRI maps of individual functional properties or stimulus contrasts[7,8,11,12,33,35,37–39,46,54–59]. Our approach

277 demonstrates that direct comparisons between electrophysiological and fMRI data can contribute

278 meaningfully to this line of research. By leveraging population-level neural responses and correlating them

279 with human fMRI responses at each vertex, we were able to delineate functional alignment without

280 requiring parallel imaging experiments in both species. Although the physiological differences between

281 spiking and BOLD signals remain, our results show that image-level tuning in multi-unit activity aligns

282  meaningfully with human fMRI responses[60–62]. Indeed, the array-to-brain correlation maps show the

283  highest similarity in regions consistent with each array's location: human early visual cortex for the

284  macaque V1 array, and ventral temporal cortex (near OFA, FFA and ATL) for the IT arrays. Furthermore, our

285  results converge with previously hypothesized homologies derived from anatomical warping of fMRI-

286  localized macaque face patches onto a human flat map[33,34]. Overall, the fact that our framework captured

287  meaningful alignment across both coarse and fine levels of cortical organization – from early versus high-

288  level visual areas to distinctions among neighboring face regions – suggests that this approach can

289  generalize to systems beyond the ventral stream, wherever shared representations exist across species.

290  Our study has several limitations. First, we recorded from only one array location per monkey, which

291  means that potential inter-individual differences are confounded with between-area differences. To

292  mitigate this, we repeated the main analyses while systematically leaving out one monkey at a time,

293  confirming that our results are not driven by any single individual. Second, although we focused on two

294  array locations targeting ML and AL, it remains to be seen whether similar cross-species correspondences

295  can be established in other high-level visual areas or cognitive systems. Third, while natural scenes cover

296  a rich and diverse stimulus space, they do not capture the full range of real-world dynamics and behaviors

297  that shape neural activity. Future work could extend this approach by incorporating dynamic stimuli and

298  behavioral context, by sampling simultaneously and more densely from a larger swath of IT or even

299  multiple visual areas[9], or by applying monkey fMRI to enable full-brain functional warping. Finally, while

300  our approach addresses the question of functional alignment, it does not, by itself, establish phylogenetic

301  homology in the strict evolutionary sense. Such claims require complementary evidence from

302  cytoarchitecture, connectivity, and developmental lineage, which is often inaccessible in humans or for

303  brain regions with ambiguous structural boundaries. In such contexts, our method provides functional

304  evidence to inform hypotheses about homology.

305  In conclusion, our results point toward a flexible, data-driven framework for establishing functional

306  correspondences across species and modalities. By using naturalistic stimuli to capture rich, high-

307  dimensional and graded response patterns, this approach overcomes the limitations of traditional

308  alignment methods that rely on narrow stimulus sets or predefined functional heuristics. Although we

309  demonstrated this in the ventral visual face patch system, this method is broadly applicable, offering a

310  general framework for establishing cross-species alignment in domains where naturalistic data can be

311  collected. As technological advances continue to enable more large-scale and brain-wide recordings in

312  non-human primates[9], frameworks that scale with this complexity and support cross-species comparisons

313    of increasingly abstract, hard-to-define functions will be essential. Our findings highlight the value of

314    naturalistic, data-driven alignment as a foundation for such efforts, offering both a principled path toward

315    unified models of brain function and a powerful complement to anatomical and developmental

316    approaches in the study of brain evolution.

## Methods

### NSD data

319    The human data analyzed in this study were obtained from the Natural Scenes Dataset (NSD), a large-scale

320    fMRI dataset comprising whole-brain, high-resolution measurements collected from eight subjects who

321    viewed natural scenes while engaged in a continuous recognition task. Details regarding data collection

322    and preprocessing can be found in the original publication[44]. Face-selective regions were defined based

323    on the results of the category-selective functional localizer included in the NSD experiments. Labels for

324    these regions were assigned by the NSD authors. For the subject-averaged human brain maps, the

325    provided probabilistic localizer maps were used with a threshold of 0.33 (i.e., only vertices corresponding

326    to a face-selective region in at least one-third of the subjects were included). We additionally defined each

327    subject's face-selective ROIs individually, based on their native-space face localizer t-maps, using a

328    threshold of t > 3. One subject was excluded because no anterior face region could be defined based on

329    the functional localizer, resulting in a final sample of seven subjects. In addition, a medial temporal lobe

330    face area that could be defined in only two remaining subjects was excluded.

