Evidence for Integrated Visual Face and Body Representations in the Anterior Temporal Lobes

Bronson B. Harry[1,2], Katja Umla-Runge[3], Andrew D. Lawrence[3],
Kim S. Graham[3], and Paul E. Downing[2]

Abstract

■ Research on visual face perception has revealed a region in the ventral anterior temporal lobes, often referred to as the anterior temporal face patch (ATFP), which responds strongly to images of faces. To date, the selectivity of the ATFP has been examined by contrasting responses to faces against a small selection of categories. Here, we assess the selectivity of the ATFP in humans with a broad range of visual control stimuli to provide a stronger test of face selectivity in this region. In Experiment 1, participants viewed images from 20 stimulus categories in an event-related fMRI design. Faces evoked more activity than all other 19 categories in the left ATFP. In the right ATFP, equally strong responses were observed for both faces and headless bodies. To pursue this unexpected finding, in Experiment 2, we used multivoxel pattern analysis to examine whether the strong response to face and body stimuli reflects a common coding of both classes or instead overlapping but distinct representations. On a voxel-by-voxel basis, face and whole-body responses were significantly positively correlated in the right ATFP, but face and body-part responses were not. This finding suggests that there is shared neural coding of faces and whole bodies in the right ATFP that does not extend to individual body parts. In contrast, the same approach revealed distinct face and body representations in the right fusiform gyrus. These results are indicative of an increasing convergence of distinct sources of person-related perceptual information proceeding from the posterior to the anterior temporal cortex. ■

INTRODUCTION

fMRI studies of humans, Old World monkeys (macaques), and New World monkeys (marmosets) have uncovered several face-selective regions in the occipital and temporal lobes (Hung et al., 2015; Tsao & Livingstone, 2008; Kanwisher & Yovel, 2006; Haxby, Hoffman, & Gobbini, 2000). Although cross-species homology has not yet been clearly established, this network of face-selective regions shows a strikingly similar organization across human and nonhuman primates (Hung et al., 2015; McMahon, Russ, Elnaiem, Kurnikova, & Leopold, 2015; Rajimehr, Young, & Tootell, 2009; Tsao, Moeller, & Freiwald, 2008) and consists of several ventral regions spanning the occipital cortex, inferior temporal lobes, and STS.

An influential theoretical perspective (Haxby et al., 2000), based on human functional imaging studies, divides face-selective regions into a "core" system comprising extrastriate nodes for the visual analysis of faces and an "extended" system incorporating additional neural regions that work in concert with the core system to extract various types of social information from faces. The core regions include the occipital face area (OFA; Pitcher, Dilks, Saxe, Triantafyllou, & Kanwisher, 2011; Gauthier et al., 2000), the fusiform face area (FFA; Weiner &

Grill-Spector, 2010; Kanwisher & Yovel, 2006), and the posterior STS (Pitcher et al., 2011; Puce, Allison, Bentin, Gore, & McCarthy, 1998). OFA and FFA are proposed to process static facial form, with the OFA more engaged in part-based processing and the FFA more engaged in processing the configuration of individual parts (Harris & Aguirre, 2010; Schiltz, Dricot, Goebel, & Rossion, 2010; Liu, Harris, & Kanwisher, 2009; Yovel & Kanwisher, 2005), whereas posterior STS processes changeable aspects of faces (e.g., eye gaze; Haxby et al., 2000). In contrast, the extended system is proposed to include regions such as the amygdala and the anterior temporal cortex, areas that are argued to be important in appraising emotional facial expressions (Calder, Lawrence, & Young, 2001; but see Mende-Siedlecki, Verosky, Turk-Browne, & Todorov, 2013) and encoding person-specific semantic knowledge (Quiroga, Kreiman, Koch, & Fried, 2008; Thompson et al., 2004), respectively.

Recent evidence prompts consideration of whether there are anterior temporal regions that should also be included as a part of the core system (Duchaine & Yovel, 2015; Collins & Olson, 2014; Haxby & Gobbini, 2011). A number of reports (Ku, Tolias, Logothetis, & Goense, 2011; Nestor, Plaut, & Behrmann, 2011; Pinsk et al., 2009; Rajimehr et al., 2009; Tsao et al., 2008) provide evidence of at least one face-selective region in the anterior temporal lobes in humans and macaques. In macaques, electrical

---

stimulation of the anterior temporal face patches (ATFPs) selectively induces activity in the posterior network (Moeller, Freiwald, & Tsao, 2008), suggesting that these areas are functionally connected. Moreover, the most ATFP is unique in that neurons in this region respond invariantly to different face views (Meyers, Borzello, Freiwald, & Tsao, 2015; Freiwald & Tsao, 2010), suggesting that it forms higher-level representations that are needed for identification (for similar evidence in human participants, see Yang, Susilo, & Duchaine, 2016; Anzellotti, Fairhall, & Caramazza, 2014). In support of the view that human ATFP captures a similarly abstract representation, Nasr and Tootell (2012) found in human participants that fMRI activity in the ATFP closely mirrored changes in recognition performance brought about by image manipulations such as face inversion and contrast reversal.

A limitation of previous studies examining visual selectivity in the human ATFP has been the use of relatively few visual control categories (e.g., Nasr & Tootell, 2012; Rajimehr et al., 2009; Tsao et al., 2008). A wide assessment of responses to items from a range of categories is vital for determining the selectivity of a region's response profile (Desimone, Albright, Gross, & Bruce, 1984) and thus for making inferences about its functional role(s). There have been multiple-category surveys of the inferior temporal cortex (Mur et al., 2012; Vul, Lashkari, Hsieh, Golland, & Kanwisher, 2012; Downing, Chan, Peelen, Dodds, & Kanwisher, 2006), but these studies only examined responses in posterior temporal regions.

This study attempts to resolve the aforementioned limitations by measuring the profile of responses in the functionally defined human ATFP to a wide range of visually presented stimulus categories. A further aim is to compare this profile with that of more posterior face-selective regions (OFA, FFA) in an effort to reveal how categorical information (particularly about people) emerges over the span of the temporal lobes.

To begin addressing these aims, in Experiment 1, we used a blocked-design functional localizer to first identify the ATFP, as well as OFA and FFA, in individual participants. We used a simultaneous odd-one-out visual discrimination ("oddity") task as a localizer. This task was selected on the grounds that it has been found to be effective in previous fMRI research at selectively activating anterior temporal regions (Barense, Henson, Lee, & Graham, 2010; O'Neil, Cate, & Köhler, 2009; Lee, Scahill, & Graham, 2008) and that performance on this paradigm is sensitive to selective lesions of anterior temporal regions (i.e., perirhinal cortex [PrC]) in monkeys (Buckley, Booth, Rolls, & Gaffan, 2001) and humans (Barense, Gaffan, & Graham, 2007; Lee et al., 2005). In the main experiment, the same participants were presented with images of items of 20 different kinds in an event-related design, while they performed a 1-back task to maintain attention to the stimuli. In this way, we were able to assess in detail the selectivity profile of ATFP and compare it with more posterior face-selective regions.

EXPERIMENT 1

Methods

Participants

Twenty healthy postgraduate volunteers (mean age = 25 years, range = 22–30 years; 13 women) were recruited from Bangor University. All participants were screened for MRI exclusion criteria and gave written informed consent for participation in the experiment, which was approved by the research ethics committee of the School of Psychology at Bangor University, United Kingdom.

Materials

Stimuli for the localizer runs (oddity task) consisted of 96 grayscale images of faces, natural scenes, and common handheld objects (Figure 1, top). Images were organized into 32 triplets for each category (Barense et al., 2010). Each triplet was presented in a triangular formation, consisting of a pair of foil images and a target image, on a 1200 × 840 pixel white background. The foil images were two pictures of the same face, scene, or object taken from different viewpoints. The target was another image from the same category as the foil pair and was selected to be highly similar in appearance to the other pictures. Thirty-two triplets consisting of three black squares were also constructed to appear as an active baseline condition. One of the squares (the target) was slightly larger or smaller than the other two shapes.

Stimuli for the main experimental runs (1-back task) consisted of 48 color images from each of 20 different categories (Figure 1, bottom). Categories consisted of birds, (headless) human bodies, cars, chairs, clothes, crystals, faces, fish, flowers, fruit and vegetables, insects, instruments, (nonhuman) mammals, prepared food, reptiles, spiders, tools, weapons, indoor scenes, and outdoor scenes (Downing et al., 2006). These stimuli were selected to capture a range of object category distinctions (i.e., animate vs. inanimate, large vs. small, natural vs. manmade) that modulate responses in the ventral temporal lobes (Konkle & Oliva, 2012; Mahon & Caramazza, 2011). Stimuli were centred on a white 400 × 400 pixel background, except for scenes, which were cropped to completely fill the image dimensions.

