

Parallel, multi-stage processing of colors, faces and shapes in macaque inferior temporal cortex

Rosa Lafer-Sousa¹ & Bevil R Conway^{1,2}

Visual-object processing culminates in inferior temporal cortex (IT). To assess the organization of IT, we measured functional magnetic resonance imaging responses in alert monkeys to achromatic images (faces, fruit, bodies and places) and colored gratings. IT contained multiple color-biased regions, which were typically ventral to face patches and yoked to them, spaced regularly at four locations predicted by known anatomy. Color and face selectivity increased for more anterior regions, indicative of a broad hierarchical arrangement. Responses to non-face shapes were found across IT, but were stronger outside color-biased regions and face patches, consistent with multiple parallel streams. IT also contained multiple coarse eccentricity maps: face patches overlapped central representations, color-biased regions spanned mid-peripheral representations and place-biased regions overlapped peripheral representations. These results show that IT comprises parallel, multi-stage processing networks subject to one organizing principle.

IT is a large expanse of tissue that has been implicated in object perception. Anatomical studies commonly partition IT into posterior, central and anterior parts, but it is not clear whether these constitute distinct areas¹. One popular theory is that the parts represent a hierarchical organization of information processing. This idea is supported by functional magnetic resonance imaging (fMRI) in monkeys, which shows a prominent face bias at three to four locations along the posterior-anterior axis^{2–4}. Consistent with a hierarchical model, more anterior face patches putatively further along the processing chain show more complex selectivity⁵. It is not known whether the face-patch system is an exceptional case or is indicative of a general hierarchical organizational principle⁶. Previous reports have attempted to address this issue by testing for a systematic relationship among fMRI response patterns to different classes of objects^{2,7,8}. In such experiments, controlling low-level attributes is challenging, and contextual interactions may further complicate interpretation⁹. Color provides a useful tool for tackling the issue, as it has little feature similarity with shapes: any relationship between color-responsive and shape-responsive regions should reflect fundamental organizational principles. Monkeys are an ideal model in which to address these issues because fMRI signals in them can be enhanced by experimental intravenous contrast agents. Psychophysical chromatic mechanisms have been determined in monkeys¹⁰ and, as in humans, color stimuli activate multiple foci at several locations across IT^{11–15}. However, the relationship between activation patterns elicited by colors and those elicited by objects, particularly faces, has not been investigated.

In addition to the possibility of a multi-staged arrangement, anatomical data suggest that IT comprises multiple parallel routes along the posterior-anterior axis^{16–18}. Functional evidence of parallel processing is provided by imaging experiments in humans that reveal distinct foci selective for faces, other objects and colors^{19–22}.

The organization of these functionally biased regions appears to be dictated by a global eccentricity map, in which central, mid-peripheral and peripheral visual-field representations correlate with peak activation to faces, non-face objects and places²³. A coarse retinotopy has been found in posterior IT (PIT) of macaque monkey using micro-electrode recording^{13,24,25}, but it is not thought that this organization extends to central and anterior IT¹. It remains unclear whether the organizational principles found in human apply to monkey, an important issue that bears on cortical evolution. We sought to clarify the functional organization of IT in monkey, to address the extent to which IT is organized by a common principle of multiple processing stages (as suggested by the face-patch system) and by parallel-processing channels (as suggested by monkey anatomy and human imaging). We did so by testing fMRI responses in alert monkeys to shapes, carefully calibrated color stimuli and retinotopic stimuli (eccentricity and meridian mapping).

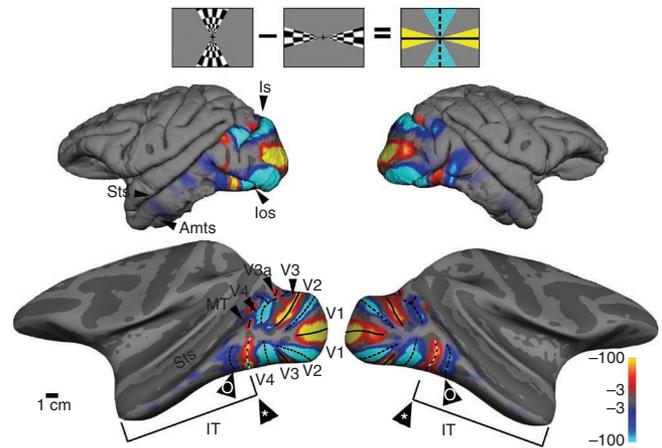
RESULTS

The posterior boundary of IT corresponds to the anterior boundary of the mid-tier visual area V4. To determine this boundary, we used fMRI to map the visual meridian representations (Fig. 1 and Supplementary Fig. 1a). To show the cortical activity buried in sulci, we computationally inflated the brain, revealing an alternating set of stripes corresponding to the boundaries of classic retinotopic areas (V1, V2 and V3). The maps also revealed a clear horizontal meridian representation at the anterior boundary of V4 along both ventral and dorsal subdivisions. In addition, the maps showed a vertical meridian representation in PIT, providing, to the best of our knowledge, the first fMRI confirmation of a retinotopic area in PIT²⁴. IT extends to the anterior tip of the temporal lobe, constituting a swath of tissue that is comparable to V1, V2 and V3 combined. In the following experiments, we sought to clarify the organizational principles that govern this region.

¹Neuroscience Program, Wellesley College, Wellesley, Massachusetts, USA. ²Department of Neurobiology, Harvard Medical School, Cambridge, Massachusetts, USA. Correspondence should be addressed to B.R.C. (bconway@wellesley.edu).

Received 21 June; accepted 24 September; published online 20 October 2013; doi:10.1038/nn.3555

Figure 1 Identification of the boundaries of IT and retinotopic visual areas using fMRI and retinotopic mapping. Images show lateral views of the left and right hemispheres of subject M1. Stronger responses to stimulation along vertical meridians are shown in blue and cyan, and stronger responses to stimulation along the horizontal meridian are shown in orange and red. Icon at top shows the visual stimulus. Top, pial surface view. Bottom, computationally inflated view. In addition to a clear V4-IT boundary (*), the maps reveal a vertical meridian representation in IT (o). Color scale bars represent significance as the common logarithm of the probability of error. Amts, anterior middle temporal sulcus; los, inferior occipital sulcus; Ls, lunate sulcus.



Functional architecture for color

Using fMRI, we determined the functional architecture for color and related it to the functional landmarks of the face-patch system. We defined color-biased regions as those showing stronger activation to drifting equiluminant colored gratings than to achromatic gratings (Fig. 2). The color and luminance stimuli were matched in cone contrast and yielded balanced activation in V1 (Fig. 2e). Colors evenly sampled the perimeter of the equiluminant plane in DKL color space²⁶ and were adjusted to yield roughly equal saturations in the perceptually uniform Commission Internationale de l'Éclairage (C.I.E.) space (Supplementary Fig. 1b). Luminance-biased responses were largely restricted to V3, V3A, area MT and parts of V4 and IT (Supplementary Fig. 2a). Prominent color-biased responses were found in V2, V4 and at several locations spanning IT.

IT has been carved into various components on the basis of cytoarchitecture and anatomical connections with V4, the amygdala, hippocampus and frontal cortex^{16–18}. In one scheme, IT comprises a posterior portion (sometimes called TEO) and an anterior portion

(also known as TE). PIT has been subdivided into dorsal and ventral components (PITd, PITv), whereas TE has been carved along both the posterior-anterior axis (central/anterior) and the dorsal-ventral axis, yielding four parts (CITd, CITv, AITd and AITv). A synthesis of connectivity studies suggests that IT comprises at least four somewhat parallel routes¹⁸; the route along the dorsal and lateral strip of IT, the focus of the present study, contains four stages, which include PIT, CIT, AIT and a region at the very tip of the temporal pole, anterior to TE. The anatomical designations for these stages are TEOd, TEpd, TEad and TGv.

The locations of the color-biased regions were generally symmetric across hemispheres in a given monkey and somewhat stereotyped

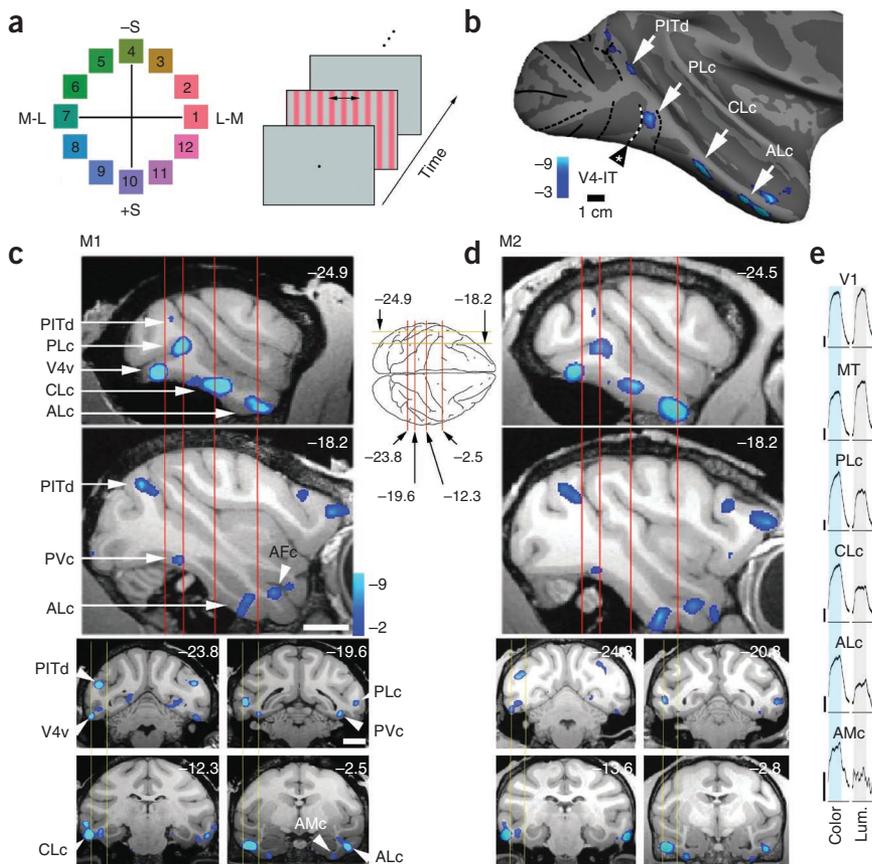
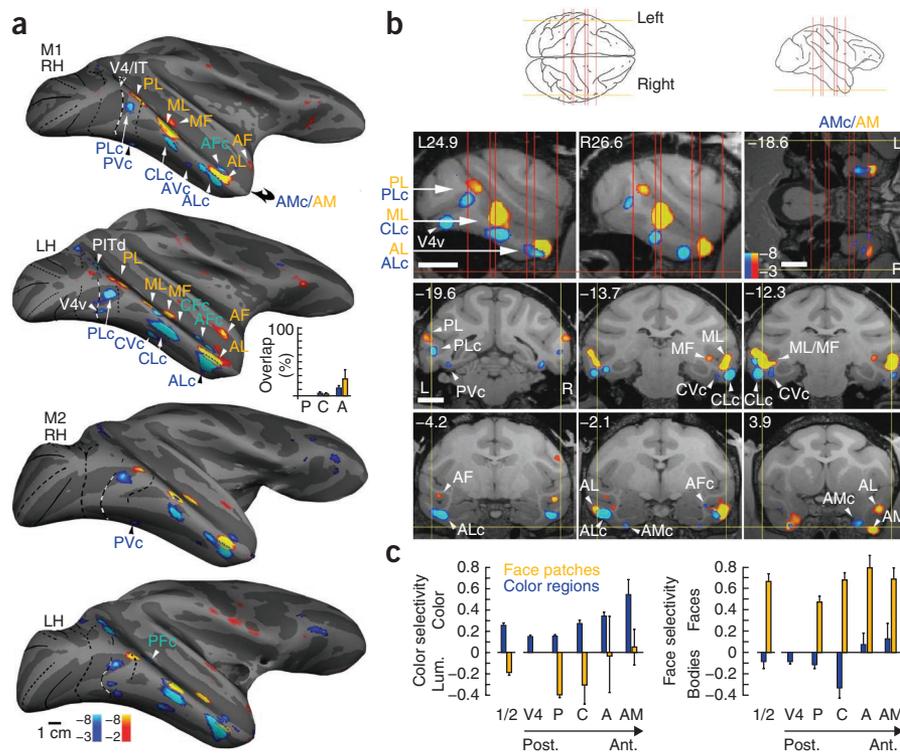