### Animals and arrays

332    Five adult male macaques were used in this experiment: four rhesus macaques (*Macaca mulatta*; initials

333    A, OG, P, and B1—referred to as M1, M2, M3, and M5, respectively) aged 8-16 years and one pigtailed

334    macaque (*Macaca nemestrina*; initial R, referred to as M4), aged 13 years. All five monkeys were implanted

335    with chronic microelectrode arrays: one in V1 and four in the lower bank of the superior temporal sulcus.

336    Specifically, two monkeys were implanted with 32-channel floating microelectrode arrays (FMA;

337    Microprobes for Life Sciences, Gaithersburg, MD): V1 of M1 and CIT of M3. Three monkeys were implanted

338    with 64-channel NiCr microwire bundle arrays (Microprobes for Life Sciences, Gaithersburg, MD)[63]: CIT of

339    M2 and AIT of M4 and M5. The target location for the face patch arrays was identified using fMRI face

340    localizers (M3, M4, M5) or anatomical landmarks (M2; STS 'bumps': Arcaro et al[64]). All procedures were

341    approved by the Harvard Medical School Institutional Animal Care and Use Committee and conformed to

342    National Institutes of Health guidelines provided in the Guide for the Care and Use of Laboratory Animals.

### Experiments

During recordings, the monkeys were performing a fixation task in which they were rewarded with drops of juice to fix their gaze on a spot in the middle of a 53-cm LCD monitor. The gaze position was monitored using an ISCAN system (ISCAN, Woburn, MA). The experiments were controlled with MonkeyLogic (https://monkeylogic.nimh.nih.gov/). During fixation, images were presented at a size of 8 visual degrees and a rate of 100 ms on and 100 ms off. The images were presented at the center of the mapped receptive field, with a jitter of ±1 to ±2 visual degrees for IT arrays. An average of 32 trials were presented per stimulus.

### Stimuli

The stimulus set consisted of a subset of 1000 images that were shared across human subjects in the NSD dataset[44]. The NSD image set contains a rich variety of photographs taken from Microsoft's Common Objects in Context (MS COCO[65]) image database. For this paper, we focus on 700 images for which there were at least two presentations for each human subject, to be able to compute a split-half reliability using the exact same images for each subject.

### fMRI-guided array targeting

The target location for the array placement for 3 out of 4 monkeys (M3, M4, M5) was identified using an fMRI face localizer. The details of the fMRI experiments are described in Arcaro and Livingstone[66] and will only be summarized here briefly. The monkeys were scanned using custom-made four-channel surface coils (by A. Maryam at the Martinos Imaging Center), in a 3T TIM Trio scanner with an AC88 gradient insert. Functional images were acquired using a repetition time = 2s, echo time = 13ms, flip angle = 72°, iPAT = 2, matrix size = 96 x 96 mm, resolution = 1-mm isotropic, and 67 contiguous sagittal slices. To enhance signal-to-noise ratio and increase contrast [67], the monkeys were injected with 12 mg/kg of monocrystalline iron oxide nanoparticles (Feraheme, AMAG Pharmaceuticals, Cambridge, MA, USA) before each scanning session. The face localizer consisted of randomly shuffled 20s blocks of face or nonface objects, interleaved with 20s of a neutral gray screen.

### Data analysis

#### *Firing rates*

To compute average firing rates, we counted the number of spikes in a 200 ms window following stimulus presentation onset. The latency of the response was selected based on a visual assessment of the global average (across channels and stimuli) peristimulus time histogram, ranging from 65 ms to 125 ms.

373    Following previous procedures[29,30], we used an a-priori response reliability criterion of >0.4 to include only

374    visually driven, selective neural units for further analysis. This yielded 11 multiunit sites from V1

375    recordings, 48 multiunit sites from CIT recordings and 116 multiunit sites from AIT recordings.