Procedure

To localize functional ROIs, participants completed four runs of an oddity task, each comprising 21 blocks of 15 sec. Blocks 1, 6, 11, 16, and 21 were fixation-only rest conditions. Each of the four stimulus blocks (faces, scenes, objects, and shapes) was presented once between each pair of rest blocks. Stimulation blocks consisted of three oddity trials, each 5 sec in duration. Participants indicated the location of the target stimulus (the odd item out) by pressing one of three buttons. Block order for each set of stimulation conditions was randomly determined between
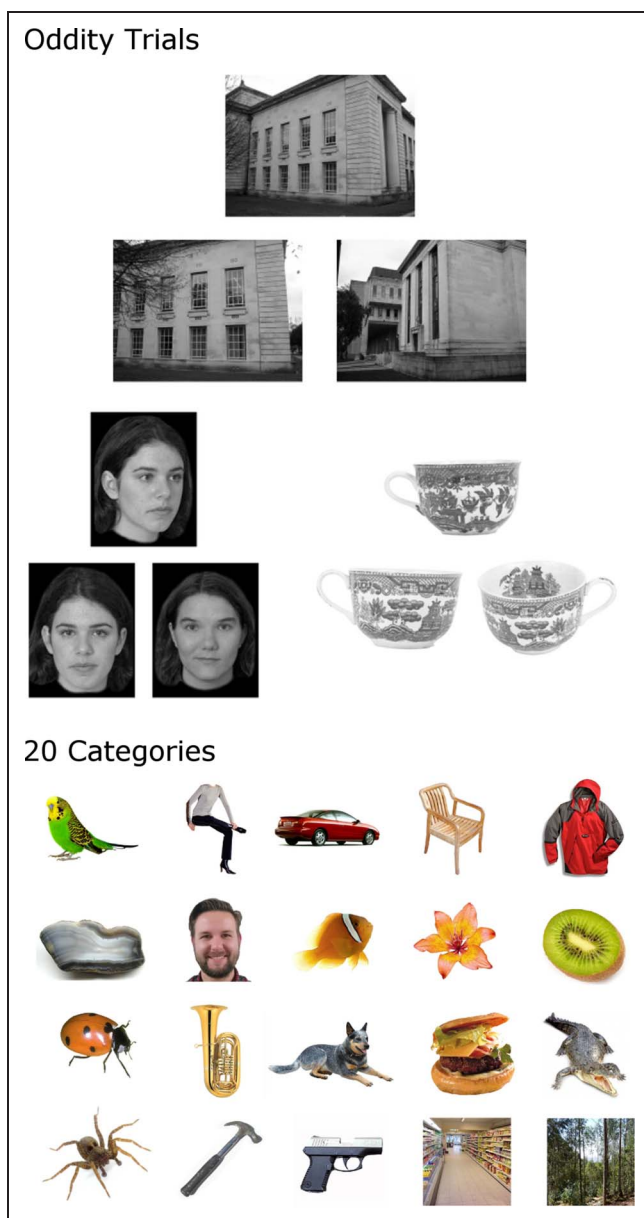
Figure 1. Example of stimuli presented in Experiment 1: localizer task (top) and main experimental task (bottom).

runs and counterbalanced within runs according to a Latin square design.

For the main experimental runs, all images from each of the 20 stimulus categories were presented once in a rapid, event-related design. Participants completed a 1-back task by pressing a button whenever a stimulus was immediately repeated. Stimulus order was determined with a first-order counterbalanced, optimized, *n* = 24, Type 1, Index 1 sequence (Aguirre, 2007). This procedure generated a sequence of 1153 events, including an initial event to establish sequence context. The 20 stimulus categories were assigned to Event types 1–20 for each participant. Event 21 was assigned to target events, whereby the previous item was repeated,

and Events 22–24 were assigned to fixation-only rest conditions. The full sequence was divided into eight separate runs. For Runs 2–8, the final item from the preceding run was presented at the beginning of the next run to reestablish sequence context (making 145 events per run). Participants were assigned to one of five counterbalanced sequences. Stimulus and target events were presented for 300 msec followed by an ISI of 1200 msec that consisted of a central fixation cross. For rest events, a fixation cross appeared for 1500 msec. Fixation-only rest blocks (duration = 16 sec) were presented at the beginning and end of each run.

Localizer runs were interspersed throughout the scanning session so that participants completed one run of the oddity task after every two runs of the main experimental task.

Brain images were acquired with a Philips Achieva 3.0-T scanner with a 32-channel head coil. BOLD contrast functional images were collected with a T2*-weighted, gradient EPI sequence (repetition time = 2000 msec, echo time = 35 msec, flip angle = 90°, field of view = 240 mm × 240 mm, acquisition matrix = 96 × 96, in-plane resolution = 2.5 mm × 2.5 mm, slice thickness = 2.5 mm, no slice gap). Volumes consisted of 28 slices angled −30° from the AC–PC plane to maximize signal over the medial-temporal lobes. Volumes were positioned to completely cover the temporal and occipital lobes at the expense of the dorsal parietal cortex. A high-resolution T1-weighted anatomical image was also acquired for each participant (3-D magnetization prepared rapid gradient-echo sequence; 175 slices, voxel size = 1 mm isotropic, field of view = 256 mm × 256 mm, repetition time = 8.4 msec, echo time = 3.8 msec, flip angle = 8°). Stimuli were displayed on a Cambridge Research Systems BOLDScreen located behind the scanner bore and were viewed via a mirror fixed to the head coil. Presentation of the stimuli was controlled by Psychtoolbox (Brainard, 1997) running on MATLAB (The MathWorks, Natick, MA).

Image Preprocessing and Analysis

Functional MRI data were preprocessed with SPM8 (Wellcome Department of Imaging Neuroscience, London, UK; www.fil.ion.ucl.ac.uk/spm/software/spm8/) and included rigid body realignment, coregistration, tissue segmentation, normalization to the Montreal Neurological Institute (MNI) 152 template with DARTEL (Ashburner, 2007) and spatial smoothing (6-mm FWHM Gaussian kernel).

We localized face-selective regions for each individual with data collected from the oddity task. Estimates of the BOLD response in each voxel and category were derived by entering the boxcar function of stimulation that was convolved with the canonical hemodynamic response into a fixed effects general linear model. Face selectivity in each voxel was calculated by contrasting activity evoked by faces against the average of scenes and objects.

Table 1. Mean MNI Peak Coordinates for Each ROI

	Mean MNI Coordinates			SD		
	x	y	z	x	y	z
rOFA	42.4	−78.0	−6.1	6.7	6.1	4.2
lOFA	−41.1	−80.9	−7.63	6.4	4.8	5.9
rFFA	43.3	−49.6	−20.5	3.8	5.7	4.2
lFFA	−41.9	−51.6	−21.0	3.1	4.9	2.7
rAFP	37.3	−14.0	−38.5	5.0	2.5	7.6
lAFP	−38.5	−13.2	−33.0	4.3	5.8	7.4

Face-selective ROIs were localized by finding the most face-selective voxel within expected regions of cortex (OFA, inferior or mid-occipital gyrus; FFA, mid-fusiform gyrus; ATFP, anterior occipito-temporal sulcus or anterior collateral sulcus) near to typical MNI coordinates identified in previous studies (Julian, Fedorenko, Webster, & Kanwisher, 2012; right OFA [rOFA]: 44, −76, −12; left OFA [lOFA]: −40, −76, −18; right FFA [rFFA]: 38, −42, −22; left FFA: −40, −52, −18; Axelrod & Yovel, 2013; right ATFP [rATFP]: 34, −10, −39; left ATFP: −34, −11, −35). ROIs were defined by selecting all significant ($p < .001$, uncorrected), contiguous voxels centered around the peak voxel closest to the coordinates provided above. For analyses of response profiles in the main experiment, ROI size was limited to 50 voxels because previous studies have shown that regions larger than this do not fully capture category selectivity (Mur et al., 2012).

Estimates of the response to each of the 20 categories presented in the 1-back task were modeled separately as instantaneous neural events (i.e., duration = 0 msec) convolved with the canonical hemodynamic response. An additional nuisance regressor of no interest was included to model responses to the initial trial and to all target trials. The values of the beta estimates for each category were averaged over all voxels included in each ROI.

Results

Behavioral Performance

Average performance on the oddity task was 81% (*SEM* = 3%) correct for faces, 83% (*SEM* = 3%) correct for scenes, 82% (*SEM* = 3%) correct for objects, and 76% (*SEM* = 3%) correct for shapes. One-way ANOVA showed no significant differences in performance between categories of stimuli ($p > .2$). Average performance on the 1-back task was 83% (5%).

Definition of ROIs

The right-hemisphere ATFP was localized in 13 of 20 participants, and the left anterior face patch (AFP) was local-

ized in 15 of 20 participants. Localizing the ATFP is problematic because of signal loss in the anterior temporal lobes (because of proximity to air-filled spaces such as the ear canal and the sinus cavity), and finding this region in 60–75% of participants is consistent with previous studies that used a single-session protocol (Axelrod & Yovel, 2013; Rajimehr et al., 2009). FFA was localized bilaterally in all 20 participants, whereas lOFA or rOFA was localized in 19 of 20 participants (bilaterally in 18 participants). Mean (±*SD*) peak coordinates for each ROI are presented in Table 1. Individual MNI coordinates for the peak voxels within the ATFP are provided in Table 2. Figure 2 illustrates the location of the ATFP in four representative participants.

Event-related Response Profiles

Of the 20 stimulus categories tested in the main experiment, faces evoked the maximal response in all of the independently defined functional ROIs (OFA, Figure 3; FFA, Figure 4; ATFP, Figure 5). In line with previous fMRI studies of category selectivity (Mur et al., 2012; Downing

Table 2. Individual MNI Peak Coordinates for the Left and Right AFP

Participant	Left AFP			Right AFP		
	x	y	z	x	y	z
1	−40	−12.5	−35	40	−17.5	−22.5
2	−37.5	−20	−22.5	30	−15	−37.5
3	−	−	−	45	−17.5	−22.5
4	−40	−12.5	−30	32.5	−12.5	−42.5
5	−30	−22.5	−32.5	−	−	−
6	−35	−7.5	−32.5	−	−	−
8	−37.5	−17.5	−25	35	−15	−37.5
11	−45	−12.5	−45	45	−15	−42.5
12	−35	−10	−32.5	37.5	−12.5	−50
13	−40	−17.5	−32.5	37.5	−10	−40
14	−35	−5	−47.5	37.5	−12.5	−42.5
15	−37.5	−12.5	−32.5	40	−12.5	−40
16	−37.5	−7.5	−42.5	−	−	−
17	−	−	−	27.5	−12.5	−40
18	−42.5	−17.5	−22.5	−	−	−
19	−45	−5	−35	40	−12.5	−40
20	−40	−17.5	−27.5	37.5	−17.5	−42.5
Mean	−38.5	−13.2	−33.0	37.3	−14.0	−38.5
SD	4.3	5.8	7.4	5	2.5	7.6