Figure 2 Functional architecture of color-biased regions in alert macaque IT. (a) Color stimuli shown in the equiluminant plane of the DKL color space²⁶; colors 1 and 7 modulate only the L and M cones, colors 10 and 4 modulate only the S cones, and other colors modulate all cone classes (Supplementary Fig. 1b). Right, stimulus procedure. Color-gray gratings were presented in 32-s blocks, maintaining mean luminance of 55 cd m⁻², interleaved with neutral gray, in two stimulus orders. (b) Regions (blue and cyan) of M1, right hemisphere, showed greater activation to colors 7 and 8 than achromatic contrast gratings (PLc, posterior lateral color; CLc, central lateral color; ALc, anterior lateral color). (c) Top two panels show color-biased activation in sagittal slices at locations indicated by the yellow lines on the top-down view of the brain schematic. Bottom four panels show color-biased activation in coronal sections corresponding to the red planes of section in the sagittal slices. Scale bars represent 1 cm; slices are given in Talairach coordinates (mm). (d) Data are presented as in c for a second monkey. Functional data is superimposed on high-resolution anatomical scans. (e) Time course traces averaging activity during all color-grating blocks and achromatic-luminance-grating blocks in each visual region. Vertical scale is 1% fMRI response (upward deflections correspond to increases in neural activity). Total number of runs for M1, 49; for M2, 21; 16 blocks per run (272 measurements per run); 16 repetition time per block; repetition time = 2 s. See Supplementary Figure 3a for the time course of the response to individual colors in each region of IT.

Figure 3 Color-biased fMRI activation (blue and cyan) found at corresponding locations to face patches (orange and red).

(a) Activity in two monkeys (M1, M2) on an inflated brain, lateral view, rotated 20° up to show the ventral surface. Left hemispheres (LH) have been horizontally flipped. Area boundaries are presented as in **Figure 1**. Inset shows the percent of face patches that were color-biased (orange bars), and the percent of color-biased regions that were part of face patches (blue bars); overlap increased from posterior to anterior (orange bars, $P = 0.04$; blue bars, $P = 0.0008$, multiple linear regression, $N = 4$). RH, right hemisphere. (b) Slices through M1 (Talairach coordinates), showing color-biased and face-biased activation. Top left, sagittal slices (L, left; R, right; section plane: yellow lines, left schematic). Top right, horizontal slice (section plane: yellow line, right schematic), and horizontal red line, sagittal slices). Bottom six panels show coronal sections (plane of section: red lines in schematics). Scale bars represent 1 cm. (c) Color selectivity and face selectivity for face patches (orange bars) and color-biased regions (blue bars). Color regions were significantly more color selective than face regions ($P = 9 \times 10^{-4}$), and face regions were significantly more face selective than color regions ($P = 4 \times 10^{-10}$) (unpaired two-tailed t tests, $N = 16$; 4 color regions per hemisphere, 4 face regions per hemisphere, 4 hemispheres). Color selectivity increased along the P→A axis ($R^2 = 0.055$, $P = 0.0002$; multiple linear regression, $N = 4$). Error bars show s.e.m. The '1/2' bars show selectivity averaged across ROIs, in ROIs defined by 1/2 the data set.



across the two monkeys (M1 and M2; **Fig. 2c,d**). In all four of the hemispheres tested, a prominent color-biased region was found in PIT, in the neighborhood of TEOd, in a region we call PLc (posterior lateral color); color-biased activity was also found in CIT (CLc, central lateral color, near TEpd) and AIT near the anterior middle temporal sulcus (ALc, anterior lateral color, near TEad). Reproducible color-biased responses were also found at the ventral-anterior tip of the temporal lobe (AMc, anterior medial color, near TGv), in the frontal cortex, and in several regions along the PIT and V4 border, as described previously^{12,13} (PITd and V4v; **Fig. 2**). PLc, CLc and ALc were often accompanied by color-biased activity in the fundus of the superior temporal sulcus (sts) and on the ventral surface. For example, there was a color-biased region at the same posterior-anterior location as PLc, but ventral and medial to it (Pvc, posterior ventral color, near TEOv), more clearly seen in the sections than in the inflated brains (**Fig. 2c**). Pvc was found in both hemispheres of both monkeys.

Despite gross stereotypy, we observed some inter-hemispheric and inter-animal variability. Both monkeys showed greater color-biased activity in the left hemisphere (and greater face-biased activation in the right hemisphere; **Fig. 3**). There were other subtle asymmetries, such as the prominence of V4v in the left hemisphere. A color-biased region in the sts (Pfc) was prominent in the left hemisphere of both monkeys, but appeared to be weak or non-existent in the right hemisphere of either monkey. Cfc was only clear in one hemisphere, the left hemisphere of M1. Cvc was found in both hemispheres of both animals, but was more prominent in the left hemisphere. In addition, in three of the four hemispheres, CLc appeared to have two lobes. A region we call AFc was found in all four hemispheres and a region that might be considered AVc was found in the right hemisphere of both monkeys (**Fig. 3a**). The time courses of the

responses to individual colors within color-biased regions are given in **Supplementary Figure 3a**.

The color-biased activity in central and anterior IT was more distinct than we found previously^{12,13} (see Online Methods). At the same time, the pattern of color-biased activity in V4 and surrounding areas (including V2; **Fig. 2**) appeared to be weaker than we previously reported¹³. We attributed this discrepancy to differences in smoothing and in stimuli used to define color-biased regions. Using all colors besides colors 7 and 8 to generate the color-response pattern (rather than just colors 7 and 8 as in **Fig. 2**), and no spatial smoothing, we found a comparable pattern in V2 and V4 to that obtained previously (**Supplementary Fig. 2c**). The pattern of activation in V4 comprises smaller scale regions (globes)¹² than those in IT, and the V2 pattern in some hemispheres appeared to be an alternating set of color-biased and luminance-biased stripes running perpendicular to the V1-V2 border. The pattern of color-biased activation in IT is consistent when using the different sets of colors, although IT shows globally stronger responses to the larger set of colors (**Supplementary Fig. 2b**).

The pattern of color-biased regions in IT is consistent with the general location of color-selective cells identified in microelectrode studies of monkey visual cortex. The location of ALc corresponds to a rich pocket of color-tuned neurons^{27,28}, a region that has also been identified with fMRI¹⁴ and positron emission tomography¹⁵. Similarly, PITd and Pvc contain predominantly color-tuned neurons, as shown by fMRI-guided microelectrode recordings^{13,29,30}. Prominent clusters of color-tuned neurons have also been described more centrally in IT²⁵, possibly corresponding to PLc or CLc. Taken together, these results suggest that the fMRI responses correspond to color selectivity of the underlying neurons and attest to the coarse stereotypy of this organization across individual monkeys.

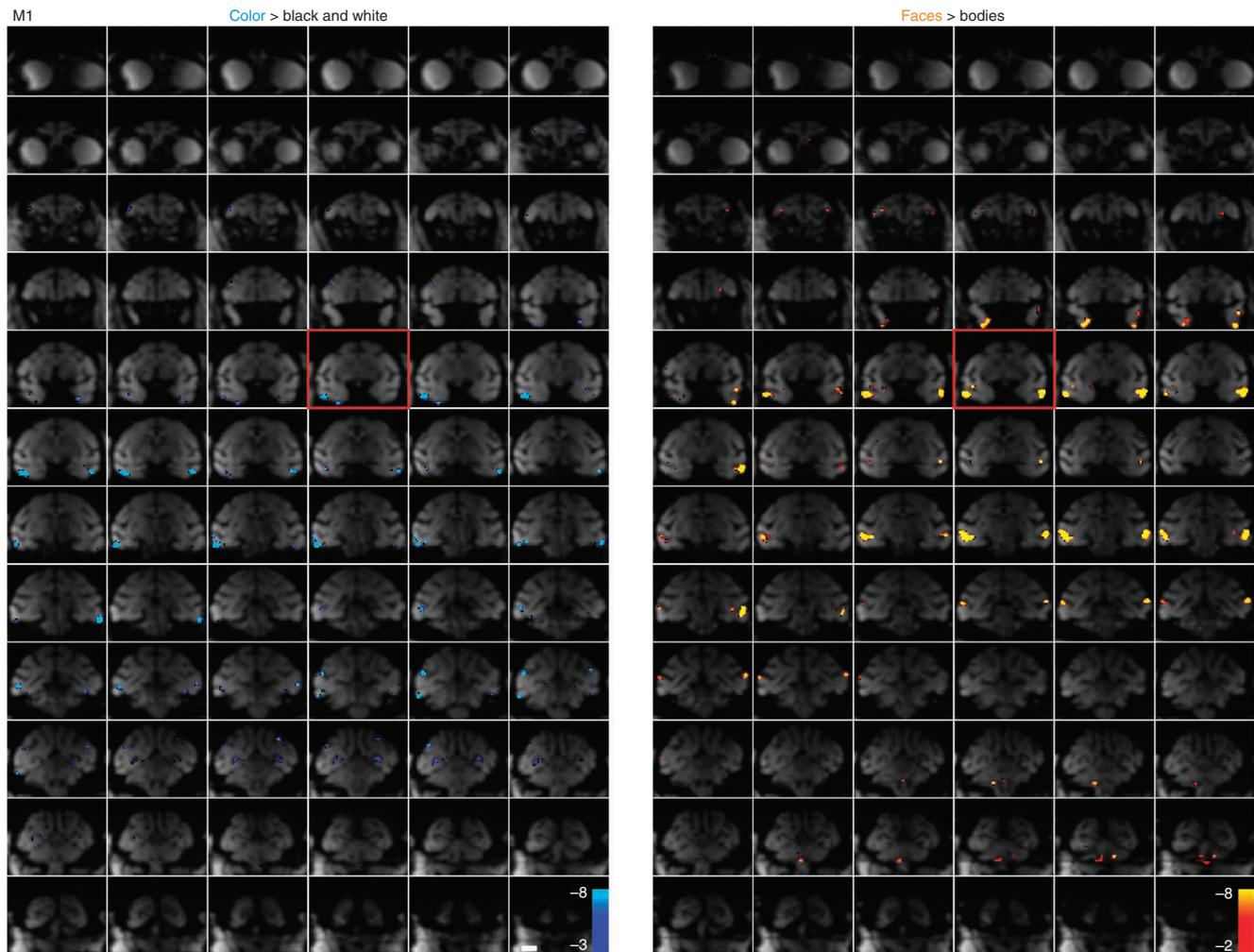


Figure 4 Raw functional echo planar coronal images showing MION activation in 1-mm-thick slices through the whole brain; color-biased activation (blue, left) and face-biased activation (orange, right) of M1. See **Supplementary Figure 5** for M2 data. Spurious activation outside the brain has been masked. The slice corresponding to the Talairach coordinate of 0 along the posterior-anterior axis is boxed in red. Scale bar represents 1 cm.