376    *Response reliability*

377    We determined a trial-wise split-half reliability per neural unit and per voxel in each NSD subject brain,

378    and we determined a subject-wise split-half reliability per vertex in the NSD 'fsaverage' space. To obtain

379    trial-wise split-half reliability, we first computed the correlation $r$ between the average response vector

380    based on odd trials and the one based on even trials. We then applied the Spearman-Brown correction to

381    obtain a reliability $\rho = 2r/(1 + r)$.

382    To obtain subject-wise split-half reliability, $r$ was computed as the correlation obtained from the trial-and-

383    subject averaged response vectors for every possible way to split the NSD subjects in two, averaged across

384    splits.

385    *Face selectivity*

386    We assessed face selectivity by computing the face versus no-face d' sensitivity index comparing trial-

387    averaged responses to a subset of images with prominent primate faces (N=26 out of all 700 images)

388    versus all images without faces or animals (N=255 out of all 700 images):

$$d' = \frac{\mu_F - \mu_{NF}}{\sqrt{(\sigma_F^2 + \sigma_{NF}^2)/2}}$$

390    where $\mu_F$ and $\mu_{NF}$ are the across-stimulus means for faces and non-faces, and $\sigma_F$ and $\sigma_{NF}$ are the across-

391    stimulus SDs.

392    *Array-to-fMRI similarity*

393    We quantified the similarity between microelectrode array responses and fMRI vertex/voxel responses as

394    follows. For each fMRI vertex/voxel $v$ and array multiunit $u$, we computed the Pearson correlation $r_{uv}$

395    between their respective response vectors across images. For the flat-maps, we then averaged the unit-

396    to-vertex correlation values across all array units $u$ to obtain a single correlation value for each vertex $v$.

397    For the ROIs, we averaged across all units and voxels to obtain a single correlation value for each ROI.

398    To estimate a noise ceiling, we computed joint reliability values as $\sqrt{\rho_u \times \rho_v}$, where $\rho_u$ and $\rho_v$ represent

399    the response reliability of each unit and each vertex/voxel, respectively (see Methods, Response

15

400  reliability). Similar to the correlation values, we averaged these joint reliability values across units (and

401  voxels in the case of ROIs) to obtain a single noise ceiling value for each vertex $v$ or each ROI.

402  Finally, we normalized each array-to-vertex/ROI correlation by its corresponding noise ceiling to obtain a

403  single array-to-vertex/ROI similarity value.

404  *Fisher Z-transformation*

405  The correlation-based array-to-fMRI similarity values are inherently bounded between -1 and 1 and exhibit

406  a skewed distribution, especially near the extremes. Therefore, for statistical comparison between

407  similarity values associated with a given vertex/ROI, we applied the Fisher Z-transformation, $Z(r) =$

408  $\text{atanh}(r)$, to normalize the distribution and stabilize variance across the range of similarity values[68].

## Acknowledgments

412  Author contributions: Conceptualization: K.V. and M.S.L. Data curation: K.V. Formal analysis: K.V.

413  Methodology: K.V. Investigation: S.S. and M.S.L. Visualization: K.V. Supervision: M.S.L. Writing—original

414  draft: K.V. and S.S. Writing—review and editing: K.V., S.S. and M.S.L.

415  Competing interests: The authors declare no competing interests.

## References

417  1.  Malach, R., Reppas, J.B., Benson, R.R., Kwong, K.K., Jiang, H., Kennedy, W.A., Ledden, P.J.,
418      Brady, T.J., Rosen, B.R., and Tootell, R.B. (1995). Object-related activity revealed by functional
419      magnetic resonance imaging in human occipital cortex. Proceedings of the National Academy
420      of Sciences *92*, 8135–8139. https://doi.org/10.1073/pnas.92.18.8135.

421  2.  Kourtzi, Z., and Kanwisher, N. (2000). Cortical Regions Involved in Perceiving Object Shape.
422      The Journal of Neuroscience *20*, 3310–3318. https://doi.org/10.1523/JNEUROSCI.20-09-
423      03310.2000.