Figure 2. Location of the AFP (red; faces > scenes + objects, $p < .001$).
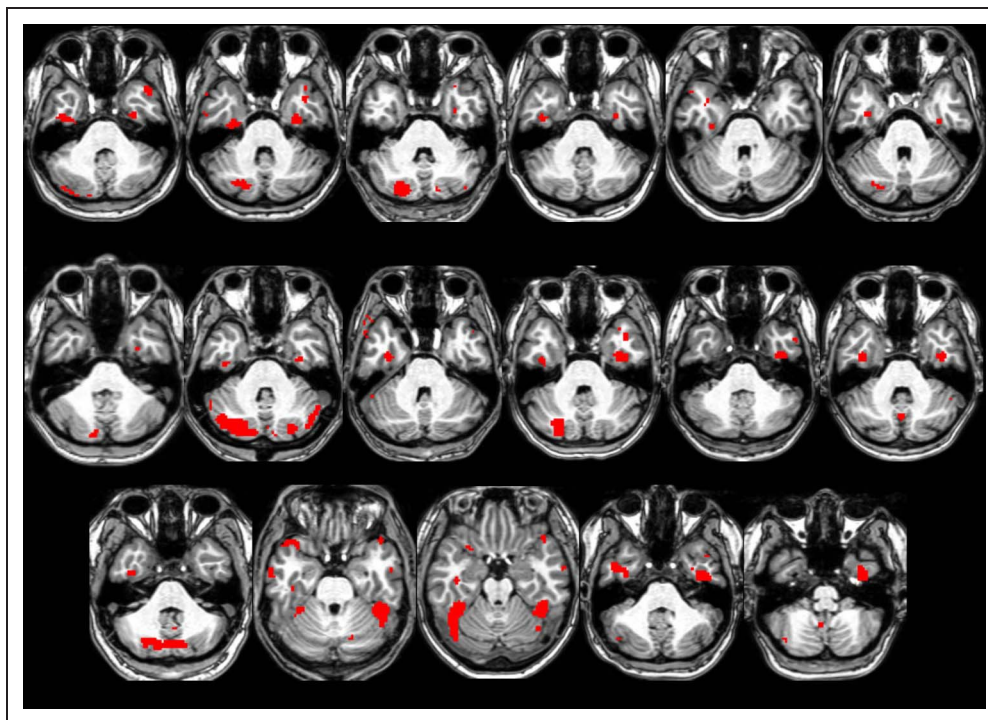
et al., 2006), we expected significantly stronger responses to faces compared with all other categories. To test the selectivity of each ROI, we compared the response to faces against the response to the next most effective stimulus category in a $3 \times 2 \times 2$ repeated-measures ANOVA with ROI, Hemisphere, and Stimulus category as factors. This analysis revealed a significant three-way interaction between ROI, Hemisphere, and Stimulus category ($F(2, 18) = 4.5, p < .05$), a significant two-way interaction between ROI and Stimulus category ($F(2, 18) = 6.03, p < .01$), and significant main effects for ROI ($F(2, 18) = 29.83, p < .001$), Hemisphere ($F(1, 9) = 11.70, p < .01$), and Stimulus category ($F(1, 9) = 8.19, p < .05$).

To interpret the three-way interaction, we carried out two separate 3×2 repeated-measures ANOVAs for each hemisphere with ROI (OFA, FFA, ATFP) and Stimulus category (face vs. next best category) as factors. For the left hemisphere, this analysis found only significant main effects of ROI ($F(2, 26) = 22.75, p < .001$, Bonferroni-corrected) and Stimulus category ($F(1, 13) = 9.34, p < .05$, Bonferroni-corrected). Analysis of the right hemisphere revealed a significant interaction between ROI and Stimulus category ($F(2, 24) = 8.1, p < .01$, Bonferroni-corrected) and significant main effects of ROI ($F(2, 24) = 17.97, p < .001$, Bonferroni-corrected) and Stimulus category ($F(1, 12) = 10.18, p < .01$, Bonferroni-corrected). Simple effects analysis revealed a main effect of stimulus category in the rOFA ($F(1, 18) = 7.2, p < .05$, Bonferroni-corrected) and FFA ($F(1, 19) = 7.2, p < .05$, Bonferroni-corrected), but not in the rATFP ($F(1, 13) = 0.15$). Post hoc tests revealed that the response averaged across faces and bodies in the rATFP was significantly higher

than the next most effective category (spiders; $F(1, 13) = 9.3, p < .01$).

A final analysis examined whether the structure of nonpreferred responses in the ATFP was similar to that found in the posterior face-selective regions. It is a known property of the latter regions that they generally respond more strongly to animate compared with inanimate categories of stimuli (Wiggett, Pritchard, & Downing, 2009; Downing et al., 2006). Thus, for each ROI, we compared the response averaged over all nonhuman animate categories (birds, fish, insects, mammals, reptiles) against the response averaged over all inanimate categories (cars, chairs, clothes, crystals, flowers, fruit and vegetables, instruments, prepared foods, tools, weapons). To test the preference of each ROI, we compared the response to animates and inanimates in a $3 \times 2 \times 2$ repeated-measures ANOVA with ROI, Hemisphere, and Animacy (animate, inanimate) as factors. This analysis revealed significant two-way interactions between ROI and Animacy ($F(2, 18) = 18.7, p < .001$) and between Hemisphere and Category ($F(2, 9) = 10.4, p = .01$) and significant main effects for ROI ($F(2, 18) = 29.01, p < .001$), Hemisphere ($F(1, 9) = 6.8, p < .05$), and Animacy (animates > inanimates, $F(1, 9) = 32.6, p < .001$).

To interpret the interaction effects, we carried out three separate 2×2 repeated-measures ANOVAs for each ROI (OFA, FFA, ATFP) with Hemisphere and Animacy as factors. This analysis found only significant main effects for Animacy in the FFA ($F(1, 19) = 59.13, p < .001$, Bonferroni-corrected) and ATFP ($F(1, 10) = 8.7, p < .05$, Bonferroni-corrected). In contrast, analysis of the OFA revealed a significant main effect of Hemisphere ($F(1, 18) = 11.84$,

$p < .01$, Bonferroni-corrected) and Animacy ($F(1, 18) = 62.3, p < .001$, Bonferroni-corrected) and a two-way interaction between Hemisphere and Animacy ($F(1, 18) = 11.13, p < .01$, Bonferroni-corrected). Simple effects analysis revealed a significant effect of Animacy in both the rOFA ($F(1, 18) = 54.6, p < .001$, Bonferroni-corrected) and lOFA ($F(1, 18) = 40.0, p < .001$, Bonferroni-corrected).

Discussion

Face-selective regions across the temporal lobes showed a similar profile of activity, across a wide range of stimulus kinds, consistent with a model in which these regions cooperate functionally (Moeller et al., 2008). There was variation, however, in the pattern of responses across the regions tested; in particular, the right-hemisphere ATFP showed a strong response to both faces and bodies (statistically different from the response to spiders, the next most effective category) despite the fact that the ROI was localized independently on the basis of a contrast of faces versus scenes and objects. This pattern of colocalized significant response to bodies and faces, along with a weaker response to other kinds of objects, is highly similar to that found in the right fusiform gyrus. In that region, strong fMRI responses to faces and bodies have been found to overlap closely (Weiner & Grill-Spector, 2010; Peelen & Downing, 2005). One account of this general finding is that it reflects the common co-occurrence of faces and bodies in the visual input and the need to jointly process the socially relevant information they provide (Peelen & Downing, 2007).

Several studies have attempted to determine whether the fusiform face- and body-selective responses reflect a single neural system or rather two distinct ones. Schwarzlose, Baker, and Kanwisher (2005) used high-resolution imaging to show that, in many participants, alongside "shared" voxels that respond to both categories, it is possible to identify distinct, but adjacent, highly selective patches for faces and bodies, referring to these accordingly as FFA and "fusiform body area" (FBA). Another approach