Relationship between color architecture and face patches

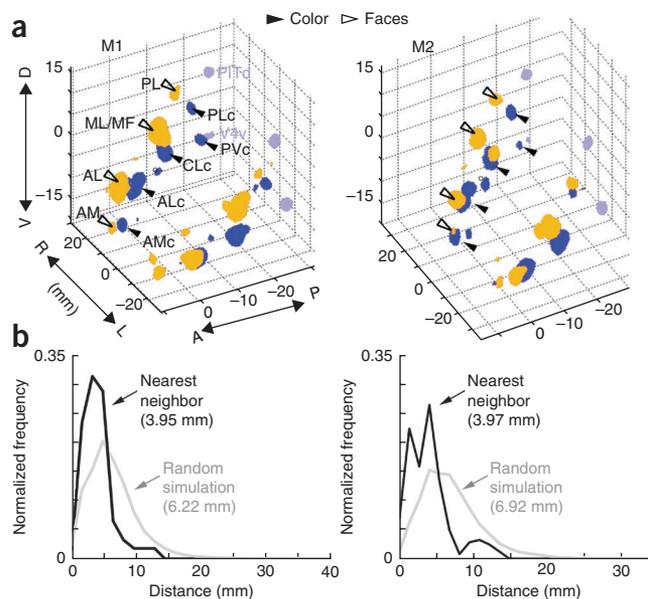
We compared the responses to color with the responses to achromatic images of faces, objects (fruits and vegetables) and places in the same subjects. As in previous reports, the fMRI maps showed three prominent clusters of face patches: PL, ML/MF, AL/AF² and several other less prominent face patches, including one at the anterior-ventral pole of the temporal lobe (AM)^{4,31} (**Fig. 3**). The time course of the responses in the face patches reflects the strong selectivity for faces (**Supplementary Fig. 3b**) and lack of color selectivity (**Supplementary Fig. 3a**). We also found face activation in several middle-ventral sub-cortical areas (**Supplementary Fig. 4**), as previously documented using higher field strengths⁴, which indicates that our techniques allow for satisfactory assessment of activity over much of the brain.

Notably, color-biased regions were found near, and typically ventral, to each cortical face patch: PLC near PL, CLc near ML, ALc near AL and AMc near AM. The face patches often extended into the sts, as described previously; color-biased regions similarly extended into the sts (for example, CFc and AFC; **Fig. 3a**). The correspondence of color regions and face patches was clear when the activation patterns were displayed on the inflated brains, but the computational algorithms involved in inflating the brain necessarily introduced distortion. In sagittal slices from each hemisphere of one monkey through PL, ML and

AL, there was a corresponding color-biased region beneath it (**Fig. 3b**). The sagittal slices showed one exception, the orphan color-biased region V4v in the left hemisphere. The horizontal section through the anterior ventral pole of the temporal lobe showed that the AM face patch was also accompanied by a color-biased region, which we call AMc (anterior medial color; **Fig. 3b**). Coronal slices through the corresponding face patches provided confirmation of the relationship between color-biased regions and face patches (**Fig. 3b**). The yoking of the color-biased activation and the face patches across the temporal lobe was evident in the mosaics showing MION (microparticulate iron oxide agent) activation on the raw functional echo planar coronal sections across the whole brain (M1 and M2; **Fig. 4** and **Supplementary Fig. 5**, stereotaxic coordinates for the center of the functionally defined regions are given in **Supplementary Fig. 5b**).

The color and face regions were defined independently; thus, it is possible for the two sets of regions to overlap completely. Despite this possibility, the two sets of regions showed little overlap (**Fig. 3a**) and little cross-stimulus selectivity (**Fig. 3c**). The color-biased regions showed little face or body bias, the posterior face patches (PL and ML/MF) showed a luminance bias, and the anterior face patches (AL/AF and AM) showed balanced color and luminance activation. Averaging across all face patches revealed a strong luminance bias, which predicts

Figure 5 Quantification of the spatial relationship of color-biased regions and face patches in alert macaque IT. **(a)** Three-dimensional plot showing color-biased regions (blue, solid arrowheads) and face patches (orange, open arrowheads) in both hemispheres for both animals (M1, M2); color-biased regions in the V4 complex are shown in light blue. Axes show Talairach coordinates (D, dorsal; V, ventral; R, right; L, left; A, anterior; P, posterior). **(b)** Histograms of distances from each IT color-selective voxel to its nearest face-selective voxel (M1, left; M2, right), and a simulation of the outcome if the regions were distributed randomly, allowing for overlap. The average separation (~4 mm) was significantly closer than expected by chance (6–7 mm, Mann-Whitney-Wilcoxon test; M1, $P = 10^{-17}$, $N = 278$ voxel to voxel distances; M2, $P = 5 \times 10^{-10}$, $N = 247$). The observed distributions showed a lower variance than the simulated distributions (squared ranks test; M1, $P < 10^{-3}$; M2, $P < 10^{-6}$). We also computed the average distance of a face voxel to its nearest color voxel and a random simulation in each hemisphere separately ($N = 4$). The distances were significantly shorter than random in all four hemispheres (Mann-Whitney-Wilcoxon; M1 left hemisphere, $P = 10^{-13}$; M1 right hemisphere, $P = 3 \times 10^{-15}$; M2 left hemisphere, $P = 5 \times 10^{-27}$; M2 right hemisphere, $P = 9 \times 10^{-12}$). The mean of the average values obtained in the four hemispheres was 4.0 mm (s.d. = 0.2), and was significantly different from chance (6.6 mm, s.d. = 0.5, t test, $P = 5 \times 10^{-5}$).



the luminance dependence of face processing³². Moreover, the color selectivity increased for color regions at more anterior locations, consistent with the hypothesis that these regions constitute a hierarchy of processing stages similar to the hierarchy described among face patches^{5,33}.

To quantify the spatial relationship of face and color regions, we constructed a three-dimensional volume of IT and volumes corresponding to the regions defined by stringent threshold criteria that yielded roughly the same number of face-selective and color-selective voxels in each hemisphere (Fig. 5a). We ran a simulation of the predicted proximity of face-biased and color-biased voxels in IT, assuming that the two sets were arranged randomly with respect to each other. In the simulation, we retained the three-dimensional shape of the cortical sheet because axonal connections could exert a powerful constraint on cortical development and evolution³⁴. We performed the simulation 1,000 times, and computed a histogram of the distance from each color-biased voxel to the nearest face-biased voxel. The average separation (~4 mm) was significantly closer than expected by chance (6–7 mm; Mann-Whitney-Wilcoxon test: M1, $P = 10^{-17}$, $n = 278$ voxel to voxel distances; M2, $P = 5 \times 10^{-10}$, $n = 247$), a finding that was true when the analysis was performed on each hemisphere separately (Fig. 5). Irrespective of a difference in mean (or median) distance between color-biased and face-biased voxels, if the two patterns of activation are systematically related, the distribution of distances between voxels should show a smaller variance than found in the random simulation. Consistent with this prediction, we found a lower variance for the observed distributions than the simulated distributions (M1, $P < 10^{-3}$; M2, $P < 10^{-6}$; squared ranks test for non-normal distributions³⁵).

Responses to non-face objects

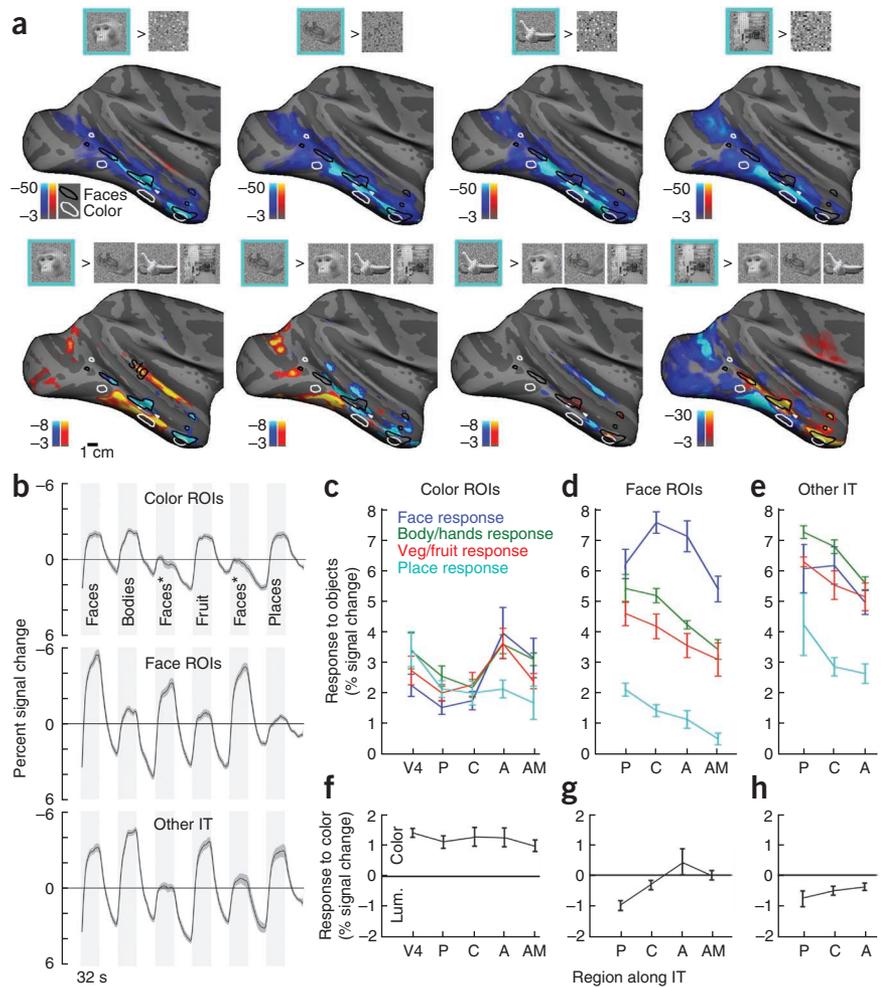
In addition to faces, we mapped the response to other types of objects (bodies and body parts, vegetables and fruit, places; Fig. 6). Significant responses to faces > scrambled faces ($P < 0.001$; Fig. 6a) were found across much of IT, although the peak significance to this stimulus contrast coincided with the face patches defined in Figure 3 (responses to faces > responses to body parts). The color-biased regions responded to all object classes (Fig. 6b,c), but the peak activation to objects (both faces and non-faces) was typically outside of the color-biased regions (Fig. 6a), providing additional support for the hypothesis that color and faces are handled by largely distinct circuits.

In addition, we found that highly significant place-biased responses tended to be ventral (and dorsal) to face patches and color-biased regions. Place-biased regions were also more prominent in PIT. The maps generated by some stimulus contrasts also showed activation in the superior temporal gyrus (Fig. 6a and Supplementary Fig. 6). The superior temporal gyrus has been found to contain neurons responsive to auditory and visual stimuli³⁶.

