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

Published in final edited form as:
Nat Neurosci. 2013 December ; 16(12): 1870–1878. doi:10.1038/nn.3555.

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

Rosa Lafer-Sousa[1] and **Bevil R. Conway**[1,2,#]

[1]Neuroscience Program, Wellesley College

[2]Department of Neurobiology, Harvard Medical School

Abstract

Visual-object processing culminates in inferior temporal (IT) cortex. To assess the organization of IT, we measured fMRI responses in alert monkey to achromatic images (faces, fruit, bodies, places) and colored gratings. IT contained multiple color-biased regions, which were typically ventral to face patches and, remarkably, 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 suggest that IT comprises parallel, multi-stage processing networks subject to one organizing principle.

Inferior temporal cortex (IT) is a large expanse of tissue implicated in object perception. Anatomical studies commonly partition IT into posterior, central and anterior parts, but it is not clear whether these constitute distinct areas[1]. One popular theory is that the parts represent a hierarchical organization of information processing. This idea is supported by fMRI in monkeys, which shows a prominent face bias at 3–4 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 selectivity[5]. It is not known whether the face-patch system is an exceptional case, or is indicative of a general hierarchical organizational principle[6]. Previous reports have attempted to address this issue by testing for a systematic relationship amongst fMRI response patterns to different classes of objects[2,7,8]. In such experiments, it is a challenge to control low-level attributes, and contextual interactions may further complicate interpretation[9]. Color provides a useful tool to tackle the issue since 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 monkeys[10], and, as in humans, color stimuli activate multiple foci at several locations across IT[11–15]. But the relationship between activation patterns elicited by colors and those elicited by objects, particularly faces, has not been investigated.

[#]correspondence: bconway@wellesley.edu.

Author Contributions: RLS carried out the experiments and analyzed the data. BRC designed the research, carried out the experiments, analyzed the data and wrote the manuscript.

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 places[23]. A coarse retinotopy has been found in posterior IT of macaque monkey using microelectrode recording[13,24,25], but it is not thought that this organization extends to central and anterior IT[1]. It remains unclear whether the organizational principles found in human apply to monkey, an important issue that bears upon 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 V4. To determine this boundary, we used fMRI to map the visual meridian representations (Fig. 1; Supplementary Fig. 1a). In order to show the cortical activity buried within sulci, we computationally inflated the brain, revealing an alternating set of stripes corresponding to the boundaries of classic retinotopic areas (V1, V2, V3). The maps also show a clear horizontal meridian representation at the anterior boundary of V4 along both ventral and dorsal subdivisions. In addition, the maps show a vertical meridian representation within posterior IT, providing fMRI confirmation of a retinotopic area within PIT[24]. 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.

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 space[26], and were adjusted to yield roughly equal saturations in the perceptually uniform C.I.E. Luv space (Supplementary Fig. 1b). Luminance-biased responses were largely restricted to V3, V3A, MT and parts of V4 and IT (Supplementary Fig. 2a). Prominent color-biased responses were found within V2, V4 and at several locations spanning IT, as described below.

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 (TE). Posterior IT has been subdivided into dorsal and ventral components (PITd, PITv), while TE has been carved along both the anterior-posterior axis (Central/Anterior) and the dorsal-ventral axis, yielding four parts (CITd, CITv, AITd, AITv). The coarse retinotopic organization of PIT[24] is not thought to extend to CIT and AIT[1]. A synthesis of connectivity studies suggests that IT comprises at least four somewhat parallel routes[18]; 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 within a given animal and somewhat stereotyped across the two animals (M1, Fig. 2c; M2, Fig. 2d). In all four 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 within 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), within the frontal cortex, and in several regions along the posterior IT/V4 border, as described previously[12,13] (Fig. 2, PITd and V4v). 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 A-P 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 (e.g. Fig. 2c, slice −19.6). PVc was found in both hemispheres of both animals.

Despite gross stereotypy, we observed some inter-hemispheric and inter-animal variability. Both animals 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 animals, but appeared to be weak or non-existent in the right hemisphere of either animal. 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 (Fig. 3, M1, left hemisphere: the region marked “CLc” and the region touching ML). 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 animals. There was also evidence for a small color-biased region between CLc and ALc in the right hemisphere of both animals (Fig. 3a, M1). The time courses of the responses to individual colors within color-biased regions are given in Supplementary Fig. 3a.

The color-biased activity in central and anterior IT was more distinct than we found previously[13]; possible causes are discussed in the Methods. 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 reported[13]. This discrepancy is attributed to differences in smoothing and in stimuli used to define color-biased regions. Using all colors besides colors 7 & 8 to generate the color-response pattern (rather than just colors 7&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 within V4 comprises smaller scale regions (“globs”) 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 also identified with fMRI[14] and positron emission tomography[15]. 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 IT[25], 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.

Relationship between color architecture and face patches

We compared the responses to color with the responses to achromatic images of faces, objects (fruits/vegetables) and places in the same subjects. As in previous reports, the fMRI maps show three prominent clusters of face patches: PL, ML/MF and AL/AF[2], 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 within 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 subcortical areas (Supplementary Fig. 4) as previously documented using higher field strengths[4], demonstrating that the present techniques allow satisfactory assessment of activity over much of the brain.

Remarkably, 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 (e.g. CFc and AFc, Fig. 3a, M1, right hemisphere). The correspondence of color regions and face patches is clear when the activation patterns are displayed on the inflated brains, but the computational algorithms involved in inflating the brain necessarily introduce distortion. In sagittal slices from each hemisphere of one monkey through PL, ML and AL, there is a corresponding color-biased region beneath it (Fig. 3b, top left panels). The sagittal slices show one exception, an orphan color-biased region, V4v in the left hemisphere. The horizontal section through the anterior ventral pole of the temporal lobe shows that the AM face patch is also accompanied by a color-biased region, which we call AMc (anterior medial color, Fig. 3b). Coronal slices through the corresponding face patches provide confirmation of the relationship between color-biased regions and face patches (bottom panels, Fig. 3b). The yoking of the color-biased activation and the face patches across the temporal lobe is evident in the mosaics showing MION activation on the raw functional echo planar (EPI) coronal sections across the whole brain (M1, Fig. 4; M2, Supplementary Fig. 5; stereotaxic coordinates for the center of the functionally defined regions is 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 (inset 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. The result of averaging across all face patches shows a strong luminance bias, which predicts the luminance dependence of face processing[32]. 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 amongst face patches[5,33].

To quantify the spatial relationship of face and color regions, we constructed a 3-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 within IT, assuming the two sets are arranged randomly with respect to each other. In the simulation, we retained the 3-D shape of the cortical sheet because axonal connections could exert a powerful constraint on cortical development and

evolution[34]. We performed the simulation 1000 times, and computed a histogram of the distance from each color-biased voxel to the nearest face-biased voxel. The average separation (~4mm) was significantly closer than expected by chance (6–7mm; Mann–Whitney–Wilcoxon test M1, $p=1\times10^{-17}$, N = 278 voxel to voxel distances; M2, $p=5\times10^{-10}$, N = 247), a finding that was true when the analysis was performed on each hemisphere separately (Fig. 5 legend). 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 distributions[35]).

Responses to non-face objects

In addition to faces, we mapped the response to other types of objects (bodies/body parts, vegetables/fruit, places; Fig. 6). Significant responses to faces > scrambled faces (Fig. 6a) were found across much of IT, although the peak significance coincided with the face patches defined in Figure 3 (responses to faces > responses to body parts). The color-biased regions responded significantly 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: the cyan regions in all the panels of Figure 6a tend to avoid the color-biased regions, providing additional support for the hypothesis that color and faces are handled by largely distinct circuits. In addition, we found that highly significant responses to places tend to be ventral (and dorsal) to face patches and color-biased regions. Place-biased regions were also more prominent in posterior IT. The maps generated by some stimulus contrasts also showed a significant response within the superior temporal gyrus (Fig. 6a; Supplementary Fig. 6). The stg has been found to contain neurons responsive to auditory and visual stimuli[36].