Figure 3. Selectivity profile for rFFA (top) and lFFA (bottom). Error bars show the *SEM*.
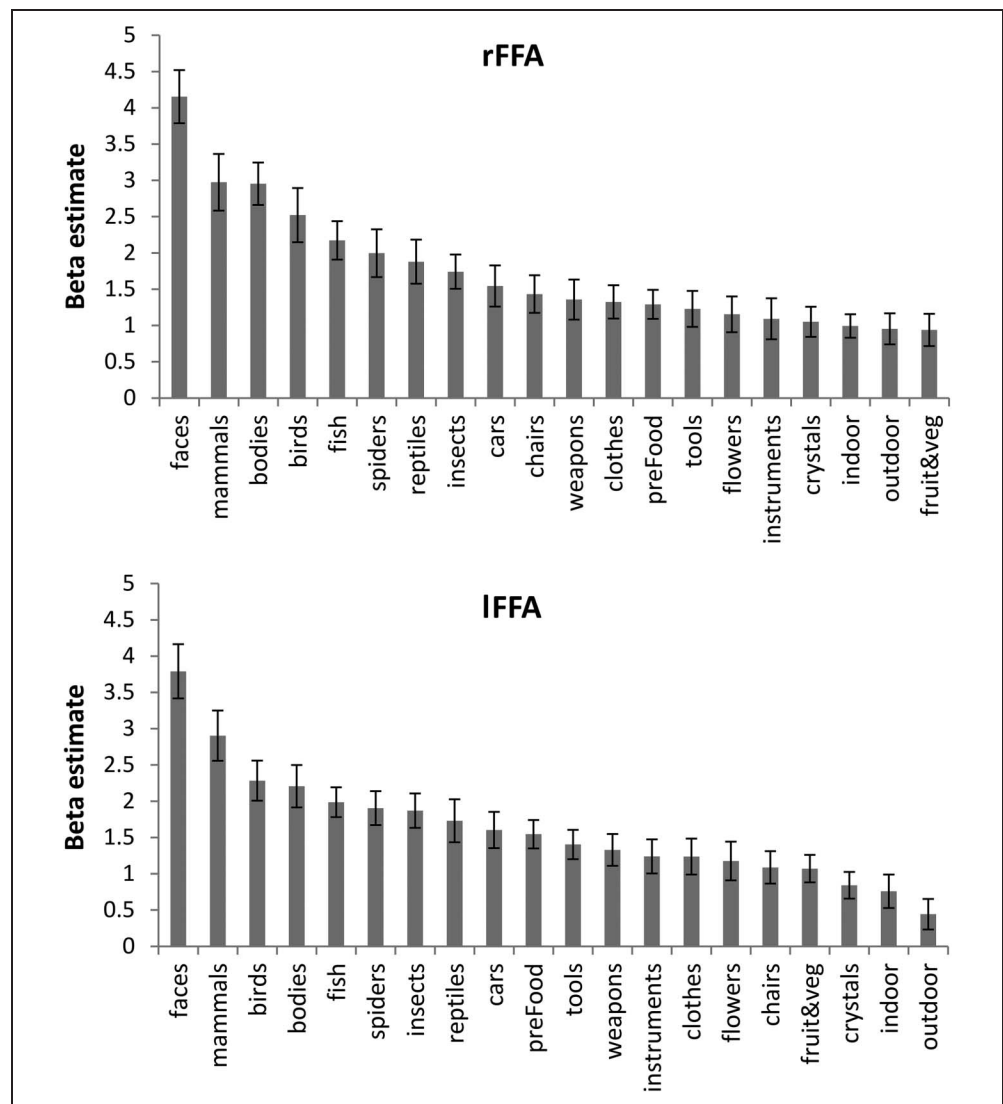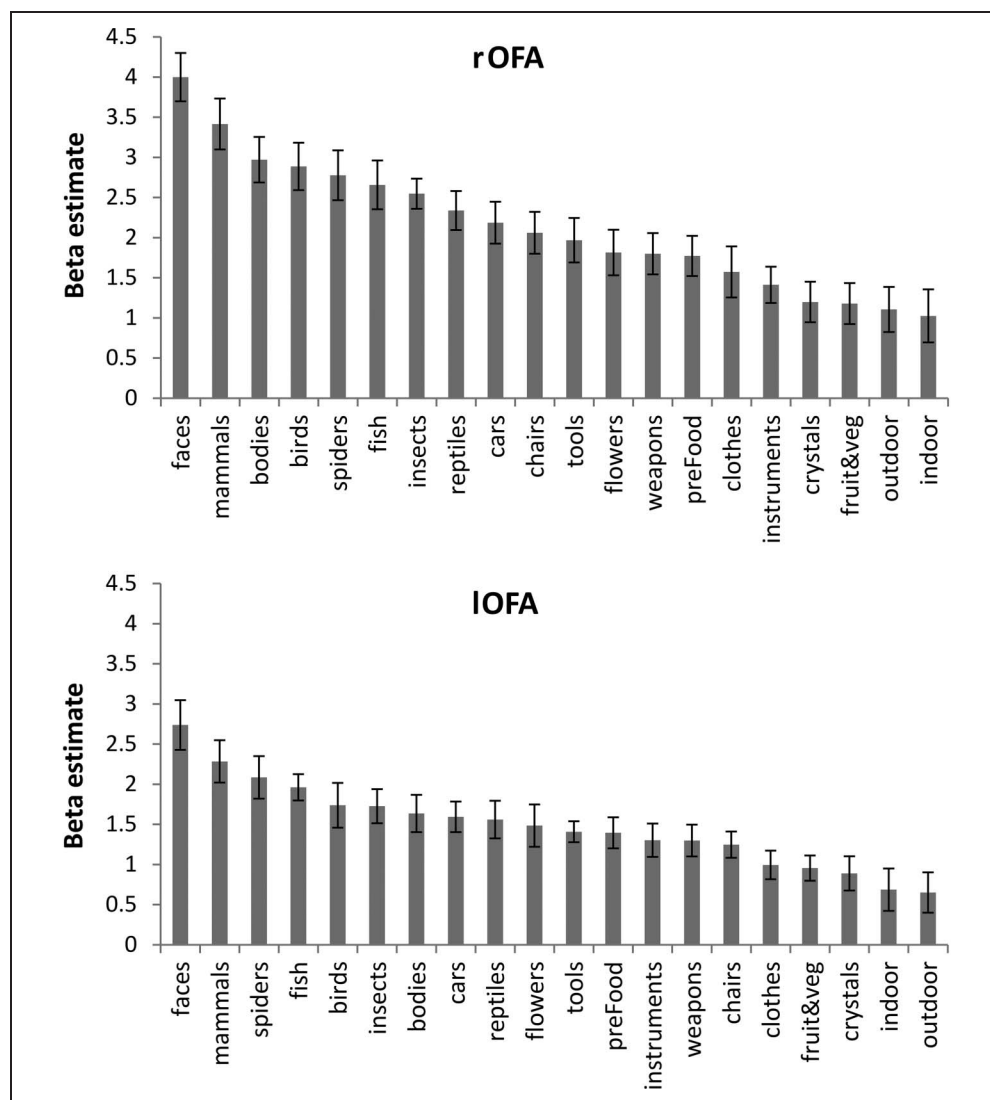
Figure 4. Selectivity profile for rOFA (top) and lOFA (bottom). Error bars show the *SEM*.

tests for distinct neural systems at the pattern level, without requiring that they be identified in a binary fashion with separate sets of voxels. The logic of this method is that overlapping voxels (at whatever resolution) need not reflect shared neural processes—an assumption commonly made in fMRI research (Peelen & Downing, 2007).

For example, in a region where there are overlapping but functionally distinct face and body representations, local patterns of selectivity to these two categories should be uncorrelated (or negatively correlated). That is, considered across a set of voxels, variability in the selectivity for bodies would not be expected to relate systematically to variability in the selectivity for faces. In contrast, where there are two overlapping and integrated representations, the variability in selectivity to these two categories would be expected to be related across voxels: Strong selectivity to one category should tend to predict strong selectivity to the other. This would result in a positive correlation between the local patterns of selectivity evoked by each category. Studies taking this approach

have found evidence for independent fusiform face- and body-selective representations (Kim, Lee, Erlendsdottir, & McCarthy, 2014; Weiner & Grill-Spector, 2010; Peelen, Wiggett, & Downing, 2006; see also Downing, Wiggett, & Peelen, 2007).

Thus, motivated by these previous findings in the extrastriate cortex and by our present results in the right-hemisphere ATFP, in Experiment 2, we used the multivoxel approach described above to examine whether faces and bodies recruit overlapping or segregated representations in the anterior temporal lobes and in the fusiform gyrus.

EXPERIMENT 2

We localized face- and body-selective regions (FFA, ATFP, FBA) in the right-hemisphere mid-fusiform and anterior temporal regions with a blocked 1-back design. Only responses in the right hemisphere were examined, as the rATFP demonstrated strong responses to both faces and bodies in Experiment 1. We opted for a block

design 1-back localizer task in Experiment 2 because a pilot oddity task including headless bodies was too difficult for participants to complete. Then, we assessed the functional responses in each ROI to six different conditions (faces, whole bodies, body parts, mammals, foods, and tools) with an event-related design. These categories were selected to assess responses across animate, natural, and manmade objects. Moreover, body parts were included to examine the breadth of body responses in ATFP. We improved the experimental protocol with a coronal slice orientation, which, compared with axial orientation (as used in Experiment 1), has been shown to maximize signal over the anterior temporal lobes (Axelrod & Yovel, 2013), and we obtained higher-resolution images (voxel size = 2 mm^3) to mitigate partial voluming effects.

Methods

Ten healthy postgraduate volunteers (mean age = 25 years, range = 24–29 years; six women) were recruited from Bangor University. All participants were screened for MRI exclusion criteria, after which they gave written informed consent for participation in the experiment, which was approved by the research ethics committee of the School of Psychology at Bangor University, United Kingdom.

Materials

Stimuli for the localizer consisted of 40 images each of faces, bodies, and chairs (Downing et al., 2006). Stimuli for the main experiment consisted of 24 images each of faces, (headless) bodies, body parts, (nonhuman) mammals, food, and tools (Figure 6). All images were prepared in a similar manner as Experiment 1. None of the images presented in the localizer task appeared in the event-related runs.

Procedure

Participants completed five runs of a 1-back localizer task, each consisting of 25 blocks. Blocks 1, 5, 9, 13, 17, 21,