We quantified the magnitude of the responses to the different stimuli in regions of interest (ROIs) along the posterior-anterior axis (Fig. 6c–e). All color ROIs responded to all images; the responses to each image class were comparable in a given color-biased region, with the exception of color-biased regions of AIT, which showed a weaker response to places than to faces, body parts, and vegetables and fruits (Fig. 6c). The responses in the face patches were greatest to faces, with weaker, but still strong, responses to bodies and body parts, and vegetables and fruits; places elicited a substantially weaker response in the face patches (Fig. 6d). We also quantified the responses in regions of IT outside of the face patches and color-biased regions; these other regions were defined as the most visually active voxels in PIT, CIT and AIT, excluding color-biased regions and face patches, but comprising the same number of voxels as the face patches. These regions were most strongly modulated by images of body parts, followed by vegetables and fruits, faces, and then places (Fig. 6e). They were more strongly modulated than the color-biased regions by all image classes. As expected, the color-biased regions showed strong responses to color (Fig. 6f). The face patches showed either a luminance bias or a balanced response to color and luminance (Fig. 6g); the intervening tissue showed a luminance bias (Fig. 6h).

These results indicate that, to a substantial extent, color and object-shape information is carried by segregated channels through IT: color-biased regions showed strong responses to color and modest responses to object shape, and regions outside of the color-biased regions showed substantially stronger responses to object shape. But the extent to which non-face object-shape information is segregated into discrete channels for different object classes is less clear. The regions showing the strongest activation to bodies versus scrambled bodies were not the same as those showing preferential activation to bodies over other objects such as vegetables and fruit (Fig. 6a and Supplementary Fig. 6), although both contrasts yielded three clusters of body-biased regions along the temporal lobe. This was not true

Figure 6 Responses to images of faces, body parts, fruit and vegetables, and places across IT. **(a)** Activation maps showing regions that were more responsive to intact pictures (blue-cyan) than to scrambled versions (orange and red, top), and maps showing regions that were more responsive to one class of pictures over all other intact pictures (bottom). From left to right: faces, body parts, fruits and vegetables, places. **Supplementary Figure 6** shows results for individual comparisons. stg, superior temporal gyrus. **(b)** Time course of response in color-biased regions (top), face patches (middle) and other visually responsive parts of IT (bottom). The face stimuli indicated by an asterisk comprised familiar faces and were not included in the quantification. Time courses are the average of four hemispheres; shading indicates s.e.m. **(c)** Quantification of the responses to different image classes in the color-biased regions along IT: posterior (P), central (C), anterior (A) and anterior-ventral-medial (AM). **(d)** Quantification of responses in face patches. **(e)** Quantification of responses within other parts of IT. **(f–h)** Quantification of the response to color within the color-biased regions, face patches and other regions along IT. Responses in **f–h** were computed as the percentage signal during color (colors 7 and 8) minus the percentage signal during the achromatic grating. Negative values indicate a luminance bias and positive values indicate a color bias. Error bars (**c–h**) represent s.e.m. ($N = 4$ hemispheres). All quantification was performed on data sets obtained independently of the data used to generate the ROIs.



for faces: regions showing strongest activation to intact faces were also the same regions that showed activation to faces versus any of the other stimuli used, including bodies, which share the most similarity to faces. The face selectivity of face patches is underscored by maps of responses to all non-face objects versus scrambled non-face objects, which showed substantial activation throughout most of IT except conspicuous islands that correspond to the face patches (**Supplementary Fig. 6**). These results are consistent with present consensus that the most selective cortex in IT is selective for faces, and selectivity in the remaining object-sensitive cortex for other object categories is weaker. This tissue may not be hard-wired to be selective for a particular stimulus feature, as evidenced by the development of word-form regions and other regions of behavioral relevance^{6,37}.

Eccentricity maps in IT

Functional imaging of humans shows an eccentricity bias in IT²³. We sought to test this in monkeys to determine the extent to which the eccentricity bias represents an organizing principle across primates. The results revealed a bias for the central visual field along the lateral margin of the temporal lobe, with a peripheral bias extending dorsal (inside the sts) and ventral (**Fig. 7a** and **Supplementary Fig. 7**). Close inspection suggests three distinct representations of the central visual field, which coincide with the posterior-anterior location of PL, ML and AL. One caveat in interpreting these results is that the small disc used to identify the central visual field representations may have weakly activated face patches: some observers have reported a vague face-like quality in the checkers of the disc. Responses in V1 were comparable for stimuli at

all eccentricities (**Fig. 7b**), which was by design: the eccentricity rings were scaled by the V1 magnification factor so that each eccentricity ring activated about the same amount of V1 tissue. The responses in IT to the different eccentricity rings showed a similar pattern as that observed in V1 (**Fig. 7b**), suggesting that IT does not possess a stronger central bias than that found in V1. The face patches, however, showed a strong central bias (t test, $P = 4 \times 10^{-7}$, $n = 4$; **Fig. 7b,c**), which was evident in each face patch assessed independently (**Supplementary Fig. 8b**). The color-biased regions did not show a consistent bias for center or periphery (**Fig. 7b,c** and **Supplementary Fig. 8a**). The place-biased regions showed a peripheral bias (t test, $P = 0.01$, $n = 4$; **Fig. 7b,c**). For comparison, and to evaluate how reliable the eccentricity mapping signals were, we computed two regions of interest from results obtained during one half of the stimulus presentations, one defined as being more responsive to the central stimulation and the other defined as being more responsive to peripheral stimulation (**Fig. 7c**). We quantified the response bias in these regions using results obtained during the other half of the stimulus presentations. As predicted, both regions showed significant response bias (t tests, central representation, $P = 2 \times 10^{-4}$; peripheral representation, $P = 2 \times 10^{-6}$; $n = 4$). Taken together, these results predict that the separation between face-biased voxels and color-biased voxels should be shorter than the separation between face-biased voxels and place-biased voxels, which was the case (**Supplementary Fig. 8c**).

Finally, we quantified the responses to different images (**Fig. 7d**) and colors (**Fig. 7e**) in ROIs defined by the eccentricity mapping. We found a distinct bias for face stimuli in the central visual field

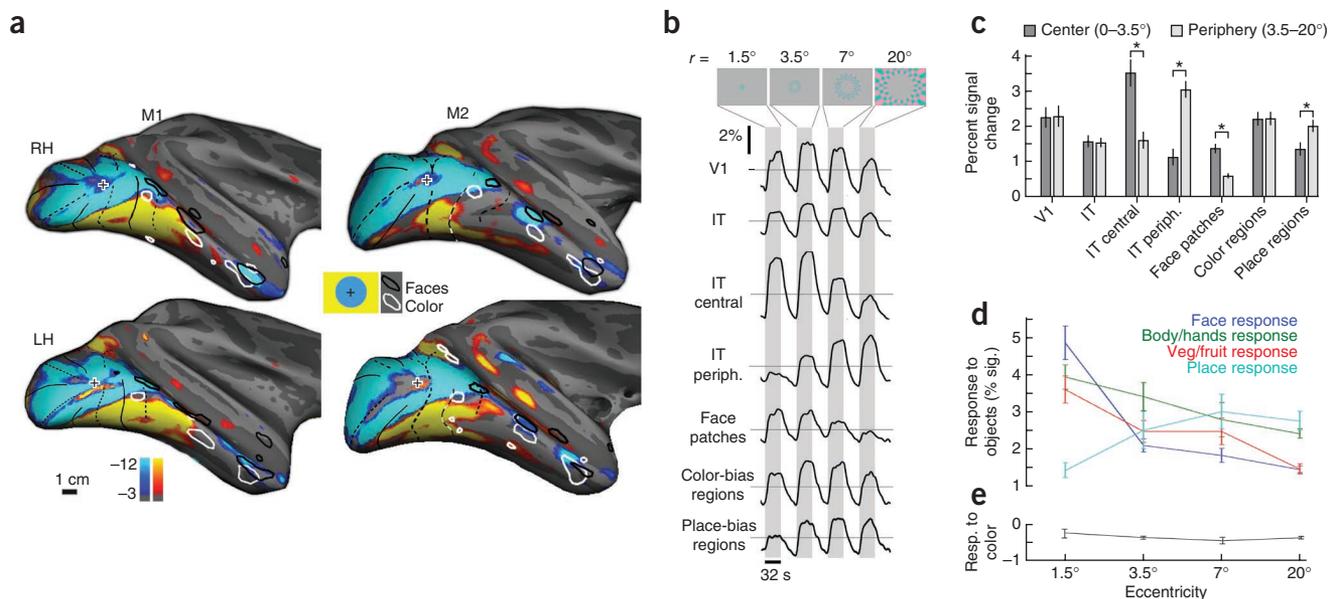


Figure 7 IT contains multiple representations of the visual field, which correlate with responses to faces, non-face objects and places. **(a)** Eccentricity mapping with location of face patches (black contours) and color-biased regions (white contours) overlaid. Regions that were more responsive to a central flickering checkerboard disc (radius = 3.5°) are shown in blue and cyan; regions that were more responsive to stimulation with an annulus of flickering checkers (radius = 3.5–20°) shown in orange and red. **(b)** Representative time courses for the eccentricity stimulus in V1, IT, the central representation in IT, the peripheral representation in IT, the face patches, the color-biased regions and the place-biased regions. **(c)** Bar plots quantifying the responses to central and peripheral stimulation in the various ROIs. Error bars represent s.e.m. ($N = 4$ hemispheres). **(d,e)** Responses to objects **(d)** and colors **(e)** in ROIs defined using the eccentricity mapping. Error bars represent s.e.m. ($N = 4$ hemispheres).

representation, a clear increase in modulation to place stimuli in the peripheral representation, and an intermediate response to non-face and non-place images in the mid-peripheral representation, confirming results obtained in humans²³. Across the eccentricity representations, responses to color consistently showed a subtle luminance bias (all values are below 0; **Fig. 7e**). These results are consistent with the existence of at least three coarse representations of the visual field arranged sequentially along the posterior-anterior axis of the temporal lobe and suggest that the visual field maps found in early visual areas extended into IT, where they have adopted a functional specialization: central representations taking on computations dependent on high acuity such as face processing, mid-peripheral representations performing computations on non-face and non-place objects (and color), and peripheral representations executing computations related to processing places.

DISCUSSION

We addressed the functional organization of IT by measuring fMRI responses to colored gratings, achromatic images (faces, vegetables and fruit, body parts, and places) and retinotopic checkerboard stimuli in two monkey subjects. Color-biased regions and face patches were mostly non-overlapping, consistent with neurophysiological and psychophysical evidence suggesting that color and faces are processed by largely independent networks^{38,39}. Faces and colors share little low-level similarity, so the limited overlap may not be surprising. However, color-biased regions appeared to be yoked to face patches, with color-biased regions typically being ventral to face patches. Clusters of color-biased regions and face patches were located at four main sites along the posterior-anterior axis of IT (**Fig. 3**). **Figure 8** shows nodes identified in anatomical studies¹⁸ superimposed on a lateral view of the brain showing the functional activation. The correspondence is notable, suggesting that anatomical architecture underlies the observed functional organization.