We quantified the magnitude of the responses to the different stimuli within regions of interest along the posterior-anterior axis (Fig. 6c,d,e). Color regions at all A-P locations showed a significant response to all images; the responses to each image class were not significantly different from each other within 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/fruits (Fig. 6c). The responses within the face patches were greatest to faces, with weaker but still strong responses to bodies/body parts and vegetables/fruits; places elicited a substantially weaker response in the face patches (Fig. 6d). We also quantified the responses within regions of IT outside of the face patches and color-biased regions; these “other IT” regions were defined as the most visually active voxels within 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/fruits and faces, and then by places (Fig. 6e). They were more strongly modulated than the color-biased regions by all image classes; and they showed a substantial response to faces. As expected, the color-biased regions showed significantly higher 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 show that to a significant 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; 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/fruit (Fig. 6a; Supplementary Fig. 6), although both contrasts yielded 3 clusters of body-biased regions along the temporal lobe. This was not true 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 show significant activation throughout most of IT except conspicuous islands that correspond to the face patches (insets, Supplementary Fig. 6). These results are consistent with present consensus that the most selective cortex in IT is selective for faces, and selectivity within 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 within IT[23]. 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, Supplementary Fig. 7). Close inspection suggests three distinct representations of the central visual field, which coincide with the anterior-posterior 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 found in V1 (Fig. 7b), suggesting that IT does not possess a stronger central bias than found in V1. The face patches, however, showed a strong central bias (Fig. 7b,c; t-test, $p=4\times10^{-7}$, N=4), which was evident in each face patch assessed independently (Supplementary Fig. 8a). The color-biased regions did not show a bias for center or periphery (Fig. 7b,c; Supplementary Fig. 8b). The place-biased regions showed a peripheral bias (Fig. 7b,c; t-test, p=0.01, N=4). For comparison, and to evaluate how reliable the eccentricity mapping signals were, from results obtained during one half of the stimulus presentations we computed two regions of interest, one defined as more responsive to the central stimulation and the other defined as more responsive to peripheral stimulation (Fig. 7c). We quantified the response bias within these regions using results obtained during the other half of the stimulus presentations. As predicted, both regions show significant response bias (t-tests, central representation, $p=2\times10^{-4}$; peripheral representation, $p=2\times10^{-6}$; N=4). 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) within regions-of-interest defined by the eccentricity mapping. These results show a distinct bias for face stimuli within the central visual field representation; a clear increase in modulation to place stimuli in the peripheral representation; and an intermediate response to non-face/non-place images in the mid-peripheral representation, confirming results obtained in humans[23]. Responses to color consistently showed a subtle but significant luminance bias (Fig. 7e, all values are below 0). 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/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/fruit, body parts, 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. But remarkably, color-biased regions appear to be yoked to face patches, with color-biased regions typically 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 studies[18] superimposed on a lateral view of the brain showing the functional activation. The correspondence is striking, suggesting that the anatomical connections underlie the observed functional organization.

Although anatomical evidence suggests the stages are not linked in a strictly feed-forward fashion[18] (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 patches[5]. The systematic relationship of the face patches and color-biased regions, their regular spacing, and the progressive increase in selectivity within 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”. The results suggest that these areas each comprise a complete functional representation of objects, and are possibly homologous to the discrete areas found within early visual cortex. 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 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 within each stage and contextual interactions[9], 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 posterior lesions in humans[40].

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 neuroscience[41]. Although we found evidence for parallel routes for color and form in IT, the color-biased regions were significantly 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 (e.g. yellow=banana). The relative contribution to shape

processing of the neurons in color-biased regions compared to neurons in intervening tissue awaits further study, as does the specific contribution to color made at each stage within the color-biased regions 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 hippocampus[42], 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 cortex[18] provide tantalizing hints to the neural basis for the behavioral significance 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 cortex[23]. It has not been clear whether this principle applies across primates. The experiments presented here provide fMRI confirmation of electrophysiological evidence[13,24,25] for a retinotopic map in posterior IT; moreover, they revealed additional representations of visual-field representations within CIT and AIT (Fig. 7). These representations corresponded to locations of the face patches, supporting conclusions drawn from human imaging experiments that central representations within object-processing cortex have evolved a specialization for computations that depend on high spatial acuity[23]. The present work also confirms the result obtained in humans that peripheral representations overlap with regions showing stronger responses to places[23]. Color-biased regions were located in an intermediate zone in the eccentricity maps. Together with work in humans, the results suggest that eccentricity maps are the unifying rule that governs the organization of IT across primates.

The results of the eccentricity mapping experiment are surprising because IT 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 within 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 field[43]. These results suggest that IT does not possess an explicit retinotopic representation, but rather one that mirrors the scale of objects recorded in cognition. The eccentricity results presented here 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).

The pattern of results described presently 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 cortex[23]. 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 within 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 neurophysiology[46] underlying this idea led to the notion that acquired cerebral achromatopsia comes about because of a lesion of the human homologue of monkey V4 (for review see[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 color[47]; second, lesions of V4 in monkeys do not cause selective (or substantial) long-term impairments in color (for review see[48]); and third, the cerebral achromatopsia in humans is often accompanied by other visual deficits, specifically prosopagnosia[49]. That achromatopsia can manifest in the context of largely preserved visual object perception suggests some degree of functional specialization within the cortex. But it has been difficult to reconcile such functional specialization with the striking 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 described presently could account the clinical observations if humans share a similar IT organization to monkeys, since 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.

Experimental Procedures

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 (MGH) in a 3-T Tim Trio scanner (Siemens, New York, NY). MR images were acquired with a custom-built four-channel MR coil system with AC88 gradient insert, which increases the signal-to-noise ratio (SNR) by allowing very short echo times (TE) [50], providing 1 mm^3 spatial resolution and good coverage of the temporal lobe (Supplementary Figure 13). We used standard echo planar imaging (repetition time (TR) = 2 s, 96 × 96 × 50 matrix, 1 mm^3 voxels; TE-13 ms). In all the experiments described here, data were obtained after administration of an intravenous contrast agent, which further enhances SNR [13] (Supplementary Figure 1). 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, animals 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, Burlington, MA) was used to monitor eye movements, and animals were only rewarded for maintaining their gaze within ~1 degree of the central fixation target. MR signal contrast was enhanced using a microparticular iron oxide agent (MION, Feraheme; 8–10 mg/kg, diluted in saline, AMAG Pharmaceuticals), injected intravenously into the femoral vein just prior to 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 the local and 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 animals were used in color-detection experiments, which show macaque monkeys have similar psychophysical chromatic mechanisms to humans[10].

Surgery

In order to maintain head stabilization during testing, plastic (Delrin) headposts were implanted using standard sterile surgical procedures. Animals were anesthetized with ketamine (15 mg/kg i.m.) and xylazine (2 mg/kg i.m.) and given atropine (0.05 mg/kg i.m.) to reduce salivary fluid production. Depth of anesthesia was maintained with 1–2% isoflurane in oxygen administered with a tracheal tube. Animals were given a pre-emptive dose of buprenorphine (0.005 mg/kg i.m.) and flunixin (1.0 mg/kg i.m) as analgesics and a prophylactic dose of an antibiotic (Baytril, 5 mg/kg i.m.). Antibiotic was administered again 1.5 hours into surgery; buprenorphine and flunixin were given for 48 hour 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 (1024 × 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 paradigm. Responses to color, objects, eccentricity and meridians were run in separate sessions. Color stimuli were generated in DKL color space[26], and calibrated using spectral readings taken with a PR-655 spectroradiometer (Photo Research Inc., Chatsworth, CA) using procedures developed by Thorsten 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 primaries [51], and cone excitation was calculated using the Smith and Pokorny cone fundamentals [52]. Twelve color directions were chosen, evenly sampling the azimuth of the equiluminant plane of DKL space (Figure 2a, Supplementary Figure 3). Chromaticity coordinates and cone contrasts differed slightly between experiments due to projector bulb performance changing over time. Stimuli were re-calibrated one or two days before each experiment.

Color stimuli were presented as equiluminant single-color vertically oriented trapezoid-wave gratings (Figure 2a). The gratings (2.9 cycles/degree; drifting 0.75 cycle/s) were drifted back and forth, switching directions every 2s, in 32 second blocks interleaved with equiluminant neutral full-field gray (32 s), in two stimulus orders: run 1: colors 1, 3, 5, bw, 7, 9, 11; and run 2: 2, 4, 6, bw, 8, 10, 12, bw, where bw=black&white achromatic luminance gratings of 10% contrast. Responses were also measured to an achromatic black and white grating of 10% luminance contrast.