Figure 5. Selectivity profile for right (top) and left (bottom) AFP. Error bars show the *SEM*.
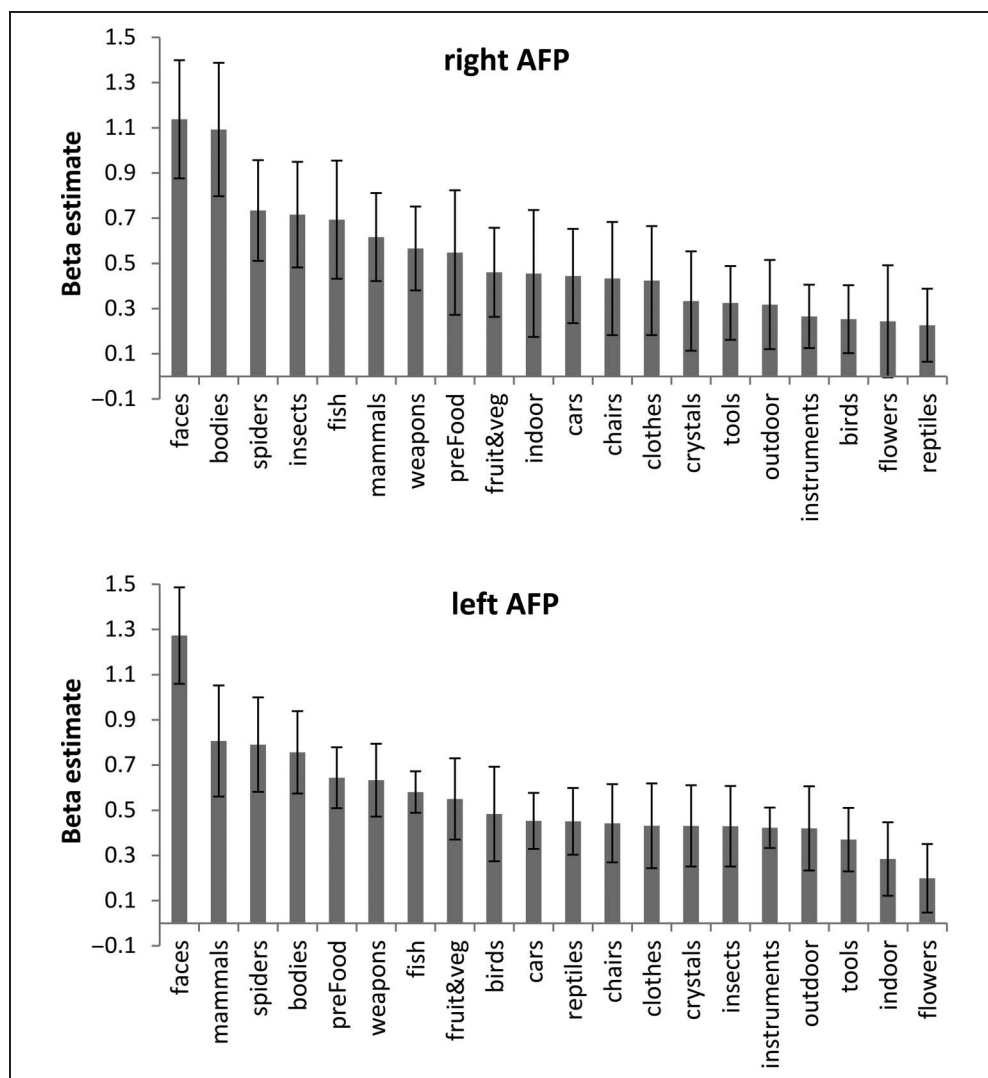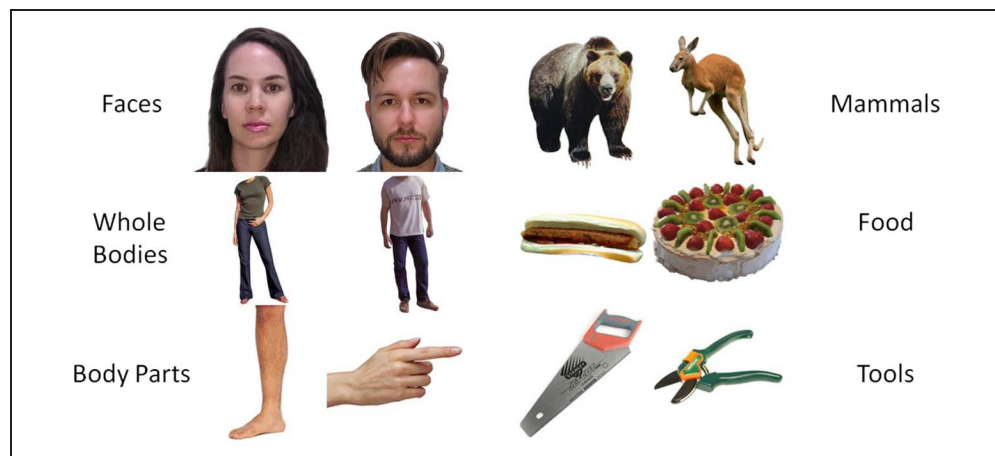
Figure 6. Example stimuli presented in Experiment 2.

and 25 were fixation-only rest conditions lasting 10 sec in duration. Each of the three stimulus blocks (faces, bodies, and chairs) was presented once between each pair of rest blocks. Stimulation blocks were composed of 15 stimulus exemplars drawn from a pool of 40 images. Stimuli were presented sequentially and appeared for 300 msec followed by a 700-msec ISI. Repetitions occurred twice per block.

For the main experimental runs, all images from each of the six stimulus categories were presented in an event-related design; each stimulus was presented once per run. Participants performed a 1-back task. Stimulus order for each run was determined with a first-order counterbalanced, optimized, $n = 8$, Type 1, Index 1 sequence. This procedure generated a sequence of 193 events of eight types, including an initial event to establish sequence context. Event types 1–6 were assigned to the six stimulus classes, Event type 7 was assigned to target events (1-back, stimulus repetition), and Event type 8 was assigned to fixation-only rest condition. Participants completed three runs of the main experiment, resulting in 72 trials per condition. Stimulus and target events were presented for 300 msec followed by an ISI of 1700 msec consisting of a fixation cross. For rest events, a fixation cross was presented for 2000 msec. Fixation blocks (duration = 16 sec) were presented at the beginning and end of each run. Runs of the localizer and the main experiment were completed in an alternating sequence.

Brain images were acquired with a Philips Achieva 3.0-T scanner with a 32-channel head coil. BOLD contrast functional images were collected with a T2*-weighted, gradient EPI sequence (repetition time = 2500 msec, echo time = 35 msec, flip angle = 90°, field of view = 240 mm × 240 mm, acquisition matrix = 120 × 120, in-plane resolution = 2 mm × 2 mm, slice thickness = 2 mm, no slice gap). Volumes were composed of 28 slices in coronal orientation that were split into two separate stacks of 14 slices to cover the anterior temporal lobes and the mid-fusiform gyrus (Axelrod & Yovel, 2013). A high-resolution T1-weighted anatomical image

was also acquired for each participant. All other aspects of the experimental setup were the same as in Experiment 1.

Image Preprocessing and Analysis

Preprocessing was similar to Experiment 1 except that, to better preserve the local spatial patterns of brain activity, images were not normalized and spatial smoothing was performed with a 3-mm Gaussian kernel. Responses to each category in the blocked localizer and event-related runs were derived in a similar manner as Experiment 1. To evaluate mean univariate response profiles in selective regions similar to Experiment 1, the FFA, ATFP, and FBA were defined from the localizer by contrasting each category against chairs (faces > chairs, bodies > chairs; $p <$ .001, uncorrected).

For the pattern analysis, localizer data were further used to identify two functional ROIs in both the mid-fusiform and anterior temporal lobes for the purpose of pattern analyses. First, a broad "human form" selective ROI was defined as the union of all face- and body-selective voxels within the mid-fusiform and collateral sulcus. This combined ROI was examined to ensure that the results of the pattern analysis would not be biased toward one stimulus category owing to unbalanced voxel selection. Second, we examined only the face-selective voxels corresponding to the ATFP. For this analysis, we defined the ATFP, as in Experiment 1, as the 30 most face-selective voxels (i.e., faces > bodies + chairs, $p <$ .001) that were contiguous with the peak voxel residing within the collateral sulcus. For both ROIs, an independent measure of face, whole body, and body part selectivity was calculated from the main experiment data by contrasting the response of each of these conditions against the average of mammals, food, and tools. The resultant t values were extracted for all voxels residing within each ROI. For each participant and ROI, the extracted pattern of t values quantifying face selectivity was correlated with the corresponding pattern of t values for whole bodies and body parts.

Results

Behavioral Performance

Average performance in the localizer scans was 75.1% (*SEM* = 4.2%), and average performance in the main experiment was 78.8% (*SEM* = 7.5%).

Univariate Analysis of Main Experiment

The three independently defined ROIs (rFFA, right FBA, and rATFP) all showed maximal responses to the ex-

pected categories: faces in FFA and ATFP and whole bodies in FBA (Figure 7). To examine whether the preferred category evoked significantly more activity than all other stimuli, the responses to the preferred category and the next most effective category were entered into a 3 × 2 repeated-measures ANOVA with ROI (FFA, FBA, ATFP) and Stimulus category (preferred category vs. next most effective) as factors. For the FFA and the ATFP, responses to faces were compared against mammals and bodies, respectively, whereas in the FBA, bodies were