Although anatomical evidence suggests the stages are not linked in a strictly feedforward fashion¹⁸ (and the functional data are not inconsistent with a rich set of interconnections between stages and pathways), we found that color selectivity increased for more anterior regions (**Fig. 3c**), matching increases in selectivity of face patches⁵. The systematic relationship of the face patches and color-biased regions, their regular spacing, and the progressive increase in selectivity in both systems along the posterior-anterior axis suggests that IT comprises not only a broad hierarchy of parallel processing streams, but also that the stages across parallel channels are subject to the same organizing principle of ~4 'areas'. Our results suggest that these areas each comprise a complete functional representation of objects, and are possibly homologous to the discrete areas found in early visual cortex.

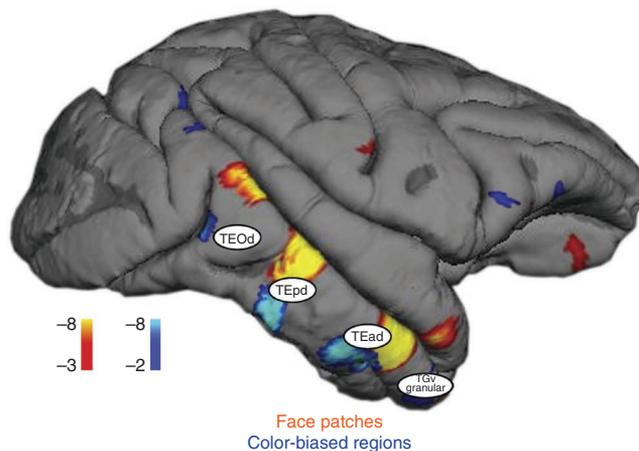


Figure 8 Projection of the previously described anatomical designations¹⁸ on the lateral surface of M1 on which has been painted the face-biased (orange and red) and color-biased activation (blue and cyan) patterns.

The findings have implications for understanding how IT evolved and shed light on a neurological puzzle, as discussed below.

Responses to images of non-face objects were significant across much of IT, but were greatest in regions outside of face patches and color-biased regions; moreover, place-biased activation was strongest in regions outside of body-biased, face-biased and color-biased regions. These results provide additional support for the notion of several parallel routes through IT. We acknowledge, however, that one might not expect the multi-stage conception of IT to manifest as three or four patches of activation to every image class tested—other factors, such as the nature of the computations performed in each stage and contextual interactions⁹, will influence the degree to which a given stage is activated. Nonetheless, regions showing response biases to images of bodies/body parts were found in three main clusters, overlapping and adjacent to face patches (Fig. 6a). This arrangement predicts that impairments in perception of bodies caused by strokes should often be accompanied by impairments in face perception, which appears to be true for some lesions of posterior IT in humans⁴⁰.

Color is an important component of objects; the extent to which color and form are processed by the same neural circuits has been a long-standing question in visual neuroscience⁴¹. Although we found evidence for parallel routes for color and form in IT, the color-biased regions were modulated by object shape (Fig. 6), providing a neural basis for the oft-cited role of color in mediating shape. These results mesh well with the nuanced understanding yielded by psychophysics, suggesting that luminance cues are the primary determinant of object shape under most circumstances, but that color cues can occasionally be diagnostic (for example, yellow = banana). The relative contribution to shape processing of the neurons in color-biased regions compared with neurons in intervening tissue awaits further study, as does the specific contribution to color made at each stage across IT. One clue may come from anatomical connection studies, which shows that the different parallel routes through IT connect differentially to different subcortical structures. The most ventral route, coinciding with the peak place-biased activation, connects with the hippocampus⁴², providing a channel by which place information could contribute to navigation computations. The strong connections between anterior IT with the amygdala, nucleus accumbens, medial temporal cortex, orbitofrontal cortex and ventrolateral prefrontal cortex¹⁸ provide tantalizing hints to the neural basis for the behavioral relevance of colors and faces, including their emotional valence, ability to elicit pleasure and role in visual-object categorization.

In humans, object-processing cortex is organized according to a coarse map of eccentricity, an extension of the retinotopic map in early visual cortex²³. It has not been clear whether this principle applies across primates. Our results provide fMRI confirmation of electrophysiological evidence^{13,24,25} for a retinotopic map in PIT; moreover, we found additional representations of the visual field in CIT and AIT (Fig. 7). The representations of the central visual field corresponded to locations of the face patches, supporting conclusions drawn from human imaging experiments that central representations in object-processing cortex have evolved a specialization for computations that depend on high spatial acuity²³. Our findings also confirm the result obtained in humans that peripheral representations overlap with regions showing stronger responses to places²³. Color-biased regions were located in an intermediate zone in the eccentricity maps. Together with work in humans, our results suggest that eccentricity maps are a unifying rule that governs the organization of IT across primates.

The results of our eccentricity mapping experiment are surprising because IT in monkeys is not thought to be organized retinotopically. Experiments in humans have shown that the eccentricity representation

in IT may reflect an organization for the real-world size; the region in IT that is activated by images of large objects is offset from the region activated by images of small objects, even when the images themselves occupy the same fraction of the visual field⁴³. These results suggest that IT does not possess an explicit retinotopic representation, but rather one that mirrors the scale of objects recorded in cognition. Our eccentricity results are parsimonious with this hypothesis if one considers the peripheral eccentricity stimulus as a 'large' object and the foveal stimulus as a small one. According to this hypothesis, the cognitive scale of colors is intermediate between faces (small objects) and places (large objects).

Our results may shed light on the evolutionary mechanisms that brought about the expansion of cerebral cortex. The eccentricity map in human IT suggests that this region initially arose by extension of retinotopic cortex²³. The multi-stage functional organizational scheme uncovered here may reflect the process by which IT subsequently expanded during evolution. The alternating pattern of horizontal and vertical meridian representations that define early retinotopic visual areas (Fig. 1) is evidence that these areas arose by mirror-image duplication of an ancestral brain region^{44,45}. The systematic repetition of adjacent color- and face-responsive regions suggests that the temporal cortex expanded by a similar process, through duplication of a single ancestral area that may have contained a complete functional representation of objects. The sequence of multiple visual-field representations in IT provides additional support for this theory. This is an appealing idea because it would invoke a single mechanism to account for the expansion of much of cerebral cortex.

The historically minded reader may recall the theory that V4 is the 'color area'. The seminal neurophysiology⁴⁶ underlying this idea led to the notion that acquired cerebral achromatopsia comes about because of a lesion of the human homolog of monkey V4 (for a review, see ref. 12). This deceptively simple interpretation has been complicated by three observations. First, V4 is not a homogenous area and is not uniquely specialized for processing color⁴⁷. Second, lesions of V4 in monkeys do not cause selective (or substantial) long-term impairments in color (for a review, see ref. 48). Third, cerebral achromatopsia in humans is often accompanied by other visual deficits, specifically prosopagnosia⁴⁹. That achromatopsia can manifest in the context of largely preserved visual object perception suggests some degree of functional specialization in the cortex. But it has been difficult to reconcile such functional specialization with the marked concordance of prosopagnosia and achromatopsia, especially in light of the physiological and psychophysical evidence showing that face processing and color processing are likely handled by largely independent circuits. Although there appears to be little single-unit evidence linking color and face processing, the anatomical pairing of color and face regions that we observed could account for the clinical observations if humans share a similar IT organization as monkeys, as strokes typically involve a region of tissue that is larger than the scale of a single color-biased region. Which specific region is the one whose compromise leads to achromatopsia remains unknown.

METHODS

Methods and any associated references are available in the [online version of the paper](#).

Note: Any Supplementary Information and Source Data files are available in the [online version of the paper](#).

ACKNOWLEDGMENTS

fMRI data were analyzed using the jip toolbox generously provided by J. Mandeville (Massachusetts General Hospital, <http://www.nitrc.org/projects/jip/>), and analysis

scripts compiled by S. Moeller (RWTH Aachen University). We thank J. Maunsell, D. Hubel and M. Livingstone for research space, helpful discussions and assistance with animal protocols. A. Rehding, N. Kanwisher, J. DiCarlo and P. Mayo provided valuable comments on the manuscript. B.R.C. gratefully acknowledges seminal conversations that took place with D. Tsao, with whom the idea to map multiple stimulus dimensions in IT was conceived and initially begun. We thank Y. Liu for implementing the simulation shown in **Figure 5** and help with MRI analysis. We thank J. McCready and E. Erzsinger, who designed and built the fMRI primate chair, and C. Stoughton and G. Gagin, who trained the animals. We thank L. Wald and A. Mareyam for providing the four-channel magnetic resonance coil. We thank W. Vanduffel, J. Mandeville, J. Polimeni, T. Witzel, M. O'Hara, G. Khandewal, M. Histed and P. Mayo for help running the experiments and conducting the analyses. This paper is dedicated to the memory of our mentor and beloved friend, David Hubel, who died September 22, 2013 (age 87). This work was supported by grants from the National Science Foundation (0918064), the US National Institutes of Health (EY023322), the Whitehall Foundation, the Radcliffe Institute for Advanced Study (Harvard University) and Wellesley College. This research was carried out at the Athinoula A. Martinos Center for Biomedical Imaging at the Massachusetts General Hospital, using resources provided by the Center for Functional Neuroimaging Technologies (P41EB015896) and a P41 Biotechnology Resource grant supported by the National Institute of Biomedical Imaging and Bioengineering. This work also involved the use of instrumentation supported by the US National Institutes of Health Shared Instrumentation Grant Program and/or High-End Instrumentation Grant Program (S10RR021110).

AUTHOR CONTRIBUTIONS

R.L.-S. carried out the experiments and analyzed the data. B.R.C. designed the research, carried out the experiments, analyzed the data and wrote the manuscript.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Reprints and permissions information is available online at <http://www.nature.com/reprints/index.html>.