Color selectivity has historically been a challenge to define [14]. 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 luminance [53]; 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 (Figure 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 (Figure 2a), colors that fall on or close to the L-M axis (i.e. 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 Figure 2b); moreover, the pattern of color-biased activation was not critically determined on the basis of significance thresholds (Supplementary Figure 14).

Object stimuli

The stimuli were achromatic photographs of faces (humans and monkeys), body parts, objects (fruit/vegetables), and familiar places, along with scrambled versions of these images, similar to those used previously [2]. Example images are given in Figure 6. The images were shown in sixteen 32-second blocks (16TR/block, TR=2sec, 2 images/TR) 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 (~25cd/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*1sec/image*16images/block=32 seconds/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 ¾ 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 within 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 within all regions of interest 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 4 presentations of horizontal wedges and 4 presentations of vertical wedges. The functional border assignments obtained with this method correspond well to areal boundaries of a standard macaque atlas[56]. Eccentricity mapping was performed using discs and annuli of a colored checkerboard. Four 32s blocks, interleaved with 32s 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° (see Figure 6b). One set of stimuli comprised colors 1, 7 (activating L and M cones); 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 (Figure 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, while the locations of the color regions were randomly assigned within 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 (LH/RH, 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 3-D simulation could be closer than if the simulation were restricted to a 2-D 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 9 scan sessions in the two animals. For the color localizer: 13600 volumes in M1; 7072 in M2. For the eccentricity mapping: 1904 in M1; 4080 for M2. For the object mapping: 4608 in M1; 4608 in M2. For the meridian mapping: 3328 in M1; 1280 in M2. High-resolution anatomical scans ($0.35\times0.35\times0.35$ 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" tookit provided by Joseph 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 algorithm[57], spatially smoothed with a Gaussian kernel (fwhm = 2 mm; except for Supplementary Figure 5 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 $x(t)$ was the raw fMRI signal, and $s(t)$ was 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:

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

where $s(t)$, $t = 1, 2, \ldots N$, $\bar{s}$ is the mean of $s(t)$, and N is the number of TRs in the experiment.

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 in the section on normalizing activation maps (see below), significance thresholds to define ROIs in Figure 3c, Figure 5a, Figure 6 (b,c,d,f,g), Figure 7c, Supplementary Figure 4, Supplementary Figure 6, Supplementary Figure 11, and Supplementary Figure 12, were set stringently to yield roughly the same number of color-selective and face-selective voxels in each hemisphere (LH/RH, 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. 14). 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 13a 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 Fig. 13b 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 MR signal. The "other IT" regions described in Fig. 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 within 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

(e.g. Figure 6b). The response during a given stimulus block was calculated using the average of the last 7 TRs 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:

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

The % signal change to color was the average Rc value during presentation of colors 7 & 8; the % 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 Rc value during presentation of the largest two annuli (see Figure 7b, top), averaging across the L-M and S conditions.

The % 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, so:

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

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

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

$$\text{Selectivity} = (Rp - Rnp)/(Rp + Rnp)$$

For color, *Rp* = response to color gratings (colors 7&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, prior to 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 7 TRs within a stimulus block) were calculated against the response to the intervening gray block. In Figure 3c, responses to all patches at a given A-P location have been averaged together (e.g. “C” shows responses for MF and ML), 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/hemisphere x 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 obtain 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 (Figure 5). Importantly, 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, as shown in Supplementary Figure 14, the maps generated by using a range of p-value cut-offs does not radically alter the pattern of activation in either the face condition or the color condition.

Inter-animal variability and methodological considerations

The color-biased activity in central and anterior IT was more distinct than we found previously[13], especially in the animal whose data is shown in Figure 1 of Conway et al. (2007). This animal showed no face patches in IT. Other technical differences exist between the present study and the earlier report: the present report takes advantage of technical developments which 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 reported [13]. 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. Supplementary Figure 5 shows that using all colors besides colors 7 & 8 to generate the color-response pattern (rather than just colors 7&8 as in Figure 2), and no spatial smoothing, we find a comparable pattern in V2 and V4 to that obtained previously. The pattern of activation within 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 system [31]. A larger study will be necessary in order to determine the extent of this variability. Towards this goal, we have recently obtained data in two additional animals using the AC88 insert and four-channel coil system; the data were obtained towards a different experimental objective and we used heterochromatic gratings to determine the color-biased regions. Nonetheless, the results in IT support the conclusions reported presently ().

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/

PITd [13]; these data were obtained using a single-channel coil without AC88 insert, in four animals in which face patches had also been tested [31]. 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 MT [12]; other methodological details can be found in [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 contrast [53], 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 influence 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 since fMRI signals to color and luminance likely sum in a non-linear way. Second, it is not known to what extent the enhancement in SNR afforded by the four-channel coil and the AC88 is decisive. We intend to present a complete analysis of the data from Conway et al (2007) and Tsao et al (2008) to address these issues. Here we provide a summary of the analysis of the color-biased activation and face patches within IT in the data set used in [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 (Figure 1 of [13]). A second animal showed a close association of PLc, CLc, and ALc to PL, ML and AL in both hemispheres (Monchichi, Supplementary Figure 15; Figure 2 in [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 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.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Supported by grants from the National Science Foundation (0918064), the NIH (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) a P41 Biotechnology Resource Grant supported by the National Institute of Biomedical Imaging and Bioengineering (NIBIB), National Institutes of Health. This work also involved the use of instrumentation supported by the NIH Shared Instrumentation Grant Program and/or High-End Instrumentation Grant Program (S10RR021110). Functional MRI data were analyzed using the jip toolbox generously provided by Joseph Mandeville (http://www.nitrc.org/projects/jip/), and analysis scripts compiled by Sebastian Moeller. We thank J. Maunsell, D. Hubel and M. Livingstone for research space, helpful discussions and assistance with animal protocols. D. Tsao, A. Rehding, N. Kanwisher, J. DiCarlo, and P. Mayo provided valuable comments on the manuscript. BRC gratefully acknowledges seminal conversations that took place with Doris 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 on MRI analysis. We thank J. McCready and E. Erzinger, 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 4-channel MR 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).