424  3.  Kanwisher, N., McDermott, J., and Chun, M.M. (1997). The Fusiform Face Area: A Module in
425      Human Extrastriate Cortex Specialized for Face Perception. The Journal of Neuroscience *17*,
426      4302–4311. https://doi.org/10.1523/JNEUROSCI.17-11-04302.1997.

427  4.  Downing, P.E., Jiang, Y., Shuman, M., and Kanwisher, N. (2001). A Cortical Area Selective for
428      Visual Processing of the Human Body. Science (1979) *293*, 2470–2473.
429      https://doi.org/10.1126/science.1063414.

430  5.  Grill-Spector, K., Kushnir, T., Hendler, T., Edelman, S., Itzchak, Y., and Malach, R. (1998). A
431      sequence of object-processing stages revealed by fMRI in the human occipital lobe. Hum
432      Brain Mapp *6*, 316–328. https://doi.org/10.1002/(SICI)1097-
433      0193(1998)6:4&#x0003c;316::AID-HBM9&#x0003e;3.0.CO;2-6.

434  6.  Vinberg, J., and Grill-Spector, K. (2008). Representation of Shapes, Edges, and Surfaces Across
435      Multiple Cues in the Human Visual Cortex. J Neurophysiol *99*, 1380–1393.
436      https://doi.org/10.1152/jn.01223.2007.

437  7.  Sereno, M.I., and Tootell, R.B. (2005). From monkeys to humans: what do we now know about
438      brain homologies? Curr Opin Neurobiol *15*, 135–144.
439      https://doi.org/10.1016/j.conb.2005.03.014.

440  8.  Tootell, R.B.H., Tsao, D., and Vanduffel, W. (2003). Neuroimaging Weighs In: Humans Meet
441      Macaques in "Primate" Visual Cortex. The Journal of Neuroscience *23*, 3981–3989.
442      https://doi.org/10.1523/JNEUROSCI.23-10-03981.2003.

443  9.  Trautmann, E.M., Hesse, J.K., Stine, G.M., Xia, R., Zhu, S., O'Shea, D.J., Karsh, B., Colonell, J.,
444      Lanfranchi, F.F., Vyas, S., et al. (2023). Large-scale high-density brain-wide neural recording in
445      nonhuman primates. Preprint, https://doi.org/10.1101/2023.02.01.526664
446      https://doi.org/10.1101/2023.02.01.526664.

447  10. Tsao, D.Y., Freiwald, W.A., Tootell, R.B.H., and Livingstone, M.S. (2006). A cortical region
448      consisting entirely of face-selective cells. Science *311*, 670–674.
449      https://doi.org/10.1126/science.1119983.

450  11. Tsao, D.Y., Moeller, S., and Freiwald, W.A. (2008). Comparing face patch systems in macaques
451      and humans. Proceedings of the National Academy of Sciences *105*, 19514–19519.
452      https://doi.org/10.1073/pnas.0809662105.

453  12. Yovel, G., and Freiwald, W.A. (2013). Face recognition systems in monkey and human: are they
454      the same thing? F1000Prime Rep *5*. https://doi.org/10.12703/P5-10.

17

455   13.   Freiwald, W.A., Tsao, D.Y., and Livingstone, M.S. (2009). A face feature space in the macaque
456         temporal lobe. Nat Neurosci *12*, 1187–1196. https://doi.org/10.1038/nn.2363.

457   14.   Moeller, S., Crapse, T., Chang, L., and Tsao, D.Y. (2017). The effect of face patch
458         microstimulation on perception of faces and objects. Nat Neurosci *20*, 743–752.
459         https://doi.org/10.1038/nn.4527.

460   15.   Grimaldi, P., Saleem, K.S., and Tsao, D. (2016). Anatomical Connections of the Functionally
461         Defined "Face Patches" in the Macaque Monkey. Neuron *90*, 1325–1342.
462         https://doi.org/10.1016/j.neuron.2016.05.009.