- DiCarlo, J.J., Zoccolan, D. & Rust, N.C. How does the brain solve visual object recognition? *Neuron* **73**, 415–434 (2012).
- Tsao, D.Y.F.W., Knutsen, T.A., Mandeville, J.B. & Tootell, R.B.H. Faces and objects in macaque cerebral cortex. *Nat. Neurosci.* **6**, 989–995 (2003).
- Pinsk, M.A., DeSimone, K., Moore, T., Gross, C.G. & Kastner, S. Representations of faces and body parts in macaque temporal cortex: a functional MRI study. *Proc. Natl. Acad. Sci. USA* **102**, 6996–7001 (2005).
- Ku, S.P., Tolia, A.S., Logothetis, N.K. & Goense, J. fMRI of the face-processing network in the ventral temporal lobe of awake and anesthetized macaques. *Neuron* **70**, 352–362 (2011).
- Freiwald, W.A. & Tsao, D.Y. Functional compartmentalization and viewpoint generalization within the macaque face-processing system. *Science* **330**, 845–851 (2010).
- Srihasam, K., Mandeville, J.B., Morocz, I.A., Sullivan, K.J. & Livingstone, M.S. Behavioral and anatomical consequences of early versus late symbol training in macaques. *Neuron* **73**, 608–619 (2012).
- Op de Beeck, H.P., Deutsch, J.A., Vanduffel, W., Kanwisher, N.G. & DiCarlo, J.J. A stable topography of selectivity for unfamiliar shape classes in monkey inferior temporal cortex. *Cereb. Cortex* **18**, 1676–1694 (2008).
- Bell, A.H., Hadj-Bouziane, F., Frihauf, J.B., Tootell, R.B. & Ungerleider, L.G. Object representations in the temporal cortex of monkeys and humans as revealed by functional magnetic resonance imaging. *J. Neurophysiol.* **101**, 688–700 (2009).
- Cox, D., Meyers, E. & Sinha, P. Contextually evoked object-specific responses in human visual cortex. *Science* **304**, 115–117 (2004).
- Stoughton, C.M., Lafer-Sousa, R., Gagin, G. & Conway, B.R. Psychophysical chromatic mechanisms in macaque monkey. *J. Neurosci.* **32**, 15216–15226 (2012).
- Tootell, R.B., Nelissen, K., Vanduffel, W. & Orban, G.A. Search for color 'center(s)' in macaque visual cortex. *Cereb. Cortex* **14**, 353–363 (2004).
- Conway, B.R. & Tsao, D.Y. Color architecture in alert macaque cortex revealed by fMRI. *Cereb. Cortex* **16**, 1604–1613 (2006).
- Conway, B.R., Moeller, S. & Tsao, D.Y. Specialized color modules in macaque extrastriate cortex. *Neuron* **56**, 560–573 (2007).
- Harada, T. *et al.* Distribution of colour-selective activity in the monkey inferior temporal cortex revealed by functional magnetic resonance imaging. *Eur. J. Neurosci.* **30**, 1960–1970 (2009).
- Katsuyama, N. *et al.* Cortical activation during color discrimination task in macaques as revealed by positron emission tomography. *Neurosci. Lett.* **484**, 168–173 (2010).
- Iwai, E. & Yukie, M. Amygdalofugal and amygdalopetal connections with modality-specific visual cortical areas in macaques (*Macaca fuscata*, *M. mulatta* and *M. fascicularis*). *J. Comp. Neurol.* **261**, 362–387 (1987).
- Van Essen, D.C., Felleman, D.J., DeYoe, E.A., Olavarria, J. & Knerim, J. Modular and hierarchical organization of extrastriate visual cortex in the macaque monkey. *Cold Spring Harb. Symp. Quant. Biol.* **55**, 679–696 (1990).
- Kravitz, D.J., Saleem, K.S., Baker, C.I., Ungerleider, L.G. & Mishkin, M. The ventral visual pathway: an expanded neural framework for the processing of object quality. *Trends Cogn. Sci.* **17**, 26–49 (2013).
- Grill-Spector, K. & Malach, R. The human visual cortex. *Annu. Rev. Neurosci.* **27**, 649–677 (2004).
- Kanwisher, N. Functional specificity in the human brain: a window into the functional architecture of the mind. *Proc. Natl. Acad. Sci. USA* **107**, 11163–11170 (2010).
- Hadjikhani, N., Liu, A.K., Dale, A.M., Cavanagh, P. & Tootell, R.B. Retinotopy and color sensitivity in human visual cortical area V8. *Nat. Neurosci.* [see comments] **1**, 235–241 (1998).
- Bartels, A. & Zeki, S. The architecture of the colour centre in the human visual brain: new results and a review. *Eur. J. Neurosci.* **12**, 172–193 (2000).
- Hasson, U., Harel, M., Levy, I. & Malach, R. Large-scale mirror-symmetry organization of human occipito-temporal object areas. *Neuron* **37**, 1027–1041 (2003).
- Boussaoud, D., Desimone, R. & Ungerleider, L.G. Visual topography of area TEO in the macaque. *J. Comp. Neurol.* **306**, 554–575 (1991).
- Yasuda, M., Banno, T. & Komatsu, H. Color selectivity of neurons in the posterior inferior temporal cortex of the macaque monkey. *Cereb. Cortex* **20**, 1630–1646 (2010).
- Derrington, A.M., Krauskopf, J. & Lennie, P. Chromatic mechanisms in lateral geniculate nucleus of macaque. *J. Physiol. (Lond.)* **357**, 241–265 (1984).
- Koida, K. & Komatsu, H. Effects of task demands on the responses of color-selective neurons in the inferior temporal cortex. *Nat. Neurosci.* **10**, 108–116 (2007).
- Matsumura, T., Koida, K. & Komatsu, H. Relationship between color discrimination and neural responses in the inferior temporal cortex of the monkey. *J. Neurophysiol.* **100**, 3361–3374 (2008).
- Stoughton, C.M. & Conway, B.R. Neural basis for unique hues. *Curr. Biol.* **18**, R698–R699 (2008).
- Conway, B.R. & Tsao, D.Y. Color-tuned neurons are spatially clustered according to color preference within alert macaque posterior inferior temporal cortex. *Proc. Natl. Acad. Sci. USA* **106**, 18034–18039 (2009).
- Tsao, D.Y., Moeller, S. & Freiwald, W.A. Comparing face patch systems in macaques and humans. *Proc. Natl. Acad. Sci. USA* **105**, 19514–19519 (2008).
- Cavanagh, P. & Leclerc, Y.G. Shape from shadows. *J. Exp. Psychol. Hum. Percept. Perform.* **15**, 3–27 (1989).
- Rolls, E.T. Face processing in different brain areas, and critical band masking. *J. Neurophysiol.* **2**, 325–360 (2008).
- Van Essen, D.C. A tension-based theory of morphogenesis and compact wiring in the central nervous system. *Nature* **385**, 313–318 (1997).
- Conover, W.J. *Practical Nonparametric Statistics, 2nd edn.* (John Wiley & Sons, New York, 1980).
- Karnath, H.O. New insights into the functions of the superior temporal cortex. *Nat. Rev. Neurosci.* **2**, 568–576 (2001).
- Baker, C.I. *et al.* Visual word processing and experiential origins of functional selectivity in human extrastriate cortex. *Proc. Natl. Acad. Sci. USA* **104**, 9087–9092 (2007).
- Tanaka, K., Saito, H., Fukada, Y. & Moriya, M. Coding visual images of objects in the inferotemporal cortex of the macaque monkey. *J. Neurophysiol.* **66**, 170–189 (1991).
- Kemp, R., Pike, G., White, P. & Musselman, A. Perception and recognition of normal and negative faces: the role of shape from shading and pigmentation cues. *Perception* **25**, 37–52 (1996).
- Moro, V. *et al.* The neural basis of body form and body action agnosia. *Neuron* **60**, 235–246 (2008).
- Conway, B.R. *et al.* Advances in color science: from retina to behavior. *J. Neurosci.* **30**, 14955–14963 (2010).
- Iwai, E. & Yukie, M. A direct projection from hippocampal field CA1 to ventral area TE of inferotemporal cortex in the monkey. *Brain Res.* **444**, 397–401 (1988).
- Konkle, T. & Oliva, A. A real-world size organization of object responses in occipitotemporal cortex. *Neuron* **74**, 1114–1124 (2012).
- Allman, J.M. & Kaas, J.H. A crescent-shaped cortical visual area surrounding the middle temporal area (MT) in the owl monkey (*Aotus trivirgatus*). *Brain Res.* **81**, 199–213 (1974).
- Sereno, M.I. & Allman, J.M. Cortical visual areas in mammals. In A. G. Leventhal (ed.) *The Neural Basis of Visual Function* (ed. Leventhal, A.G.) 160–172 (Macmillan, London, 1991).
- Zeki, S. The representation of colours in the cerebral cortex. *Nature* **284**, 412–418 (1980).
- Roe, A.W. *et al.* Towards a unified theory of visual area V4. *Neuron* **74**, 12–29 (2012).
- Heywood, C.A. & Kennerly, R.W. Achromatopsia, color vision, and cortex. *Neurosci. Clin.* **21**, 483–500 (2003).
- Bouvier, S.E. & Engel, S.A. Behavioral deficits and cortical damage loci in cerebral achromatopsia. *Cereb. Cortex* **16**, 183–191 (2006).

ONLINE METHODS

Functional magnetic resonance imaging. Two male rhesus macaques (7–8 kg), M1 and M2, pair housed in standard 12:12 light-dark cycle and given food *ad libitum* were scanned at Massachusetts General Hospital Martinos Imaging Center in a 3-T Tim Trio scanner (Siemens). Magnetic resonance images were acquired with a custom-built four-channel magnetic resonance coil system with AC88 gradient insert, which increases the signal-to-noise ratio (SNR) by allowing very short echo times⁵⁰, providing 1-mm³ spatial resolution and good coverage of the temporal lobe (Supplementary Fig. 9). We used standard echo planar imaging (repetition time = 2 s, 96 × 96 × 50 matrix, 1-mm³ voxels; echo time = 13 ms). In all the experiments described here, data were obtained after administration of an intravenous contrast agent, which further enhances SNR¹³ (Supplementary Fig. 1a). No power analysis was done to determine sample size a priori. Our sample sizes are similar to those used in previous studies; data from both animals and all hemispheres was included in the analysis. Using juice rewards, monkeys were trained to sit in a sphinx position in a custom-made chair placed inside the bore of the scanner and to fixate a central spot presented on a screen 49 cm away. Head position was maintained using surgically implanted custom-made plastic head posts (see surgical details below). An infrared eye tracker (ISCAN) was used to monitor eye movements, and animals were only rewarded for maintaining their gaze within ~1 degree of the central fixation target. Magnetic resonance signal contrast was enhanced using a MION (Feraheme, 8–10 mg per kg of body weight, diluted in saline, AMAG Pharmaceuticals), injected intravenously into the femoral vein just before scanning. Decreases in MION signals correspond to increases in BOLD response; time course traces in all figures have been vertically flipped to facilitate comparison with conventional BOLD in which upward deflections correspond to increases in neural activity. All imaging and surgical procedures conformed to local and US National Institutes of Health guidelines and were approved by the Harvard Medical School and Wellesley College Institutional Animal Care and Use Committees. The same pair of monkeys were used in color-detection experiments that show macaque monkeys have similar psychophysical chromatic mechanisms to humans¹⁰.

Surgery. To maintain head stabilization during testing, plastic (Delrin) headposts were implanted using standard sterile surgical procedures. Animals were anesthetized with ketamine (15 mg per kg, intramuscular) and xylazine (2 mg per kg, intramuscular) and given atropine (0.05 mg per kg, intramuscular) to reduce salivary fluid production. Depth of anesthesia was maintained with 1–2% isoflurane (vol/vol) in oxygen administered with a tracheal tube. Animals were given a pre-emptive dose of buprenorphine (0.005 mg per kg, intramuscular) and flunixin (1.0 mg per kg, intramuscular) as analgesics and a prophylactic dose of an antibiotic (Baytril, 5 mg per kg, intramuscular). Antibiotic was administered again 1.5 h into surgery; buprenorphine and flunixin were given for 48 h post-operatively. During the surgery, the animals were placed in a stereotaxic holder and sterile techniques were used to insert ceramic screws and custom made inverted plastic “T” bolts into the skull. A head post was placed on the surface of the skull and cemented in place to the skull, anchored by the screws and T-bolts using dental acrylic. The animals were closely monitored after surgery for signs of pain or infection and treated accordingly. The animals recovered for 2–3 months before resuming training.