References

1. DiCarlo JJ, Zoccolan D, Rust NC. How does the brain solve visual object recognition? Neuron. 2012; 73:415–434. [PubMed: 22325196]
2. Tsao DY, FW, Knutsen TA, Mandeville JB, Tootell RBH. Faces and objects in macaque cerebral cortex. Nat Neurosci. 2003; 6:989–995. [PubMed: 12925854]
3. Pinsk MA, DeSimone K, Moore T, Gross CG, Kastner S. Representations of faces and body parts in macaque temporal cortex: a functional MRI study. Proc Natl Acad Sci U S A. 2005; 102:6996–7001. [PubMed: 15860578]
4. Ku SP, Tolias AS, Logothetis NK, Goense J. fMRI of the face-processing network in the ventral temporal lobe of awake and anesthetized macaques. Neuron. 2011; 70:352–362. [PubMed: 21521619]
5. Freiwald WA, Tsao DY. Functional compartmentalization and viewpoint generalization within the macaque face-processing system. Science. 2010; 330:845–851. [PubMed: 21051642]
6. Srihasam K, Mandeville JB, Morocz IA, Sullivan KJ, Livingstone MS. Behavioral and anatomical consequences of early versus late symbol training in macaques. Neuron. 2012; 73:608–619. [PubMed: 22325210]
7. Op de Beeck HP, Deutsch JA, Vanduffel W, Kanwisher NG, DiCarlo JJ. A stable topography of selectivity for unfamiliar shape classes in monkey inferior temporal cortex. Cereb Cortex. 2008; 18:1676–1694. [PubMed: 18033769]
8. Bell AH, Hadj-Bouziane F, Frihauf JB, Tootell RB, Ungerleider LG. Object representations in the temporal cortex of monkeys and humans as revealed by functional magnetic resonance imaging. J Neurophysiol. 2009; 101:688–700. [PubMed: 19052111]
9. Cox D, Meyers E, Sinha P. Contextually evoked object-specific responses in human visual cortex. Science. 2004; 304:115–117. [PubMed: 15001712]
10. Stoughton CM, Lafer-Sousa R, Gagin G, Conway BR. Psychophysical chromatic mechanisms in macaque monkey. J Neurosci. 2012; 32:15216–15226. [PubMed: 23100442]
11. Tootell RB, Nelissen K, Vanduffel W, Orban GA. Search for color 'center(s)' in macaque visual cortex. Cerebral Cortex. 2004; 14:353–363. [PubMed: 15028640]
12. Conway BR, Tsao DY. Color architecture in alert macaque cortex revealed by FMRI. Cereb Cortex. 2006; 16:1604–1613. [PubMed: 16400160]
13. Conway BR, Moeller S, Tsao DY. Specialized color modules in macaque extrastriate cortex. Neuron. 2007; 56:560–573. [PubMed: 17988638]
14. Harada T, et al. Distribution of colour-selective activity in the monkey inferior temporal cortex revealed by functional magnetic resonance imaging. Eur J Neurosci. 2009; 30:1960–1970. [PubMed: 19912328]
15. Katsuyama N, et al. Cortical activation during color discrimination task in macaques as revealed by positron emission tomography. Neurosci Lett. 2010; 484:168–173. [PubMed: 20727941]
16. 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. 1987; 261:362–387. [PubMed: 3611417]
17. Van Essen DC, Felleman DJ, DeYoe EA, Olavarria J, Knierim J. Modular and hierarchical organization of extrastriate visual cortex in the macaque monkey. Cold Spring Harb Symp Quant Biol. 1990; 55:679–696. [PubMed: 1966771]
18. Kravitz DJ, Saleem KS, Baker CI, Ungerleider LG, Mishkin M. The ventral visual pathway: an expanded neural framework for the processing of object quality. Trends Cogn Sci. 2012; 17:26–49. [PubMed: 23265839]
19. Grill-Spector K, Malach R. The human visual cortex. Annu Rev Neurosci. 2004; 27:649–677. [PubMed: 15217346]

20. Kanwisher N. Functional specificity in the human brain: a window into the functional architecture of the mind. Proc Natl Acad Sci U S A. 2010; 107:11163–11170. [PubMed: 20484679]
21. Hadjikhani N, Liu AK, Dale AM, Cavanagh P, Tootell RB. Retinotopy and color sensitivity in human visual cortical area V8 [see comments]. Nature Neuroscience. 1998; 1:235–241.
22. Bartels A, Zeki S. The architecture of the colour centre in the human visual brain: new results and a review. Eur J Neurosci. 2000; 12:172–193. [PubMed: 10651872]
23. Hasson U, Harel M, Levy I, Malach R. Large-scale mirror-symmetry organization of human occipito-temporal object areas. Neuron. 2003; 37:1027–1041. [PubMed: 12670430]
24. Boussaoud D, Desimone R, Ungerleider LG. Visual topography of area TEO in the macaque. Journal of Comparative Neurology. 1991; 306:554–575. [PubMed: 1712794]
25. Yasuda M, Banno T, Komatsu H. Color selectivity of neurons in the posterior inferior temporal cortex of the macaque monkey. Cereb Cortex. 2010; 20:1630–1646. [PubMed: 19880593]
26. Derrington AM, Krauskopf J, Lennie P. Chromatic mechanisms in lateral geniculate nucleus of macaque. Journal of Physiology. 1984; 357:241–265. [PubMed: 6512691]
27. Koida K, Komatsu H. Effects of task demands on the responses of color-selective neurons in the inferior temporal cortex. Nat Neurosci. 2007; 10:108–116. [PubMed: 17173044]
28. Matsumora T, Koida K, Komatsu H. Relationship between color discrimination and neural responses in the inferior temporal cortex of the monkey. J Neurophysiol. 2008; 100:3361–3374. [PubMed: 18922950]
29. Stoughton CM, Conway BR. Neural basis for unique hues. Current Biology. 2008; 18:R698–R699. [PubMed: 18727902]
30. Conway BR, Tsao DY. Color-tuned neurons are spatially clustered according to color preference within alert macaque posterior inferior temporal cortex. Proceedings of the National Academy of Science (USA). 2009; 106:18034–18039.
31. Tsao DY, Moeller S, Freiwald WA. Comparing face patch systems in macaques and humans. Proc Natl Acad Sci U S A. 2008; 105:19514–19519. [PubMed: 19033466]
32. Cavanagh P, Leclerc YG. Shape from shadows. J Exp Psychol Hum Percept Perform. 1989; 15:3–27. [PubMed: 2522531]
33. Rolls ET. Face processing in different brain areas, and critical band masking. J Neuropsychol. 2008; 2:325–360. [PubMed: 19824174]
34. Van Essen DC. A tension-based theory of morphogenesis and compact wiring in the central nervous system. Nature. 1997; 385:313–318. [PubMed: 9002514]
35. Conover, WJ. Practical nonparametric statistics. 2. New York: John Wiley & Sons; 1980.
36. Karnath HO. New insights into the functions of the superior temporal cortex. Nat Rev Neurosci. 2001; 2:568–576. [PubMed: 11484000]
37. Baker CI, et al. Visual word processing and experiential origins of functional selectivity in human extrastriate cortex. Proc Natl Acad Sci U S A. 2007; 104:9087–9092. [PubMed: 17502592]
38. Tanaka K, Saito H, Fukada Y, Moriya M. Coding visual images of objects in the inferotemporal cortex of the macaque monkey. J Neurophysiol. 1991; 66:170–189. [PubMed: 1919665]
39. 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. 1996; 25:37–52. [PubMed: 8861169]
40. Moro V, et al. The neural basis of body form and body action agnosia. Neuron. 2008; 60:235–246. [PubMed: 18957216]
41. Conway BR, et al. Advances in color science: from retina to behavior. Journal of Neuroscience. 2010; 30:14955–14963. [PubMed: 21068298]
42. Iwai E, Yukie M. A direct projection from hippocampal field CA1 to ventral area TE of inferotemporal cortex in the monkey. Brain Res. 1988; 444:397–401. [PubMed: 2834025]
43. Konkle T, Oliva A. A real-world size organization of object responses in occipitotemporal cortex. Neuron. 2012; 74:1114–1124. [PubMed: 22726840]
44. Allman JM, Kaas JH. A crescent-shaped cortical visual area surrounding the middle temporal area (MT) in the owl monkey (Aotus trivirgatus). Brain Res. 1974; 81:199–213. [PubMed: 4215542]

45. Sereno, Ml; Allman, JM. Cortical visual areas in mammals. In: Leventhal, AG., editor. The Neural Basis of Visual Function. London: Macmillan; 1991. p. 160-172.
46. Zeki S. The representation of colours in the cerebral cortex. Nature. 1980; 284:412–418. [PubMed: 6767195]
47. Roe AW, et al. Towards a unified theory of visual area V4. Neuron. 2012; 74:12–29. [PubMed: 22500626]
48. Heywood CA, Kentridge RW. Achromatopsia, color vision, and cortex. Neurol Clin. 2003; 21:483–500. [PubMed: 12916488]
49. Bouvier SE, Engel SA. Behavioral deficits and cortical damage loci in cerebral achromatopsia. Cereb Cortex. 2006; 16:183–191. [PubMed: 15858161]
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 Mag Reson Med. 2004; 11:1630.
51. Hansen T, Giesel M, Gegenfurtner KR. Chromatic discrimination of natural objects. J Vis. 2008; 8:2, 1–19. [PubMed: 18318605]
52. Smith VC, Pokorny J. Spectral sensitivity of color-blind observers and the cone photopigments. Vision Research. 1972; 12:2059–2071. [PubMed: 4539070]
53. Johnson EN, Hawken MJ, Shapley R. Cone inputs in macaque primary visual cortex. J Neurophysiol. 2004; 91:2501–2514. [PubMed: 14749310]
54. Fize D, et al. The retinotopic organization of primate dorsal V4 and surrounding areas: A functional magnetic resonance imaging study in awake monkeys. Journal of Neuroscience. 2003; 23:7395–7406. [PubMed: 12917375]
55. Lafer-Sousa R, Liu YO, Lafer-Sousa L, Wiest MC, Conway BR. 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. 2012; 29:657–670. [PubMed: 22561924]
56. Paxinos, G.; Huang, X-F.; Toga, AW. The Rhesus Monkey Brain In Stereotaxic Coordinates. Academic Press; San Diego, California: 2000.
57. Cox RW, Hyde JS. Software tools for analysis and visualization of fMRI data. NMR in Biomedicine. 1997; 10:171–178. [PubMed: 9430344]
58. Kriegeskorte N, Simmons WK, Bellgowan PS, Baker CI. Circular analysis in systems neuroscience: the dangers of double dipping. Nat Neurosci. 2009; 12:535–540. [PubMed: 19396166]
59. Simmons WK, Bellgowan PS, Martin A. Measuring selectivity in fMRI data. Nat Neurosci. 2007; 10:4–5. [PubMed: 17189941]
60. Baker CI, Hutchison TL, Kanwisher N. Does the fusiform face area contain subregions highly selective for nonfaces? Nat Neurosci. 2007; 10:3–4. [PubMed: 17189940]