463   16.   Freiwald, W.A., and Tsao, D.Y. (2010). Functional Compartmentalization and Viewpoint
464         Generalization Within the Macaque Face-Processing System. Science (1979) *330*, 845–851.
465         https://doi.org/10.1126/science.1194908.

466   17.   Chang, L., and Tsao, D.Y. (2017). The Code for Facial Identity in the Primate Brain. Cell *169*,
467         1013-1028.e14. https://doi.org/10.1016/j.cell.2017.05.011.

468   18.   Sadagopan, S., Zarco, W., and Freiwald, W.A. (2017). A causal relationship between face-patch
469         activity and face-detection behavior. Elife *6*. https://doi.org/10.7554/eLife.18558.

470   19.   Azadi, R., Lopez, E., Taubert, J., Patterson, A., and Afraz, A. (2024). Inactivation of face-
471         selective neurons alters eye movements when free viewing faces. Proceedings of the National
472         Academy of Sciences *121*. https://doi.org/10.1073/pnas.2309906121.

473   20.   Afraz, S.-R., Kiani, R., and Esteky, H. (2006). Microstimulation of inferotemporal cortex
474         influences face categorization. Nature *442*, 692–695. https://doi.org/10.1038/nature04982.

475   21.   Afraz, A., Boyden, E.S., and DiCarlo, J.J. (2015). Optogenetic and pharmacological suppression
476         of spatial clusters of face neurons reveal their causal role in face gender discrimination. Proc
477         Natl Acad Sci U S A *112*, 6730–6735. https://doi.org/10.1073/pnas.1423328112.

478   22.   Waidmann, E.N., Koyano, K.W., Hong, J.J., Russ, B.E., and Leopold, D.A. (2022). Local features
479         drive identity responses in macaque anterior face patches. Nat Commun *13*, 5592.
480         https://doi.org/10.1038/s41467-022-33240-w.

481    23.    Khandhadia, A.P., Murphy, A.P., Romanski, L.M., Bizley, J.K., and Leopold, D.A. (2021).
482           Audiovisual integration in macaque face patch neurons. Curr Biol *31*, 1826-1835.e3.
483           https://doi.org/10.1016/j.cub.2021.01.102.

484    24.    Premereur, E., Taubert, J., Janssen, P., Vogels, R., and Vanduffel, W. (2016). Effective
485           Connectivity Reveals Largely Independent Parallel Networks of Face and Body Patches.
486           Current Biology *26*, 3269–3279. https://doi.org/10.1016/j.cub.2016.09.059.

487    25.    Taubert, J., Van Belle, G., Vanduffel, W., Rossion, B., and Vogels, R. (2015). The effect of face
488           inversion for neurons inside and outside fMRI-defined face-selective cortical regions. J
489           Neurophysiol *113*, 1644–1655. https://doi.org/10.1152/jn.00700.2014.

490    26.    Taubert, J., Goffaux, V., Van Belle, G., Vanduffel, W., and Vogels, R. (2016). The impact of
491           orientation filtering on face-selective neurons in monkey inferior temporal cortex. Sci Rep *6*,
492           21189. https://doi.org/10.1038/srep21189.

493    27.    Bardon, A., Xiao, W., Ponce, C.R., Livingstone, M.S., and Kreiman, G. (2022). Face neurons
494           encode nonsemantic features. Proceedings of the National Academy of Sciences *119*.
495           https://doi.org/10.1073/pnas.2118705119.

496    28.    Ponce, C.R., Xiao, W., Schade, P.F., Hartmann, T.S., Kreiman, G., and Livingstone, M.S. (2019).
497           Evolving Images for Visual Neurons Using a Deep Generative Network Reveals Coding
498           Principles and Neuronal Preferences. Cell *177*, 999-1009.e10.
499           https://doi.org/10.1016/j.cell.2019.04.005.

500    29.    Sharma, S., Vinken, K., Jagadeesh, A. V., and Livingstone, M.S. (2024). Face cells encode object
501           parts more than facial configuration of illusory faces. Nat Commun *15*, 9879.
502           https://doi.org/10.1038/s41467-024-54323-w.