Color stimuli. For fMRI experiments, visual stimuli were displayed on a screen (41° × 31°) 49 cm in front of the animal using a JVC-DLA projector (1,024 × 768 pixels). All stimuli spanned the entire screen and contained a small central fixation cross to engage fixation. Stimuli were presented in a block-design procedure. Responses to color, objects, eccentricity and meridians were run in separate sessions. Color stimuli were generated in DKL color space²⁶, and calibrated using spectral readings taken with a PR-655 spectroradiometer (Photo Research) using procedures developed by T. Hansen (personal communication). The spectra were multiplied with the Judd-revised Commission Internationale de l’Eclairage (CIE) 1931 color matching functions to derive CIE xyY coordinates of the monitor primaries⁵¹, and cone excitation was calculated using the Smith and Pokorny cone fundamentals⁵². 12 color directions were chosen, evenly sampling the azimuth of the equiluminant plane of DKL space (Fig. 2a and Supplementary Fig. 1b). Chromaticity coordinates and cone contrasts differed slightly between experiments due to projector bulb performance changing over time. Stimuli were re-calibrated 1 or 2 d before each experiment.

Color stimuli were presented as equiluminant single-color vertically oriented trapezoid-wave gratings (Fig. 2a). The gratings (2.9 cycles per degree, drifting

0.75 cycles per s) were drifted back and forth, switching directions every 2 s, in 32-s blocks interleaved with equiluminant neutral full-field gray (32 s), in two stimulus orders (run 1: colors 1, 3, 5, black/white, 7, 9, 11; run 2: 2, 4, 6, black/white, 8, 10, 12, black/white; black/white achromatic luminance gratings of 10% contrast).

Color selectivity has historically been a challenge to define¹⁴. The operational definition adopted presently (responses to equiluminant color > responses to achromatic luminance) is the most widely used. The definition assumes that grays (including black and white) are not colors, yet luminance (gray value) is clearly one dimension of color space. If luminance contributes to color, then we might expect a color region (or cell) to respond to both equiluminant color and luminance⁵³; a non-color region, such as area MT, might respond only to luminance^{12,13}. Any region that shows a greater response to equiluminant color than to achromatic luminance matched in cone contrast can therefore be interpreted as likely making a contribution to hue processing (that dimension of color called by color names red, green, etc.). The challenge in defining color selectivity may account for the lower selectivity of the color regions than the face regions when assessed with fMRI, even though both sets of regions are functionally selective. For example, the color selectivity in V4 and PIT color-biased regions is relatively low (Fig. 3c), but fMRI-guided single-unit recording shows that the underlying neurons are highly color selective when assessed with rigorous color-selectivity criteria including tolerance to luminance changes^{13,30}.

The color stimuli assume that monkeys and humans have the same cone fundamentals, which is justified because the cone absorption spectra are very similar in the two species. Nonetheless, there may be subtle differences in pre-retinal pigments between the species, which could introduce luminance artifacts. To mitigate these, color responses were defined by using the two colors that elicited the weakest response in area MT, an area most sensitive to luminance contrast. These two colors were 7 and 8 (Fig. 2a), colors that cause little activation of S cones and are therefore least likely to be susceptible to chromatic aberration. Although these considerations are important, the conclusions of the paper are not affected if regions of interest were defined by the most significant activation produced by the response to the other colors tested (Supplementary Fig. 2b); moreover, the pattern of color-biased activation was not critically determined on the basis of significance thresholds (Supplementary Fig. 9c).

Object stimuli. The stimuli were achromatic photographs of faces (humans and monkeys), body parts, objects (fruit and vegetables) and familiar places, along with scrambled versions of these images, similar to those used previously². Example images are given in Figure 6. The images were shown in 16 32-s blocks (16 repetition times per block, repetition time = 2 s, 2 images per repetition) presented in one run sequence: [random noise], [frontal unfamiliar face], [faces scrambled], [bodies+hands+limbs], [bodies scrambled], [familiar frontal faces], [scrambled faces], [vegetables+fruit], [vegetables+fruit scrambled], [familiar frontal faces], [faces scrambled], [places], [places scrambled], [3/4 faces], [3/4 faces scrambled], [random noise]. The images occupied a 6° square centered at the fixation spot, and were surrounded by a neutral gray. The images were matched in average luminance to the neutral gray, maintaining roughly constant average luminance (~25 cd m⁻²) throughout the stimulus sequence. For the first frontal images block, we used 16 unique images of unfamiliar faces (8 human, 8 monkey); the image sequence was repeated during the block (2 × 1 s per image × 16 images per block = 32 s per block). The bodies/body part block comprised 32 unique images of monkey and human bodies (no heads/faces) and body parts. The second frontal faces block comprised 12 unique images of the animal’s cage mate. The fruit/vegetable block comprised 16 unique images (whole apple, cut apple, tomato, unpeeled banana, half-peeled banana, whole pair, cluster of grapes, whole lemon, half orange, whole peach with leaves, whole bell pepper, half a pear, half tomato, whole pumpkin, whole strawberry, cluster of cherry tomatoes). The third frontal faces block comprised 16 unique images of the animal itself (images were mirror reversed; mirrors were provided in the cages). The familiar places block comprised 13 unique images of the animal housing room, preparation room and laboratory. The 3/4 faces comprised 15 images of unfamiliar human and monkeys; data obtained during this block were not analyzed.

We obtained 18 total runs for each animal. As shown in Figure 6b, responses in the face patches were of comparable magnitude to the different blocks of frontal faces (unfamiliar and familiar), but the color ROIs and the other IT regions

showed a lower response to the familiar faces than the unfamiliar faces. These differences may reflect differences in the way animals perceive recognizable faces, although we do not have sufficient data to address this issue. The face patches and quantification of face responses in all ROIs were made on the basis of the responses to unfamiliar faces.

Areal boundaries and eccentricity mapping. Meridian mapping was performed using wedges of a black-and-white checkerboard (flickering 1 Hz) radiating out from the fixation spot along the vertical and horizontal meridians. Similar stimuli have been used to determine the vertical and horizontal meridians that define retinotopic visual areas^{54,55}. The stimulus sequence consisted of 32 s of horizontal wedges (99% luminance contrast, occupying 30° visual angle), followed by 32 s of uniform neutral gray, followed by 32 s of vertical wedges (occupying 60° visual angle), followed by 32 s of neutral gray, and so on for a total of four presentations of horizontal wedges and four presentations of vertical wedges. The functional border assignments obtained with this method correspond well to areal boundaries of a standard macaque atlas⁵⁶. Eccentricity mapping was performed using discs and annuli of a colored checkerboard. Four 32-s blocks, interleaved with 32-s blocks of neutral gray, included a disc with radius of 1.5°, and annuli extending from 1.5–3.5°, 3.5–7° and 7–20°. One set of stimuli comprised colors 1 and 7 (activating L and M cones) and another set of stimuli comprised colors 4 and 10 (activating S cones). **Figure 7b** shows results to the L-M stimuli; responses to the S stimuli followed a comparable eccentricity pattern, but were of lower magnitude. For the purposes of quantification (**Fig. 7c**), responses to the central two stimuli were combined, and responses to the peripheral two stimuli were combined.

Three-dimensional simulation. In the simulation, the locations of the face-biased regions were fixed according to the anatomical location determined experimentally, whereas the locations of the color regions were randomly assigned in the IT volume. The IT volume and regions of interest excluded white matter. To ensure the test was as stringent and as biologically relevant as possible, voxels within a given region remained contiguous, and two regions could overlap. Significance thresholds were set stringently to yield roughly the same number of color-selective and face-selective voxels in each hemisphere (left hemisphere/right hemisphere, M1 color regions: $P = 10^{-5}, 10^{-9}$; face patches: $P = 10^{-13}, 10^{-10}$; M2 color regions: $P = 10^{-3}, 10^{-8}$; face patches: $P = 10^{-8}, 10^{-7}$). For each face-biased voxel, we then measured the location of the nearest color-biased voxel generated by the random simulation; note that the location of the nearest voxel in the three-dimensional simulation could be closer than if the simulation were restricted to a two-dimensional cortical surface because of cortical folding. The histograms were normalized by the total number of comparisons made.

fMRI data processing. A total of 40,480 functional volumes were obtained during nine scan sessions in the two animals. For the color localizer, 13,600 volumes in M1, 7,072 in M2. For the eccentricity mapping, 1,904 in M1, 4,080 for M2. For the object mapping, 4,608 in M1, 4,608 in M2. For the meridian mapping, 3,328 in M1, 1,280 in M2. High-resolution anatomical scans ($0.35 \times 0.35 \times 0.35 \text{ mm}^3$ voxels) were obtained for each animal while it was lightly sedated. Significance maps were painted on inflated surfaces of each animal's anatomical volume.

Data analysis was performed using FREESURFER and FS-FAST software (<http://surfer.nmr.mgh.harvard.edu/>), the custom “jip” toolkit provided by J. Mandeville (<http://www.nitrc.org/projects/jip/>), and custom scripts written in Matlab. The surfaces of the high-resolution structural volumes were reconstructed and inflated using FREESURFER; functional data were motion corrected with the AFNI motion correction algorithm⁵⁷, spatially smoothed with a Gaussian kernel (full-width at half maximum = 2 mm; except for **Supplementary Fig. 2c**, which used no spatial smoothing), and registered to each animal's own anatomical volume using jip. Following standard alert monkey fMRI processing techniques, images were first normalized to correct for signal intensity changes and temporal drift, and t tests uncorrected for multiple comparisons were performed to construct statistical activation maps based on a general linear model^{2,7,8,13,14,31}. Activation was thresholded at significance levels indicated in the figures by a color scale bar (values show the common logarithm of the probability of error, and are uncorrected). Activation maps were then projected on high resolution anatomical volumes and surfaces. Time courses were calculated by first detrending the

fMRI response. The temporal drift often associated with fMRI signals was modeled by a second-order polynomial

$$x(t) = S(t) + at^2 + bt + c$$

where $x(t)$ is the raw fMRI signal, and $S(t)$ is the detrended signal. The coefficients a , b and c were calculated using the Matlab function *polyfit*. The percentage deviation of the fMRI signal, $s'(t)$, reported as the y axis values in the time course traces, was calculated by

$$s'(t) = 100 \times \frac{s(t) - \bar{s}}{\bar{s}}$$

where $s(t) = \text{smoothed}(S(t) + c)$, $t = 1, 2, \dots, N$, \bar{s} is the mean of $s(t)$, and N is the number of repetition times in the experiment. The constant c was added to $S(t)$ to obtain $s(t)$ to avoid dividing by zero errors. Smoothing was performed using the MATLAB function *smooth* with a moving average of 3.

Selectivity analysis and ROI definition. Regions of interest were defined for each hemisphere on the basis of significant activation during each a functional localizer. Color-biased ROIs were defined by comparing responses to the achromatic contrast gratings with responses to the two chromatic gratings from the stimulus set that elicited the weakest response in MT, colors 7 and 8 (these colors would be subject to the least amount of luminance artifact). Face-patches were defined by comparing responses to achromatic images of faces with responses to achromatic images of bodies^{2,31}. As discussed below, significance thresholds to define ROIs in **Figures 3c, 5a, 6b–d, f, g** and **7c** were set stringently to yield roughly the same number of color-selective and face-selective voxels in each hemisphere (left hemisphere/right hemisphere, M1 color regions: $P = 10^{-5}, 10^{-9}$; face patches: $P = 10^{-13}, 10^{-10}$; M2 color regions: $P = 10^{-3}, 10^{-8}$; face patches: $P = 10^{-8}, 10^{-7}$), and ROIs were defined as contiguous voxels with suprathreshold significance values. Note that adjusting the P value thresholds did not radically alter the pattern of activation (**Supplementary Fig. 9c**). To avoid ‘double-dipping’⁵⁸, ROIs were defined using half the data runs from each localizer and selectivity indices were computed in those ROIs using the other half of the data (indices are described below). **Supplementary Figure 9a** shows the activation maps of all significant activation to the colors and achromatic gratings versus gray; the activation extends fairly continuously down the temporal lobe. Time courses are shown in **Supplementary Figure 9b** for the color biased regions, the intervening luminance biased regions (see **Supplementary Fig. 2a**), all of visually active IT, and the visually active voxels of IT that were neither luminance nor color biased. The results show that the patchiness in the color-biased response distribution reflects spatial clustering of responses and not holes in the magnetic resonance signal. The other IT regions described in **Figure 6b** were defined as the regions that responded most significantly to all non-face images > scrambled images, outside of the face patches and color-biased regions, and were restricted to include the same total number of voxels as the face-biased regions in anatomically designated regions PIT, CIT and AIT.