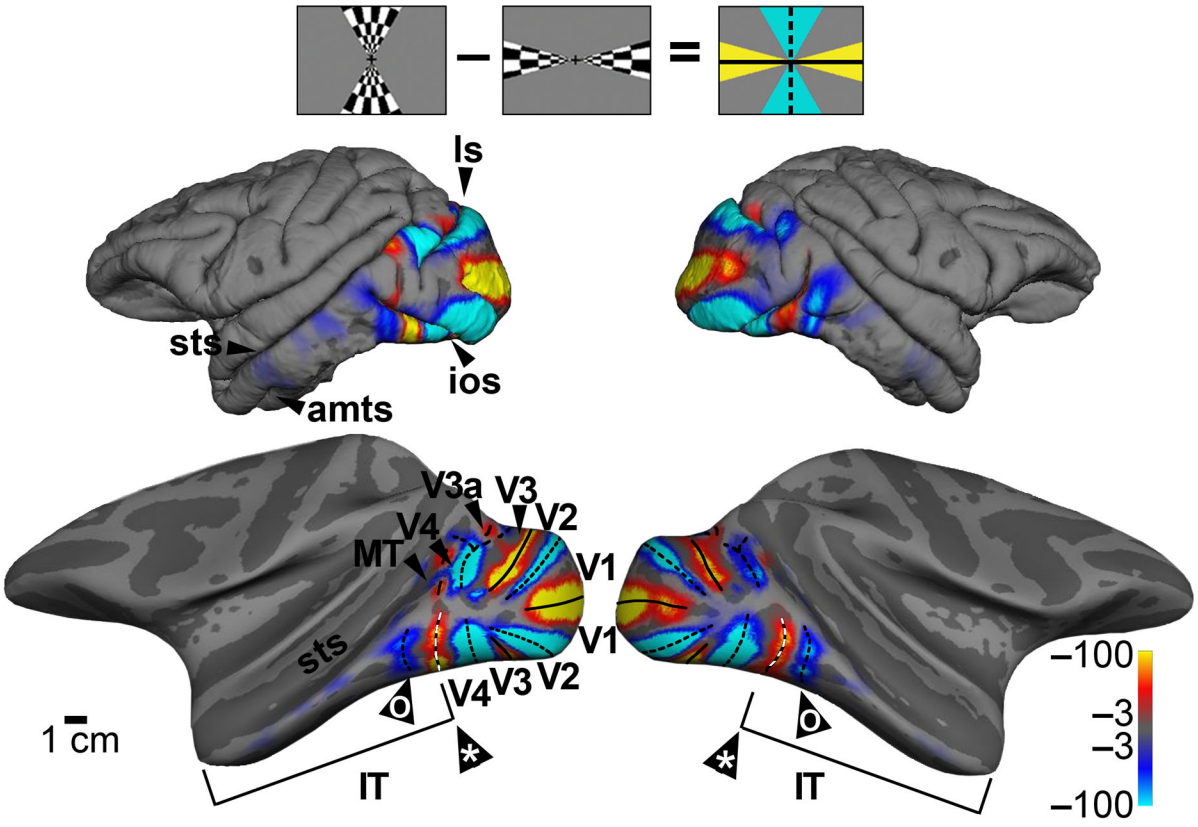

Figure 1.

Identification of the boundaries of inferior temporal cortex and retinotopic visual areas using fMRI and retinotopic mapping. Stronger responses to stimulation along vertical meridians are shown in blue-cyan; stronger responses to stimulation along the horizontal meridian are shown in orange-red. Icon at top shows the visual stimulus. Pial surface view, top; computationally inflated view, bottom (representative activation shown for M1). In addition to a clear V4/IT boundary (*), the maps reveal a vertical meridian representation within IT (o). Color scale bars show significance as the common logarithm of the probability of error. ls, lunate sulcus; sts, superior temporal sulcus; ios, inferior occipital sulcus; amts, anterior middle temporal sulcus.

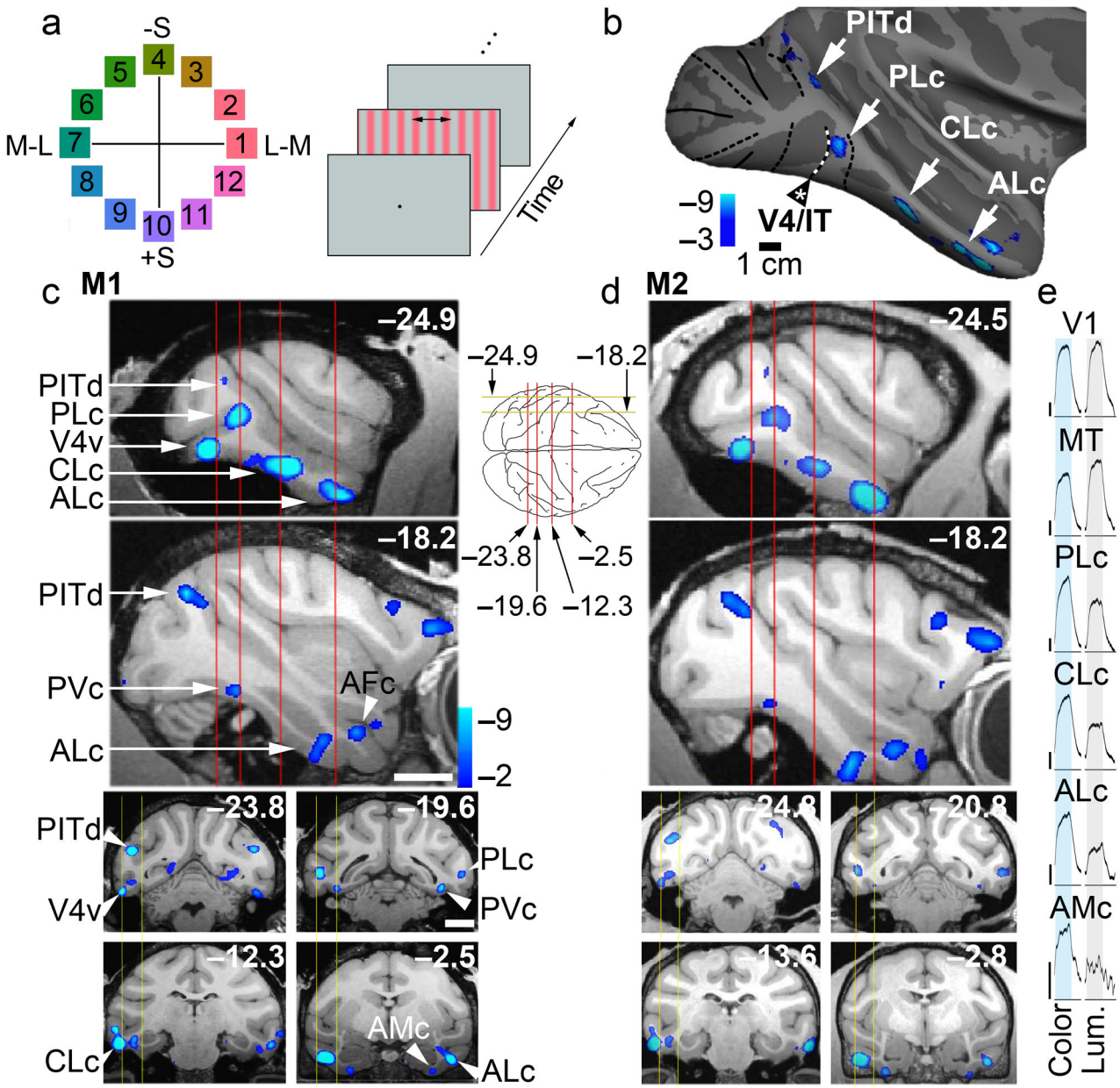

Figure 2.