503    30.    Vinken, K., Prince, J.S., Konkle, T., and Livingstone, M.S. (2023). The neural code for "face cells"
504           is not face-specific. Sci Adv *9*. https://doi.org/10.1126/sciadv.adg1736.

505    31.    Aparicio, P.L., Issa, E., and DiCarlo, J.J. (2016). Neurophysiological organization of the middle
506           face patch in macaque inferior temporal cortex. J Neurosci *36*, 12729–12745.
507           https://doi.org/10.1523/JNEUROSCI.0237-16.2016.

508    32.    Vinken, K., and Vogels, R. (2019). A behavioral face preference deficit in a monkey with an
509           incomplete    face    patch    system.    Neuroimage    *189*,    415–424.
510           https://doi.org/10.1016/j.neuroimage.2019.01.043.

511    33.    Rajimehr, R., Young, J.C., and Tootell, R.B.H. (2009). An anterior temporal face patch in human
512           cortex, predicted by macaque maps. Proceedings of the National Academy of Sciences *106*,
513           1995–2000. https://doi.org/10.1073/pnas.0807304106.

514    34.    Tsao, D.Y., Freiwald, W.A., Knutsen, T.A., Mandeville, J.B., and Tootell, R.B.H. (2003). Faces and
515           objects    in    macaque    cerebral    cortex.    Nat    Neurosci    *6*,    989–995.
516           https://doi.org/10.1038/nn1111.

517    35.    Lafer-Sousa, R., Conway, B.R., and Kanwisher, N.G. (2016). Color-Biased Regions of the Ventral
518           Visual Pathway Lie between Face- and Place-Selective Regions in Humans, as in Macaques.
519           The Journal of Neuroscience *36*, 1682–1697. https://doi.org/10.1523/JNEUROSCI.3164-
520           15.2016.

521    36.    Caspari, N., Popivanov, I.D., De Mazière, P.A., Vanduffel, W., Vogels, R., Orban, G.A., and
522           Jastorff, J. (2014). Fine-grained stimulus representations in body selective areas of human
523           occipito-temporal    cortex.    Neuroimage    *102*,    484–497.
524           https://doi.org/10.1016/j.neuroimage.2014.07.066.

525    37.    Pinsk, M.A., Arcaro, M., Weiner, K.S., Kalkus, J.F., Inati, S.J., Gross, C.G., and Kastner, S. (2009).
526           Neural Representations of Faces and Body Parts in Macaque and Human Cortex: A
527           Comparative    fMRI    Study.    J    Neurophysiol    *101*,    2581–2600.
528           https://doi.org/10.1152/jn.91198.2008.

529    38.    Orban, G.A., Van Essen, D., and Vanduffel, W. (2004). Comparative mapping of higher visual
530           areas    in    monkeys    and    humans.    Trends    Cogn    Sci    *8*,    315–324.
531           https://doi.org/10.1016/j.tics.2004.05.009.

532    39.    Meyer, E.E., Martynek, M., Kastner, S., Livingstone, M.S., and Arcaro, M.J. (2025). Expansion
533           of a conserved architecture drives the evolution of the primate visual cortex. Proceedings of
534           the National Academy of Sciences *122*. https://doi.org/10.1073/pnas.2421585122.

535    40.    Dubois, J., de Berker, A.O., and Tsao, D.Y. (2015). Single-Unit Recordings in the Macaque Face
536           Patch System Reveal Limitations of fMRI MVPA. The Journal of Neuroscience *35*, 2791–2802.
537           https://doi.org/10.1523/JNEUROSCI.4037-14.2015.

538    41.    Axelrod, V., and Yovel, G. (2012). Hierarchical Processing of Face Viewpoint in Human Visual
539           Cortex. Journal of Neuroscience *32*, 2442–2452. https://doi.org/10.1523/JNEUROSCI.4770-
540           11.2012.