Figure 7. Profile of responses to six stimulus categories, for the rFFA (top), right FBA (center), and right AFP (bottom). Error bars show the *SEM*.
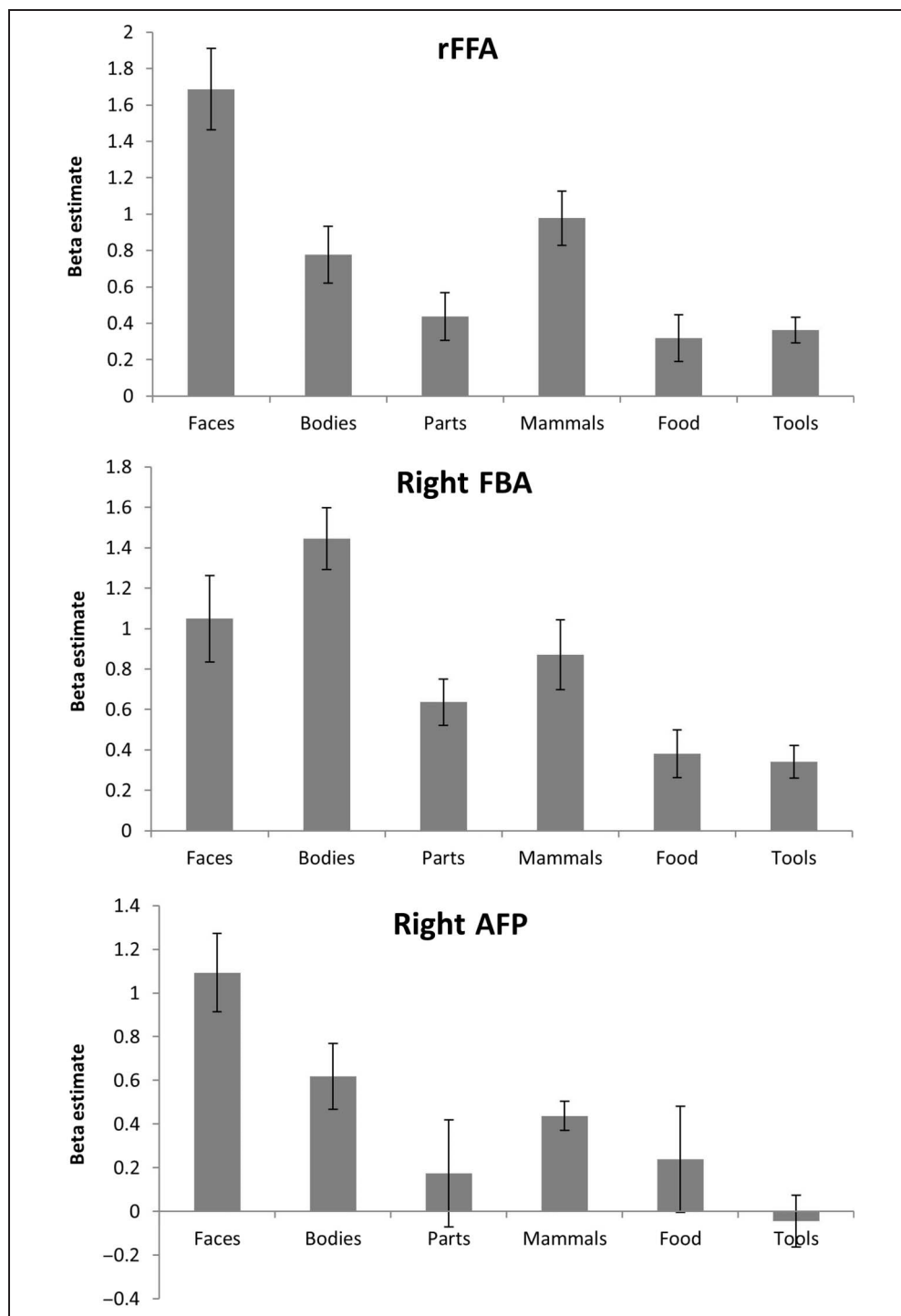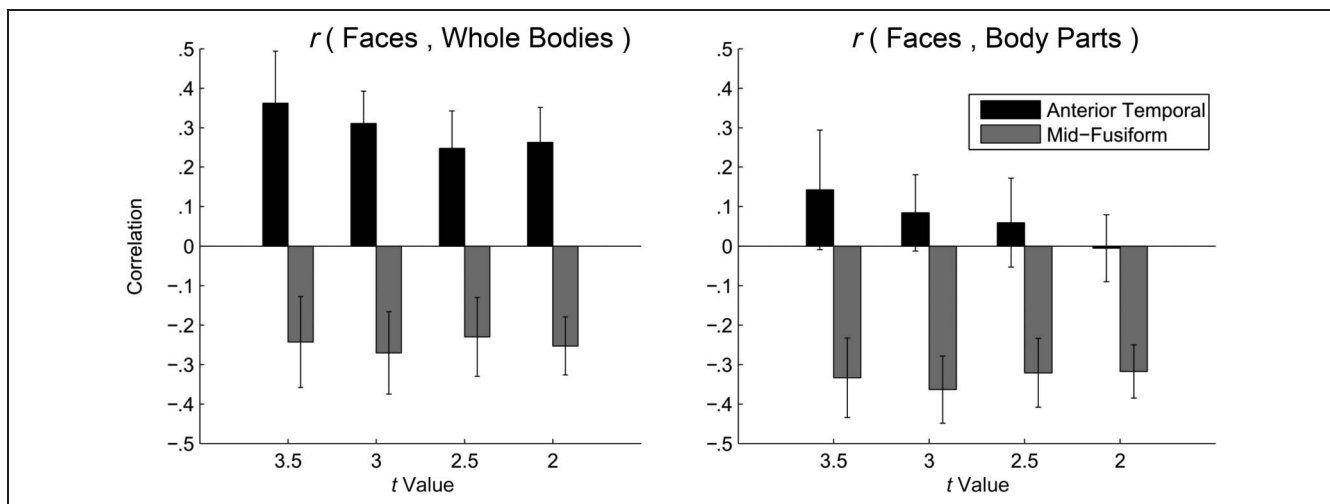
Figure 8. Correlation between face and body selectivity (left) and face and body-part selectivity (right) for the union of all independently defined face- and body-selective voxels in the mid-fusiform and anterior temporal lobes. Correlations shown were obtained for a range of thresholds for defining face- and body-selective voxels (x axis). These findings suggest that, in the anterior temporal cortex, face and body representations are integrated, in contrast to the fusiform gyrus, where they appear to remain distinct. Error bars show the *SEM*.

compared with faces. This analysis revealed only a significant main effect of Stimulus category ($F(1, 7) = 53.69$, $p < .001$), indicating that faces and whole bodies evoked more activity than the next most effective stimulus category in face- and body-selective regions, respectively.

Pattern Analysis

We performed a pattern analysis to examine the relationship of face- and body-selective populations in both the right-hemisphere mid-fusiform and anterior temporal lobes. In the first analysis, we examined all voxels that showed a preference for either faces or bodies by taking the union of the face- and body-selective regions as defined by the localizer runs. These regions were defined for a range of thresholds ($t_{ROI} > 3.5, 3, 2.5, 2$) to ensure that the results of the pattern analysis were not dependent on how stringently voxels were selected. Separately for the mid-fusiform and anterior temporal ROIs, voxel-wise patterns of selectivity in the main experiment for faces and for whole human bodies (faces > mammals + food + tools, bodies > mammals + food + tools) were correlated for each participant. Single-sample t tests comparing the Fisher-transformed correlations in each region (Figure 8, left) showed that face and body selectivity were negatively correlated in the mid-fusiform ROI (significantly at thresholds $t_{ROI} > 3, p = .0321; T_{ROI} > 2$, $p < .01$, Bonferroni-corrected). This suggests, consistent with previous findings (Kim et al., 2014; Peelen et al., 2006), that face and body representations remain distinct in the fusiform gyrus—faces and bodies elicit distinct patterns of local activity. In contrast, significantly positive correlations between face and body selectivity were observed in the anterior temporal lobes (*r*s range = 0.24–0.36, all *p*s at different selection thresholds < .05,

Bonferroni-corrected), suggestive of a shared face and body representation.

In principle, however, the positive pattern correlation between face and body selectivity in the anterior temporal lobe could be because of general factors affecting responses in this region (i.e., signal limitations, use of common baseline), rather than a specific property of face and body representations. To exclude this possibility, we performed the same analysis comparing patterns of face and body-part selectivity (Figure 8, right). This analysis revealed that face and body-part selectivity were negatively correlated in the mid-fusiform region (all *p*s < .01, Bonferroni-corrected). Critically, face and body-part selectivity were not significantly correlated in the anterior temporal lobes (all *t*s < 1). Furthermore, in the anterior temporal region, the correlation between face and whole-body selectivity was significantly greater than the correlation between face and body-part selectivity (*p*s < .05, Bonferroni-corrected). Therefore, it is not simply the case that response patterns evoked by any and all visual stimuli are positively correlated within the anterior temporal region.

Although examining the union of all face- and body-selective voxels is an unbiased method of voxel selection, this approach is less directly comparable with previous studies of the human ATFP. Therefore, in a second pattern analysis, we limited the analyzed region to only the 30 most face-selective voxels (selected from the localizer data) located within the collateral sulcus. Overall, the correlations observed for the ATFP were similar to those found in the first pattern analysis. The mean correlation between face and whole-body selectivity was significantly above zero ($r_{mean} = .38, t = 4.73, p < .01$, Bonferroni-corrected), whereas the mean correlation between faces and body parts was not ($r_{mean} = .08, t = 0.86$). Moreover,

the correlation between face and whole-body selectivity was greater than between faces and body parts ($t =$ 8.05, $p < .0001$, Bonferroni-corrected), indicating that the spatial organization of face- and body-selective responses observed for the combined ROI is also present in the ATFP.

Discussion

In Experiment 2, we did not replicate the finding of Experiment 1 that faces and headless bodies drive activity in the rATFP equally well, perhaps because, in this study, we scanned at a higher spatial resolution and optimized slice selection, resulting in more precise localization of highly face-selective voxels in the anterior temporal cortex. The multivoxel pattern analysis showed, however, that, in this region, there is a significant positive correlation between the activity patterns evoked by face and whole-body stimuli. Importantly, this was significantly greater than the relationship between faces and individual body parts, which was not greater than chance, indicating that the positive correlation was not because of general properties of the responses observed in this region. Furthermore, these results were distinct from the findings from the mid-fusiform region, where patterns of face and body selectivity were negatively correlated, indicating independent (or at least less integrated) encoding of these stimuli. Taken together, these results suggest that part of the right anterior temporal lobe, in contrast to the posterior fusiform gyrus, encodes an integrated representation of the visual appearance of the face and of the whole body—a representation that does not extend to isolated body parts.

GENERAL DISCUSSION

Our study aimed to survey the response profile of the ATFP (Rajimehr et al., 2009) to different categories of stimuli. The profile of ATFP across multiple categories appeared, to a large extent, to mirror that of the FFA and OFA. A notable finding from Experiment 1 was the high response to human bodies (without faces) in the independently defined right-hemisphere ATFP ROI. In previous work on human anterior temporal face representations, Tsao and colleagues (2008) found a reliable response to bodies that was nonetheless significantly lower than the response to faces. These findings of colocated strong responses to faces and bodies are another potentially important similarity between the visual representations of the anterior and posterior inferior temporal lobes, where face- and body-evoked activations are intertwined in the fusiform gyrus (Kim et al., 2014; Weiner & Grill-Spector, 2010; Peelen & Downing, 2005; Schwarzlose et al., 2005).

Therefore, in Experiment 2, we used a pattern analysis to examine the spatial organization of face and whole-body representations in the anterior and posterior tem-poral regions. In the fusiform gyrus, the pattern of face and body selectivity was negatively correlated, consistent with functionally distinct representations. In contrast, in the anterior temporal region, the correlation between patterns evoked by bodies and faces was positive, sug-gestive of integrated, whole-person representations. This interpretation is supported by the finding that face-evoked activity in the rATFP was more similar to whole bodies compared with isolated body parts, suggesting that anterior person-selective responses are primarily driven by whole-person form information.