Functional architecture of color-biased regions in alert macaque IT. **a,** Color stimuli shown in the equiluminant plane of the DKL color space[26]; colors 1 and 7 modulate only the L and M cones; colors 10 and 4 modulate only the S cones; other colors modulate all cone classes (Supplementary Fig. 2). Right panel shows the stimulus paradigm: color-gray gratings presented in 32 s blocks, maintaining mean luminance of 55 cd/m^2, interleaved with neutral gray, in two stimulus orders. **b,** Regions (blue-cyan) of M1, right hemisphere, showed greater activation to colors 7 & 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, 1cm;

slices given in Talairach coordinates (mm). **d,** Data as in panel c, for a second monkey. Functional data has been superimposed on high-resolution anatomical scans. **e,** Time-course traces averaging activity during all color-grating blocks and achromatic-luminance-grating blocks within 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; 17 blocks/run (272 measurements/run); 16 TR/block; TR=2sec. See Supplementary Fig. 3a for the time course of the response to individual colors within each region of IT.

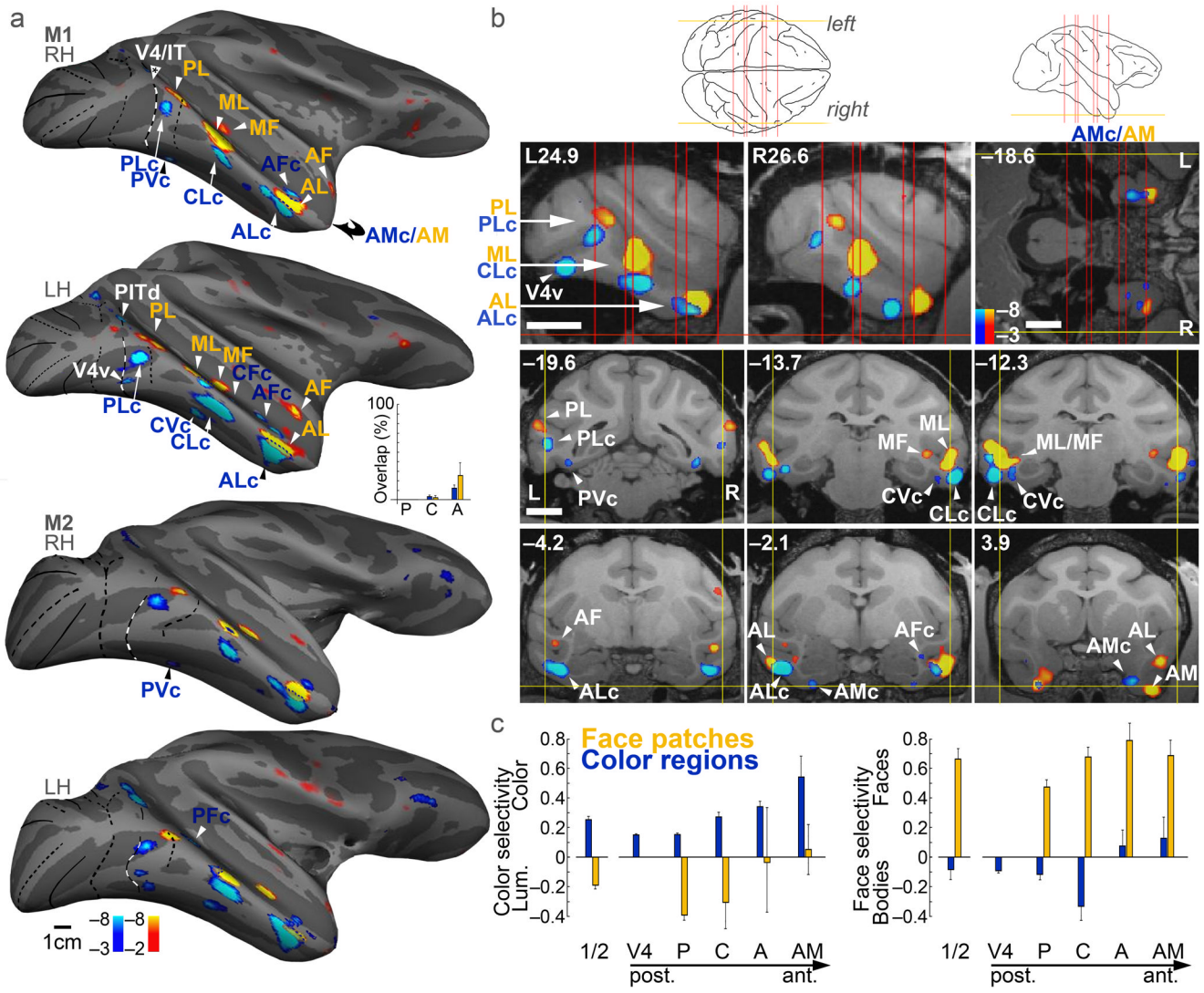

Figure 3.

Color-biased fMRI activation (blue-cyan) found at corresponding locations to face patches (orange-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 as in Figure 1. See text for naming conventions. 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 P→A (orange bars, p=0.04; blue bars, p=0.0008, multiple linear regression, N=4). **b,** Slices through M1 (Talairach coordinates), showing color-biased and face-biased activation. Top left two panels, sagittal slices (L, left; R, right; section plane: yellow lines, left schematic). Top right panel, 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 bar, 1cm. **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 ($p= 9\times10^{-4}$), and face regions were significantly more face selective than color regions ($p= 4\times10^{-10}$) (unpaired two-tailed t-tests, N=16; 4 color regions/hemi, 4 face regions/hemi, 4 hemis). 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. The “1/2” bars show selectivity averaged across ROIs, within ROIs defined by ½ the data set.

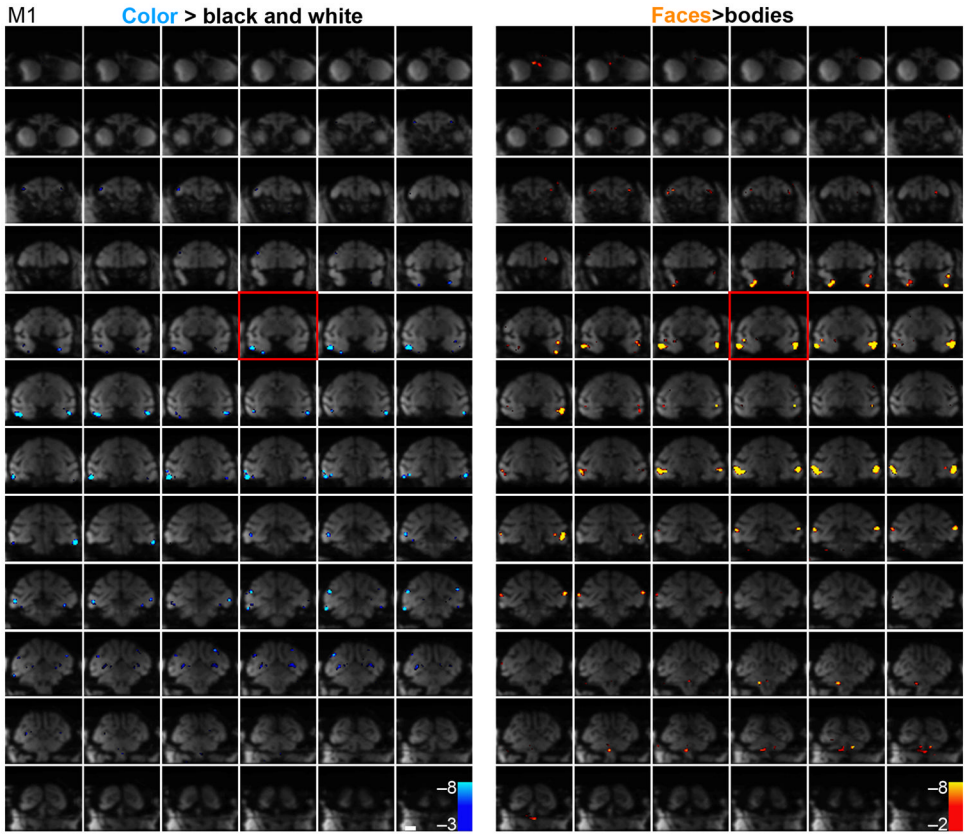

Figure 4.
Raw functional echo planar (EPI) coronal images showing MION activation every 1 mm; color-biased activation (blue, left panels) and face-biased activation (blue, right panels) of M1. See Supplementary Figure 5 for these data obtained in M2. Spurious activation outside the brain has been masked. The slice corresponding to Talairach coordinate 0 along the P-A axis is boxed in red. Scale bar, 1 cm.