541    42.    Ramírez, F.M., Cichy, R.M., Allefeld, C., and Haynes, J.-D. (2014). The Neural Code for Face
542           Orientation in the Human Fusiform Face Area. The Journal of Neuroscience *34*, 12155–12167.
543           https://doi.org/10.1523/JNEUROSCI.3156-13.2014.

544    43.    Revsine, C., Gonzalez-Castillo, J., Merriam, E.P., Bandettini, P.A., and Ramírez, F.M. (2024). A
545           unifying model for discordant and concordant results in human neuroimaging studies of facial
546           viewpoint    selectivity.    The    Journal    of    Neuroscience    *44*,    e0296232024.
547           https://doi.org/10.1523/JNEUROSCI.0296-23.2024.

548    44.    Allen, E.J., St-Yves, G., Wu, Y., Breedlove, J.L., Prince, J.S., Dowdle, L.T., Nau, M., Caron, B.,
549           Pestilli, F., Charest, I., et al. (2022). A massive 7T fMRI dataset to bridge cognitive neuroscience
550           and artificial intelligence. Nat Neurosci *25*, 116–126. https://doi.org/10.1038/s41593-021-
551           00962-x.

552    45.    Halley, A.C., and Krubitzer, L. (2019). Not all cortical expansions are the same: the coevolution
553           of the neocortex and the dorsal thalamus in mammals. Curr Opin Neurobiol *56*, 78–86.
554           https://doi.org/10.1016/j.conb.2018.12.003.

555    46.    Bell, A.H., Hadj-Bouziane, F., Frihauf, J.B., Tootell, R.B.H., and Ungerleider, L.G. (2009). Object
556           Representations in the Temporal Cortex of Monkeys and Humans as Revealed by Functional
557           Magnetic    Resonance    Imaging.    J    Neurophysiol    *101*,    688–700.
558           https://doi.org/10.1152/jn.90657.2008.

559    47.    Axelrod, V., and Yovel, G. (2013). The challenge of localizing the anterior temporal face area:
560           A    possible    solution.    Neuroimage    *81*,    371–380.
561           https://doi.org/10.1016/j.neuroimage.2013.05.015.

562    48.    Baker, C., and Kravitz, D. (2024). Insights from the Evolving Model of Two Cortical Visual
563           Pathways. J Cogn Neurosci *36*, 2568–2579. https://doi.org/10.1162/jocn_a_02192.

49. Ritchie, J.B., Wardle, S.G., Vaziri-Pashkam, M., Kravitz, D.J., and Baker, C.I. (2024). Rethinking category-selectivity in human visual cortex.

50. Doshi, F.R., and Konkle, T. (2023). Cortical topographic motifs emerge in a self-organized map of object space. Sci Adv *9*. https://doi.org/10.1126/sciadv.ade8187.

51. Bao, P., She, L., McGill, M., and Tsao, D.Y. (2020). A map of object space in primate inferotemporal cortex. Nature *583*, 103–108. https://doi.org/10.1038/s41586-020-2350-5.

52. Jagadeesh, A. V., and Gardner, J.L. (2022). Texture-like representation of objects in human visual cortex. Proceedings of the National Academy of Sciences *119*. https://doi.org/10.1073/pnas.2115302119.

53. Lugtmeijer, S., Sobolewska, A.M., de Haan, E.H.F., and Scholte, H.S. (2025). Visual feature processing in a large stroke cohort: evidence against modular organization. Brain *148*, 1144–1154. https://doi.org/10.1093/brain/awaf009.

54. Kanwisher, N. (2025). Animal models of the human brain: Successes, limitations, and alternatives. Curr Opin Neurobiol *90*, 102969. https://doi.org/10.1016/j.conb.2024.102969.

55. Vanduffel, W., Fize, D., Peuskens, H., Denys, K., Sunaert, S., Todd, J.T., and Orban, G.A. (2002). Extracting 3D from Motion: Differences in Human and Monkey Intraparietal Cortex. Science (1979) *298*, 413–415. https://doi.org/10.1126/science.1073574.