These findings, considered alongside a recent study showing super-additive responses to combined face and whole-body stimuli in the ATFP of the macaque (Fisher & Freiwald, 2015), provide support for the existence of integrated whole-agent processing regions in the primate visual system. In principle, integrated processing is not required to form whole-agent representations (Afraz, 2015), as whole-agent information is also present in the distributed response across face- and body-selective regions in the mid-fusiform. Evidence of super-additive responses and joint selectivity, however, confirms that integrated representations are indeed formed in the ante-rior temporal cortex and likely play a key role in visual recognition (Lehky & Tanaka, 2016).

More specifically, these and other findings provide some evidence for a hierarchically organized chain of face- and body-form representations along the length of the ventral occipito-temporal cortex (Grill-Spector & Weiner, 2014; Taylor & Downing, 2011; Minnebusch & Daum, 2009) that become progressively more integrated with each other (Figure 9). In the posterior occipito-temporal cortex, face and body representations (OFA, EBA) are anatomically distinct (Pitcher et al., 2009) and appear to emphasize the representation of component parts (e.g., Schiltz et al., 2010; Liu et al., 2009; Taylor, Wiggett, & Downing, 2007). More anteriorly, the FFA and FBA encode more holistic properties of their pre-ferred stimuli (Brandman & Yovel, 2016; Harris & Aguirre, 2010; Schiltz et al., 2010; Liu et al., 2009; Taylor et al., 2007; Schiltz & Rossion, 2006). A question under active inves-tigation is whether, in the fusiform regions, there is fur-ther integrated processing across faces and bodies (Bernstein, Oron, Sadeh, & Yovel, 2014; Kaiser, Strnad, Seidl, Kastner, & Peelen, 2014; Song, Luo, Li, Xu, & Liu, 2013; Schmalzl, Zopf, & Williams, 2012), with some evi-dence both for (e.g., Bernstein et al., 2014) and against (Fisher & Freiwald, 2015; Kaiser et al., 2014) this proposal. The current results argue for even closer integration of face and whole-body representation in the anterior-most reaches of the temporal cortex. On this view, the anterior temporal cortex can be seen as a core region in a broadly defined person-form processing pathway that combines domain-specific representations that are constructed in the extrastriate cortex.

This perspective integrates the current findings with previous behavioral work and suggests how the very

Figure 9. Schematic illustration of a proposed hierarchical organization of human form information in the ventral temporal lobes. EBA and OFA, which are anatomically distinct, engage in domain-specific part-based processing of their preferred categories. FFA and FBA are closely overlapping and form domain-specific holistic representations of faces and bodies, respectively. Finally, face and body information is functionally integrated in the anterior temporal lobes (including the region typically identified as the AFP) to contribute to whole-person representation. Red clusters, faces > chairs, $p < .001$; green clusters, bodies > chairs, $p < .001$.
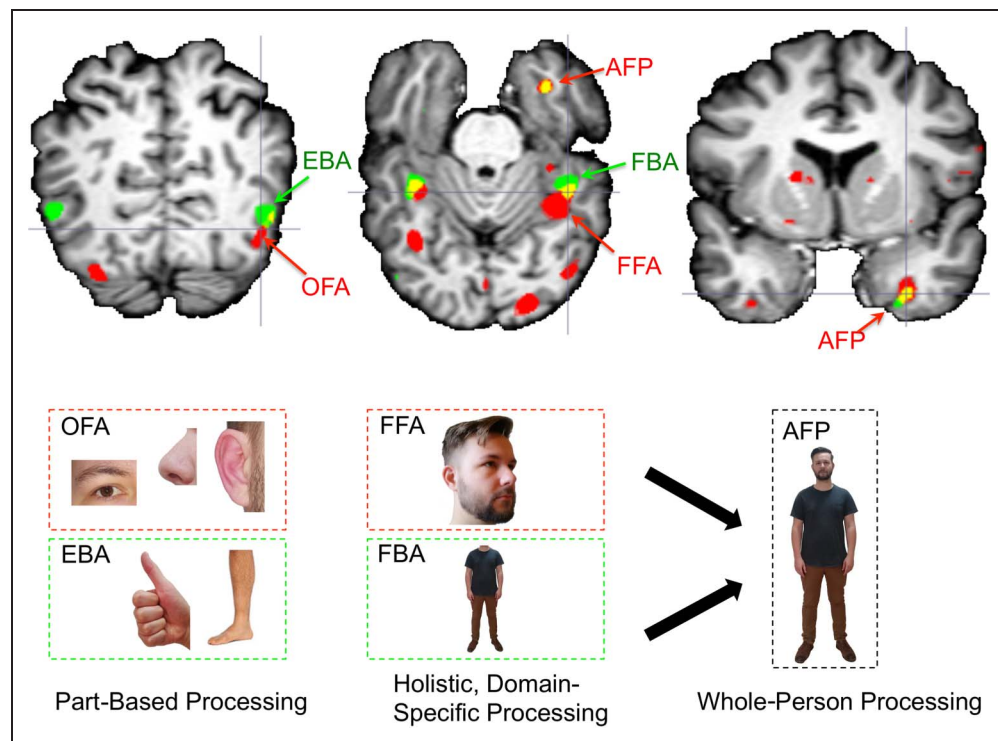
different kinds of visual cues provided by bodies and faces are brought together to represent what must be the real object of interest for social perception: whole people (Macrae, Quinn, Mason, & Quadflieg, 2005). Evidence from perceptual studies shows that judgments of identity, emotion, and gender from faces can be strongly influenced by the state of the body (Rice, Phillips, Natu, An, & O'Toole, 2013; Aviezer, Trope, & Todorov, 2012). Moreover, adaptation to pictures of bodies presented in isolation can alter the perception of subsequently viewed faces (Ghuman, McDaniel, & Martin, 2010), suggesting that faces and bodies share processing mechanisms. A very important open question is to what extent the pathways described here contribute to the integrated processing of different facial and bodily cues. Notably, extensive work by de Gelder and colleagues (2006) points to other, largely subcortical routes that are involved in rapidly extracting and integrating emotional information from faces and bodies (Meeren, van Heijnsbergen, & de Gelder, 2005). Here, we propose that the ATFP forms part of a ventral temporal pathway involved in person perception and identification that integrates static form cues from across the face and the body.

Evidence of joint face and whole-body selectivity in the ATFP also aligns with paired object associative responses found in the PrC (Fujimichi et al., 2010). PrC is an anterior temporal region located in the anterior collateral sulcus (rhinal sulcus; Suzuki & Naya, 2014), which has been argued to occupy the highest point in the ventral visual processing pathway (Murray, Bussey, & Saksida, 2007). PrC receives dense input from visual area TE (Suzuki & Naya, 2014) and also receives input from

regions of STS containing body-form selective neurons (Suzuki & Naya, 2014; Oram & Perrett, 1994). In the macaque, paired-associate learning involving two objects alters the selectivity of neurons in the PrC such that neurons that are selective for a particular object also become selective for a paired associate when repeatedly presented together (Fujimichi et al., 2010). Similar to the present findings, these "unitized" representations also show a hierarchical organization, with a gradient of increasingly overlapping responses found spanning area TE and regions within the PrC (areas A36 and A35; Hirabayashi et al., 2014; Fujimichi et al., 2010). Given that the ATFP was primarily observed in the anterior collateral sulcus, it is possible that the current evidence of integrated face and body processing reflects a specific instance of a general PrC process that forms unitized representations of highly relevant objects via paired-associate coding mechanisms.

Such a hierarchical scheme accords with proposals arising from the literature on memory and conceptual representations. These hold that the PrC is involved in forming complex, conjunctive object representations by combining feature-based representations derived in the extrastriate cortex (Clarke & Tyler, 2014; Graham, Barense, & Lee, 2010; Barense et al., 2005, 2007; Bartko, Winters, Cowell, Saksida, & Bussey, 2007; Buckley & Gaffan, 2006; Lee et al., 2005). According to this account, functionally integrating domain-specific representations serve to help buffer the perceptual system against interference, which has a larger impact on the simpler, part-based representations found in the extrastriate cortex (Watson & Lee, 2013; Barense et al., 2012; Bartko,

Cowell, Winters, Bussey, & Saksida, 2010; Fujimichi et al., 2010).

To summarize, this study demonstrated that the ATFP shares several similarities with other face-selective regions found in the extrastriate cortex. Differences were evident, however, in the way that face- and body-selective responses are organized in these distributed temporal lobe brain areas, with evidence for integrated selectivity of faces and whole bodies found only in the ATFP. In that sense, the present findings are consistent with models that posit a posterior-to-anterior gradient in perceptual representations of objects that is built on increasingly complex combinations of features (Lehky & Tanaka, 2016).

REFERENCES

Afraz, A. (2015). Head to toe, in the head. *Proceedings of the National Academy of Sciences, U.S.A., 112,* 15004–15005.

Aguirre, G. K. (2007). Continuous carry-over designs for fMRI. *Neuroimage, 35,* 1480–1494.

Anzellotti, S., Fairhall, S. L., & Caramazza, A. (2014). Decoding representations of face identity that are tolerant to rotation. *Cerebral Cortex, 24,* 1988–1995.

Ashburner, J. (2007). A fast diffeomorphic image registration algorithm. *Neuroimage, 38,* 95–113.

Aviezer, H., Trope, Y., & Todorov, A. (2012). Body cues, not facial expressions, discriminate between intense positive and negative emotions. *Science, 338,* 1225–1229.

Axelrod, V., & Yovel, G. (2013). The challenge of localizing the anterior temporal face area: A possible solution. *Neuroimage, 81,* 371–380.

Barense, M. D., Bussey, T. J., Lee, A. C. H., Rogers, T. T., Davies, R. R., Saksida, L. M., et al. (2005). Functional specialization in the human medial temporal lobe. *Journal of Neuroscience, 25,* 10239–10246.

Barense, M. D., Gaffan, D., & Graham, K. S. (2007). The human medial temporal lobe processes online representations of complex objects. *Neuropsychologia, 45,* 2963–2974.

Barense, M. D., Groen, I. I. A., Lee, A. C. H., Yeung, L.-K., Brady, S. M., Gregori, M., et al. (2012). Intact memory for irrelevant information impairs perception in amnesia. *Neuron, 75,* 157–167.