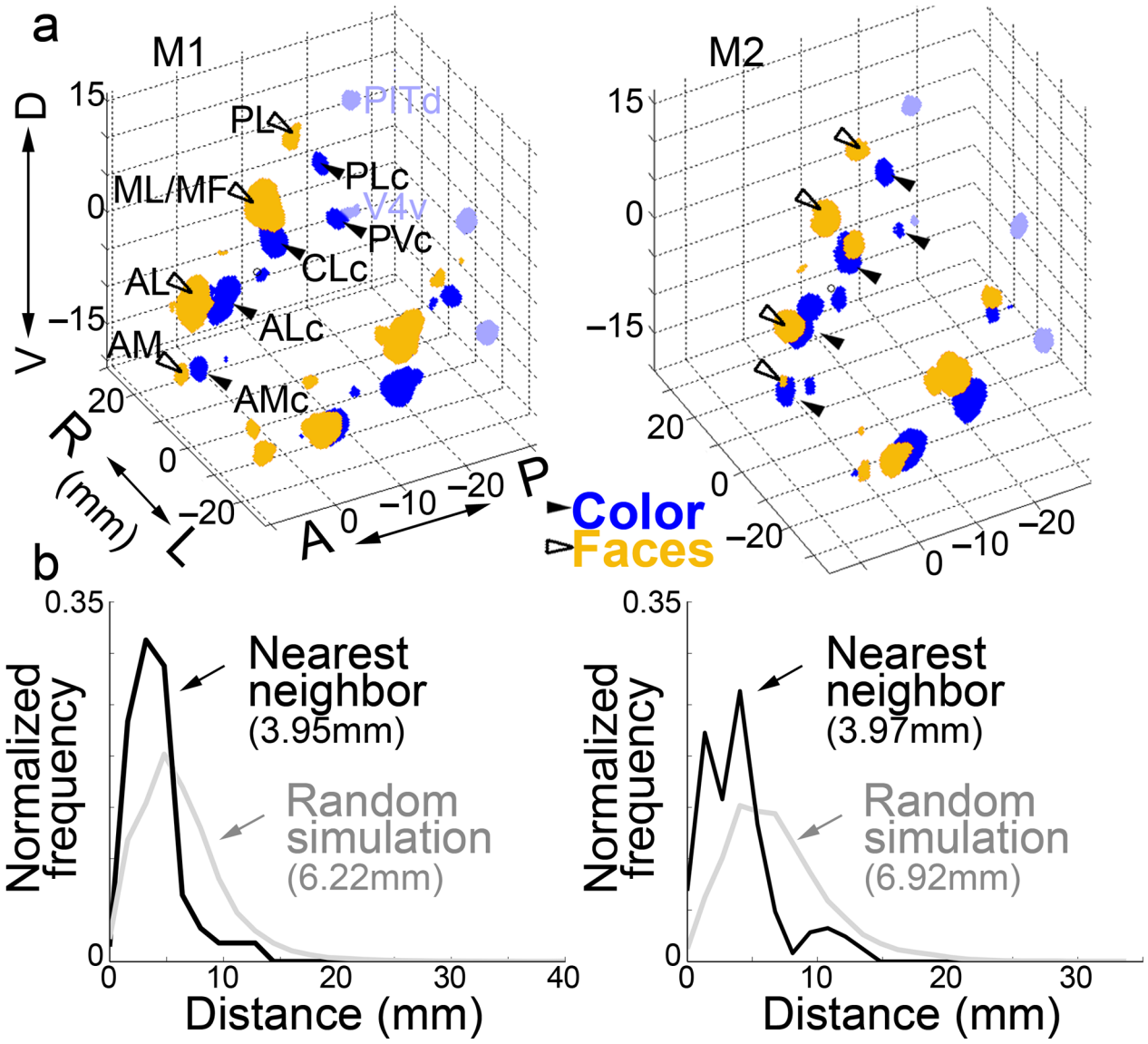

Figure 5.

Quantification of the spatial relationship of color-biased regions and face patches in alert macaque IT. **a,** 3-D plot showing color-biased regions (blue, solid arrowheads) and face patches (orange, open arrowheads) in both hemispheres for both animals (M1, M2); see text for labeling convention; color-biased regions in the V4 Complex 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 (~4mm) was significantly closer than expected by chance (6–7mm; Mann–Whitney–Wilcoxon test M1, $p=1\times10^{-17}$, N = 278 voxel to voxel distances; M2, $p=5\times10^{-10}$, N = 247). The observed distributions also show 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

LH, $p=1\times10^{-13}$; M1 RH, $p=3\times10^{-15}$; M2 LH, $p=5\times10^{-27}$; M2 RH, $p=9\times10^{-12}$). The mean of the average values obtained in the four hemispheres was 4.0 mm (s.d. 0.2), and was significantly different from random (6.6 mm, s.d. 0.5, t-test $p=5\times10^{-5}$).