56. Tsao, D.Y., Vanduffel, W., Sasaki, Y., Fize, D., Knutsen, T.A., Mandeville, J.B., Wald, L.L., Dale, A.M., Rosen, B.R., Van Essen, D.C., et al. (2003). Stereopsis Activates V3A and Caudal Intraparietal Areas in Macaques and Humans. Neuron *39*, 555–568. https://doi.org/10.1016/S0896-6273(03)00459-8.

57. Orban, G.A., Fize, D., Peuskens, H., Denys, K., Nelissen, K., Sunaert, S., Todd, J., and Vanduffel, W. (2003). Similarities and differences in motion processing between the human and macaque brain: evidence from fMRI. Neuropsychologia *41*, 1757–1768. https://doi.org/10.1016/S0028-3932(03)00177-5.

58. Orban, G.A., Claeys, K., Nelissen, K., Smans, R., Sunaert, S., Todd, J.T., Wardak, C., Durand, J.-B., and Vanduffel, W. (2006). Mapping the parietal cortex of human and non-human primates. Neuropsychologia *44*, 2647–2667. https://doi.org/10.1016/j.neuropsychologia.2005.11.001.

592     59.     Kourtzi, Z., Tolias, A.S., Altmann, C.F., Augath, M., and Logothetis, N.K. (2003). Integration of
593             local features into global shapes: monkey and human FMRI studies. Neuron *37*, 333–346.
594             https://doi.org/10.1016/s0896-6273(02)01174-1.

595     60.     Goense, J.B.M., and Logothetis, N.K. (2008). Neurophysiology of the BOLD fMRI Signal in
596             Awake Monkeys. Current Biology *18*, 631–640. https://doi.org/10.1016/j.cub.2008.03.054.

597     61.     Logothetis, N.K., and Wandell, B.A. (2004). Interpreting the BOLD Signal. Annu Rev Physiol *66*,
598             735–769. https://doi.org/10.1146/annurev.physiol.66.082602.092845.

599     62.     Logothetis, N.K., Pauls, J., Augath, M., Trinath, T., and Oeltermann, A. (2001).
600             Neurophysiological investigation of the basis of the fMRI signal. Nature *412*, 150–157.
601             https://doi.org/10.1038/35084005.

602     63.     McMahon, D.B.T., Bondar, I. V, Afuwape, O.A.T., Ide, D.C., and Leopold, D.A. (2014). One month
603             in the life of a neuron: longitudinal single-unit electrophysiology in the monkey visual system.
604             J Neurophysiol *112*, 1748–1762. https://doi.org/10.1152/jn.00052.2014.

605     64.     Arcaro, M.J., Mautz, T., Berezovskii, V.K., and Livingstone, M.S. (2020). Anatomical correlates
606             of face patches in macaque inferotemporal cortex. Proceedings of the National Academy of
607             Sciences *117*, 32667–32678. https://doi.org/10.1073/pnas.2018780117.

608     65.     Lin, T., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., and Zitnick, C.L.
609             (2014). Microsoft COCO: Common Objects in Context. In European Conference on Computer
610             Vision, T. Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, ed. (Springer), pp. 740–755.
611             https://doi.org/10.1007/978-3-319-10602-1_48.

612     66.     Arcaro, M.J., and Livingstone, M.S. (2017). A hierarchical, retinotopic proto-organization of
613             the primate visual system at birth. Elife *6*, 1–24. https://doi.org/10.7554/eLife.26196.

614     67.     Vanduffel, W., Fize, D., Mandeville, J.B., Nelissen, K., Van Hecke, P., Rosen, B.R., Tootell, R.B.H.,
615             and Orban, G.A. (2001). Visual Motion Processing Investigated Using Contrast Agent-
616             Enhanced fMRI in Awake Behaving Monkeys. Neuron *32*, 565–577.
617             https://doi.org/10.1016/S0896-6273(01)00502-5.

618   68.   Sharma, S., Schaeffer, D.J., Vinken, K., Everling, S., and Nelissen, K. (2021). Intrinsic functional

619         clustering of ventral premotor F5 in the macaque brain. Neuroimage *227*, 117647.

620         https://doi.org/10.1016/j.neuroimage.2020.117647.

621