To compute the percent signal change reported in **Figures 6** and **7**, and used to compute the selectivity indexes reported in **Figure 3**, each ROI's response time course during the stimulus run was de-trended, smoothed and normalized by the response to the entire run (for example, **Fig. 6b**). The response during a given stimulus block was calculated using the average of the last seven repetition times of the stimulus block to avoid confounds introduced by the hemodynamic delay. Suppose that the stimulus occurred during the k^{th} block in a run. For the color and eccentricity experiments, the block immediately preceding and immediately following stimulus blocks would have been gray. We denote the response to the stimulus block as R^k , which is the mean of the response during the 9th to the 16th samples, and the response to the neighboring gray blocks as R^{k-1} and R^{k+1} . The percentage signal change of the color or eccentricity stimulus block was then defined by

$$Rc = 100 \times \left(R^k - \frac{R^{k-1} + R^{k+1}}{2} \right)$$

The percentage signal change to color was the average Rc value during presentation of colors 7 and 8; the percentage signal change for the eccentricity mapping experiment for the central representation was the average Rc value during presentation of the central disc and smallest annulus, and for the peripheral

representation was the average R_c value during presentation of the largest two annuli (Fig. 7b), averaging across the L-M and S conditions.

The percentage signal change to an object block was computed by subtracting the response to the scrambled condition for the given object class; the scrambled condition for a given object block immediately followed the block

$$Ro = 100 \times (R^k - R^{k+1})$$

Note that the calculation of percentage signal change allows the freedom for the values to be negative. For example, negative percentage signal changes to color represent a luminance bias (for example, Figs. 6g,h and 7e).

Selectivity indices in Figure 3 were computed on the basis of the preferred and non-preferred stimulus responses

$$\text{Selectivity} = \frac{Rp - Rnp}{Rp + Rnp}$$

For color, Rp = response to color gratings (colors 7 and 8); Rnp = response to achromatic gratings. For faces, Rp = response to faces, Rnp = response to bodies. Valid computations of selectivity require that Rp and Rnp are zero or positive^{59,60}. Thus, before computing selectivity, if either Rp or Rnp were negative, a value was added to both to raise them such that both were zero or positive. For example, if Rp was 5 and Rnp was -1, a value of 1 was added to both Rp and Rnp such that Rp was 6 and Rnp was 0. Responses (averaged over the last seven repetition times in a stimulus block) were calculated against the response to the intervening gray block. In Figure 3c, responses to all patches at a given posterior-anterior location have been averaged together, calculated against the mean response for the stimulus run and corrected for spurious negative fMRI signals. For 1/2 bars, half the data was used to define the regions and the other half was used to calculate the selectivity indices; selectivity in the 1/2 bar was determined by averaging across all functionally defined regions in all four hemispheres.

For the 1/2 bars, $N = 16$ color regions, 16 face regions (4 ROIs per hemisphere \times 4 hemispheres = 16); distributions were normal (Jarques-Bera goodness-of-fit tests), justifying the use of the t test. The remaining bars show selectivity at each location along the posterior→anterior axis calculated with the entire data set for maximum statistical power.

Normalizing activation maps across stimulus conditions. As Supplementary Figure 2 shows, depending on the colors chosen to generate the activation, we obtained different levels of significance. The broad conclusion of all the color experiments is that IT is generally responsive to color, although not all parts of it show the same degree of responsiveness and only some portions show overt color-bias. The range of significance levels achieved by different contrasts raises a challenge, namely how to relate the lack of homogeneity that is clearly evident in the color maps (using any color stimulus contrast) with that obtained using a stimulus generated with entirely different low-level features such as faces. To normalize the responses across disparate stimulus conditions, we identified as color selective a comparable total number of voxels as were identified as face selective by adjusting the significance cut-offs of the two maps (Fig. 5). Notably, the main results are not different if we use a different overall set of P values that yield higher or lower numbers of color-biased and face-biased voxels, so long as the number of color-biased voxels roughly matched the number of face-biased voxels and neither population of voxels entirely covered IT. This normalization is valid because the maps generated by using a range of P value cut-offs do not radically alter the pattern of activation in either the face condition or the color condition (Supplementary Fig. 9c).

Inter-animal variability and methodological considerations. The color-biased activity in central and anterior IT was more distinct than we found previously¹³, especially in the animal whose data is shown in Figure 1 of ref. 13. This animal showed no face patches in IT. Other technical differences exist between the present study and the earlier report; here we took advantage of technical developments that yield higher signal-to-noise (afforded by the use of the AC88 gradient insert, and the use of a four-channel coil system). In addition, there are methodological differences between the studies in the way in which colors were defined. Moreover, the previous animals were exposed to the intravenous contrast agent over many months. Chronic exposure can affect fMRI signals especially in the temporal lobe (unpublished observations), perhaps because of the accumulation

of hemosiderin in the smooth muscle walls of blood vessels. At the same time, the pattern of color-biased activity in V4 and surrounding areas (including V2) shown in Figure 2 appears to be weaker than we previously reported¹³. This discrepancy is attributed to differences in smoothing and in stimuli used to define color-biased regions: the earlier report used less smoothing, a broader range of colors, and colors of higher saturation. Using all colors besides colors 7 and 8 to generate the color-response pattern (rather than just colors 7 and 8 as in Fig. 2), and no spatial smoothing, we found a comparable pattern in V2 and V4 to that obtained previously (Supplementary Fig. 2c). The pattern of activation in V4 comprises smaller scale regions (globs) than those in IT. Supplementary Figure 2b shows that the pattern of color-biased activation in IT is consistent when using the different sets of colors, although IT shows globally stronger responses to the larger set of colors.

The results obtained in M1 and M2 were not identical. Such inter-animal variability is predicted from studies of the face patch system³¹. A larger study will be necessary in order to determine the extent of this variability. Toward this goal, we have recently obtained data in two additional animals using the AC88 insert and four-channel coil system; the data were obtained toward a different experimental objective and we used heterochromatic gratings to determine the color-biased regions. Nonetheless, the results in IT support our conclusions.

To get a greater sense of inter-animal variability amongst a broader number of animals and to determine the extent to which methodological differences influence the observed phenomena, we performed a re-analysis of data used to identify the color globs in V4/PITd¹³; these data were obtained using a single-channel coil without AC88 insert in four animals in which face patches had also been tested³¹. The color data were obtained using red-blue trapezoidal gratings in which the relative luminance of the red and blue was defined to elicit the minimum response in area MT¹²; other methodological details can be found in ref. 13. This approach differs from that used presently, in which equiluminant colors were defined using human photometric equiluminance, colors were equated to luminance gratings in cone contrast⁵³, and responses were defined using the colors that elicited the minimum response in area MT. These colors fell on the L-M chromatic axis, which would have the least chromatic aberration and therefore be least likely to possess luminance artifacts. The extent to which these experimental differences influenced the ability to observe the relationship between color-biased regions and face patches is unclear. First, establishing a color response is not trivial (see above), and is complicated further given that fMRI signals to color and luminance likely sum in a nonlinear way. Second, the extent to which the enhancement in SNR afforded by the four-channel coil and the AC88 is decisive is not known. Here we provide a summary of the analysis of the color-biased activation and face patches within IT in the data set used in ref. 13.

One of the four animals in the earlier study did not have face patches or color patches in IT—this appears to be a congenital anomaly—but did have robust color signals in V4/PIT (Fig. 1 of ref. 13). A second animal showed a close association of PLc, CLc, and ALc to PL, ML and AL in both hemispheres (Monchichi, Supplementary Fig. 10; see Fig. 2 in ref. 13); the slice prescription did not cover the most anterior portion of the temporal lobe that encompasses AM and AMc. The third animal showed an association between PLc, Cc, ALc, and PL, ML, AL/AF (Napoleon); this association was more marked in one hemisphere. In the fourth animal (Bert), the face patches were asymmetrically positioned in the two hemispheres, with patches displaced 3 mm posterior in the left compared to the right; in this animal there was an association between PLc, CLc and ALc, with PL, ML and AL in the right hemisphere but not the left, as the color patches were bilaterally symmetric. The results from these animals confirms the central finding described presently, although in many of these animals there is globally higher levels of activation to the “color” condition throughout much of retinotopic cortex (including V1) and the LGN, suggestive of contamination by a luminance artifact, and the spatial resolution of the color-biased activation in IT appears to be lower.

50. Kimmlingen, R. *et al.* An easy to exchange high performance head gradient insert for a 3T whole body MRI system: first results. *Proc. Intl. Soc. Magn. Reson. Med. Sci. Meet. Exhib.* **11**, 1630 (2004).

51. Hansen, T., Giesel, M. & Gegenfurtner, K.R. Chromatic discrimination of natural objects. *J. Vis.* **8**, 2 1–19 (2008).

52. Smith, V.C. & Pokorny, J. Spectral sensitivity of color-blind observers and the cone photopigments. *Vision Res.* **12**, 2059–2071 (1972).

53. Johnson, E.N., Hawken, M.J. & Shapley, R. Cone inputs in macaque primary visual cortex. *J. Neurophysiol.* **91**, 2501–2514 (2004).
54. Fize, D. *et al.* The retinotopic organization of primate dorsal V4 and surrounding areas: a functional magnetic resonance imaging study in awake monkeys. *J. Neurosci.* **23**, 7395–7406 (2003).
55. Lafer-Sousa, R., Liu, Y.O., Lafer-Sousa, L., Wiest, M.C. & Conway, B.R. Color tuning in alert macaque V1 assessed with fMRI and single-unit recording shows a bias toward daylight colors. *J. Opt. Soc. Am. A Opt. Image Sci. Vis.* **29**, 657–670 (2012).
56. Paxinos, G., Huang, X.-F. & Toga, A.W. *The Rhesus Monkey Brain in Stereotaxic Coordinates* (Academic Press, San Diego, 2000).
57. Cox, R.W. & Hyde, J.S. Software tools for analysis and visualization of fMRI data. *NMR Biomed.* **10**, 171–178 (1997).
58. Kriegeskorte, N., Simmons, W.K., Bellgown, P.S. & Baker, C.I. Circular analysis in systems neuroscience: the dangers of double dipping. *Nat. Neurosci.* **12**, 535–540 (2009).
59. Simmons, W.K., Bellgown, P.S. & Martin, A. Measuring selectivity in fMRI data. *Nat. Neurosci.* **10**, 4–5 (2007).
60. Baker, C.I., Hutchison, T.L. & Kanwisher, N. Does the fusiform face area contain subregions highly selective for nonfaces? *Nat. Neurosci.* **10**, 3–4 (2007).