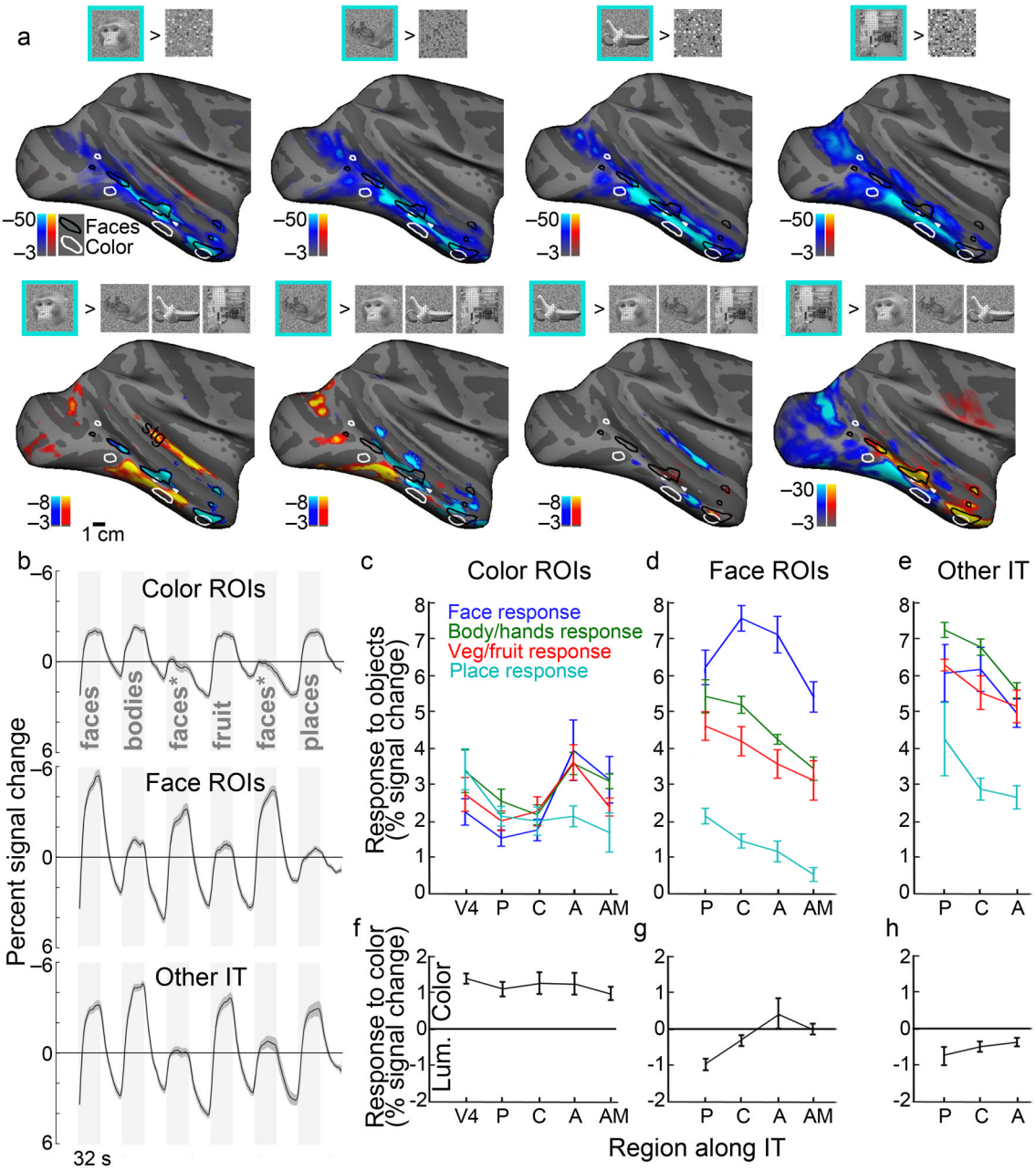

Figure 6.

Responses to images of faces, body-parts, fruit/vegetables and places across IT. **a,** Activation maps showing regions more responsive to intact pictures (blue-cyan) than to scrambled versions (orange-red)(top), and maps showing regions more responsive to one class of pictures over all other intact pictures(bottom), from left to right: faces, body parts, fruit/vegetables, places. Supplementary Figure 6 shows results for individual comparisons. stg, superior temporal gyrus. **b,** Time course of response within color-biased regions (top), face patches (middle), and other visually responsive parts of IT (bottom). The face stimuli indicated by an "*" comprised familiar faces and were not included in the quantification. Time courses are the average of four hemispheres; shading indicates standard error. **c,** Quantification of the responses different image classes within the color-biased regions along

IT, posterior (P), central (C), anterior (A) and anterior-ventral-medial (AM). **d,** Quantification of responses within 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 % signal during color (colors 7&8) minus the % signal during the achromatic grating. Negative values indicate a luminance bias, while positive values indicate a color bias. Error bars s.e. (N=4 hemispheres). All quantification was performed on data sets obtained independently of the data used to generate the regions of interest.

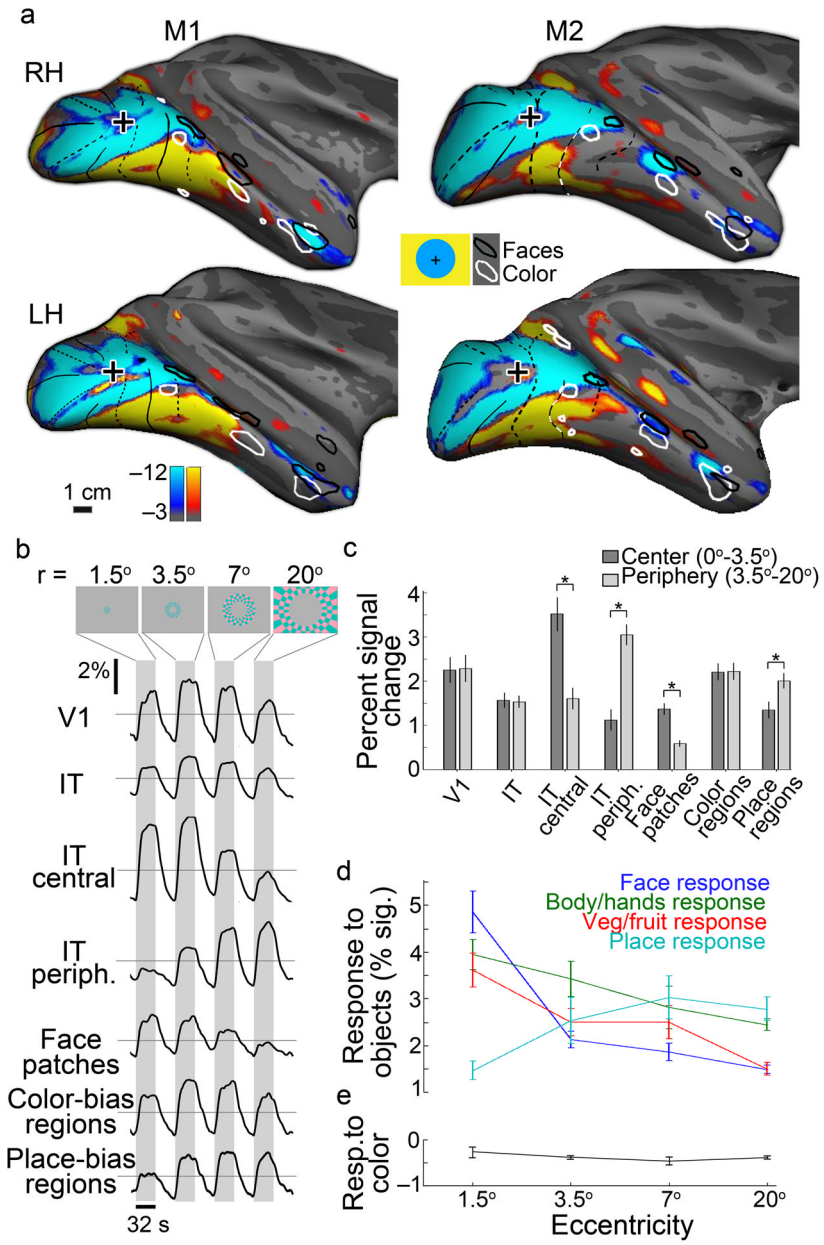

Figure 7.

Inferior temporal cortex 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 more responsive to a central flickering checkered disc (radius 3.5°) shown in blue-cyan; regions more responsive to stimulation with an annulus of flickering checkers (radius 3.5°–20°) shown in orange-red. **b,** Representative time courses for the eccentricity stimulus within V1, IT, the central representation within IT, the peripheral representation within IT, the face patches, the color-biased regions and the place-biased regions. **c,** Bar plots quantifying the responses to central and peripheral stimulation within the various regions of interest. Error bars s.e. (N= 4 hemispheres). **d,e,** Responses to objects (d) and colors (e)

within regions of interest defined using the eccentricity mapping. Error bars s.e. (N= 4 hemispheres).

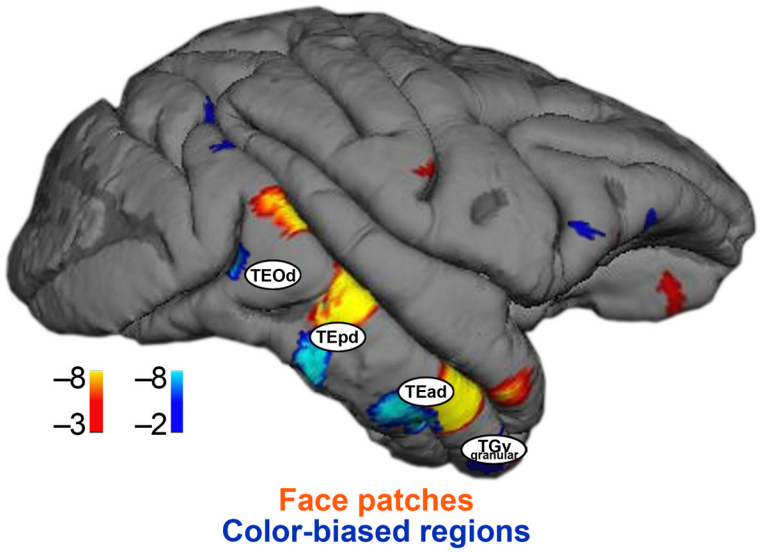

Figure 8.

Projection of the anatomical designations described by Kravitz et al (2012) on the lateral surface of M1 on which has been painted the face-biased (orange-red) and color-biased activation (blue-cyan) patterns.