Barense, M. D., Henson, R. N. A., Lee, A. C. H., & Graham, K. S. (2010). Medial temporal lobe activity during complex discrimination of faces, objects, and scenes: Effects of viewpoint. *Hippocampus, 20,* 389–401.

Bartko, S. J., Cowell, R. A., Winters, B. D., Bussey, T. J., & Saksida, L. M. (2010). Heightened susceptibility to interference in an animal model of amnesia: Impairment in encoding, storage, retrieval—Or all three? *Neuropsychologia, 48,* 2987–2997.

Bartko, S. J., Winters, B. D., Cowell, R. A., Saksida, L. M., & Bussey, T. J. (2007). Perceptual functions of perirhinal cortex in rats: Zero-delay object recognition and simultaneous oddity discriminations. *Journal of Neuroscience, 27,* 2548–2559.

Bernstein, M., Oron, J., Sadeh, B., & Yovel, G. (2014). An integrated face–body representation in the fusiform gyrus but not the lateral occipital cortex. *Journal of Cognitive Neuroscience, 26,* 2469–2478.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10,* 433–436.

Brandman, T., & Yovel, G. (2016). Bodies are represented as wholes rather than their sum of parts in the occipital–temporal cortex. *Cerebral Cortex, 26,* 530–543.

Buckley, M. J., Booth, M. C. A., Rolls, E. T., & Gaffan, D. (2001). Selective perceptual impairments after perirhinal cortex ablation. *Journal of Neuroscience, 21,* 9824–9836.

Buckley, M. J., & Gaffan, D. (2006). Perirhinal cortical contributions to object perception. *Trends in Cognitive Sciences, 10,* 100–107.

Calder, A. J., Lawrence, A. D., & Young, A. W. (2001). Neuropsychology of fear and loathing. *Nature Reviews Neuroscience, 2,* 352–363.

Clarke, A., & Tyler, L. K. (2014). Object-specific semantic coding in human perirhinal cortex. *Journal of Neuroscience, 34,* 4766–4775.

Collins, J. A., & Olson, I. R. (2014). Beyond the FFA: The role of the ventral anterior temporal lobes in face processing. *Neuropsychologia, 61,* 65–79.

de Gelder, B. (2006). Towards the neurobiology of emotional body language. *Nature Reviews Neuroscience, 7,* 242–249.

Desimone, R., Albright, T. D., Gross, C. G., & Bruce, C. (1984). Stimulus-selective properties of inferior temporal neurons in the macaque. *Journal of Neuroscience, 4,* 2051–2062.

Downing, P. E., Chan, A. W.-Y., Peelen, M. V., Dodds, C. M., & Kanwisher, N. (2006). Domain specificity in visual cortex. *Cerebral Cortex, 16,* 1453–1461.

Downing, P. E., Wiggett, A. J., & Peelen, M. V. (2007). Functional magnetic resonance imaging investigation of overlapping lateral occipitotemporal activations using multi-voxel pattern analysis. *Journal of Neuroscience, 27,* 226–233.

Duchaine, B., & Yovel, G. (2015). A revised neural framework for face processing. *Annual Review of Vision Science, 1,* 393–416.

Fisher, C., & Freiwald, W. A. (2015). Whole-agent selectivity within the macaque face-processing system. *Proceedings of the National Academy of Sciences, U.S.A., 112,* 14717–14722.

Freiwald, W. A., & Tsao, D. Y. (2010). Functional compartmentalization and viewpoint generalization within the macaque face-processing system. *Science, 330,* 845–851.

Fujimichi, R., Naya, Y., Koyano, K. W., Takeda, M., Takeuchi, D., & Miyashita, Y. (2010). Unitized representation of paired objects in area 35 of the macaque perirhinal cortex. *European Journal of Neuroscience, 32,* 659–667.

Gauthier, I., Tarr, M. J., Moylan, J., Skudlarski, P., Gore, J. C., & Anderson, A. W. (2000). The fusiform "face area" is part of a network that processes faces at the individual level. *Journal of Cognitive Neuroscience, 12,* 495–504.

Ghuman, A. S., McDaniel, J. R., & Martin, A. (2010). Face adaptation without a face. *Current Biology, 20,* 32–36.

Graham, K. S., Barense, M. D., & Lee, A. C. H. (2010). Going beyond LTM in the MTL: A synthesis of neuropsychological and neuroimaging findings on the role of the medial temporal lobe in memory and perception. *Neuropsychologia, 48,* 831–853.

Grill-Spector, K., & Weiner, K. S. (2014). The functional architecture of the ventral temporal cortex and its role in categorization. *Nature Reviews Neuroscience, 15,* 536–548.

Harris, A., & Aguirre, G. K. (2010). Neural tuning for face wholes and parts in human fusiform gyrus revealed by fMRI adaptation. *Journal of Neurophysiology, 104,* 336–345.

Haxby, J. V., & Gobbini, M. I. (2011). Distributed neural systems for face perception. In A. J. Calder, G. Rhodes, M. H. Johnson, & J. V. Haxby (Eds.), *The Oxford handbook of face perception* (pp. 93–110). Oxford, UK: OUP.

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences, 4,* 223–233.

Hirabayashi, T., Tamura, K., Takeuchi, D., Takeda, M., Koyano, K. W., & Miyashita, Y. (2014). Distinct neuronal interactions in anterior inferotemporal areas of macaque monkeys during retrieval of object association memory. *Journal of Neuroscience, 34,* 9377–9388.

Hung, C.-C., Yen, C. C., Ciuchta, J. L., Papoti, D., Bock, N. A., Leopold, D. A., et al. (2015). Functional mapping of face-selective regions in the extrastriate visual cortex of the marmoset. *Journal of Neuroscience, 35,* 1160–1172.

Julian, J. B., Fedorenko, E., Webster, J., & Kanwisher, N. (2012). An algorithmic method for functionally defining regions of interest in the ventral visual pathway. *Neuroimage, 60,* 2357–2364.

Kaiser, D., Strnad, L., Seidl, K. N., Kastner, S., & Peelen, M. V. (2014). Whole person-evoked fMRI activity patterns in human fusiform gyrus are accurately modeled by a linear combination of face- and body-evoked activity patterns. *Journal of Neurophysiology, 111,* 82–90.

Kanwisher, N., & Yovel, G. (2006). The fusiform face area: A cortical region specialized for the perception of faces. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences, 361,* 2109–2128.

Kim, N. Y., Lee, S. M., Erlendsdottir, M. C., & McCarthy, G. (2014). Discriminable spatial patterns of activation for faces and bodies in the fusiform gyrus. *Frontiers in Human Neuroscience, 8,* 632.

Konkle, T., & Oliva, A. (2012). A real-world size organization of object responses in occipitotemporal cortex. *Neuron, 74,* 1114–1124.

Ku, S.-P., Tolias, A. S., Logothetis, N. K., & Goense, J. (2011). fMRI of the face-processing network in the ventral temporal lobe of awake and anesthetized macaques. *Neuron, 70,* 352–362.

Lee, A. C. H., Buckley, M. J., Pegman, S. J., Spiers, H., Scahill, V. L., Gaffan, D., et al. (2005). Specialization in the medial temporal lobe for processing of objects and scenes. *Hippocampus, 15,* 782–797.

Lee, A. C. H., Scahill, V. L., & Graham, K. S. (2008). Activating the medial temporal lobe during oddity judgment for faces and scenes. *Cerebral Cortex, 18,* 683–696.

Lehky, S. R., & Tanaka, K. (2016). Neural representation for object recognition in inferotemporal cortex. *Current Opinion in Neurobiology, 37,* 23–35.

Liu, J., Harris, A., & Kanwisher, N. (2009). Perception of face parts and face configurations: An fMRI study. *Journal of Cognitive Neuroscience, 22,* 203–211.

Macrae, C. N., Quinn, K. A., Mason, M. F., & Quadflieg, S. (2005). Understanding others: The face and person construal. *Journal of Personality and Social Psychology, 89,* 686–695.

Mahon, B. Z., & Caramazza, A. (2011). What drives the organization of object knowledge in the brain? *Trends in Cognitive Sciences, 15,* 97–103.

McMahon, D. B. T., Russ, B. E., Elnaiem, H. D., Kurnikova, A. I., & Leopold, D. A. (2015). Single-unit activity during natural vision: Diversity, consistency, and spatial sensitivity among AF face patch neurons. *Journal of Neuroscience, 35,* 5537–5548.

Meeren, H. K. M., van Heijnsbergen, C. C. R. J., & de Gelder, B. (2005). Rapid perceptual integration of facial expression and emotional body language. *Proceedings of the National Academy of Sciences, U.S.A., 102,* 16518–16523.

Mende-Siedlecki, P., Verosky, S. C., Turk-Browne, N. B., & Todorov, A. (2013). Robust selectivity for faces in the human amygdala in the absence of expressions. *Journal of Cognitive Neuroscience, 25,* 2086–2106.

Meyers, E. M., Borzello, M., Freiwald, W. A., & Tsao, D. (2015). Intelligent information loss: The coding of facial identity, head pose, and non-face information in the macaque face patch system. *Journal of Neuroscience, 35,* 7069–7081.

Minnebusch, D. A., & Daum, I. (2009). Neuropsychological mechanisms of visual face and body perception. *Neuroscience & Biobehavioral Reviews, 33,* 1133–1144.

Moeller, S., Freiwald, W. A., & Tsao, D. Y. (2008). Patches with links: A unified system for processing faces in the macaque temporal lobe. *Science, 320,* 1355–1359.

Mur, M., Ruff, D. A., Bodurka, J., Weerd, P. D., Bandettini, P. A., & Kriegeskorte, N. (2012). Categorical, yet graded—Single-image activation profiles of human category-selective cortical regions. *Journal of Neuroscience, 32,* 8649–8662.

Murray, E. A., Bussey, T. J., & Saksida, L. M. (2007). Visual perception and memory: A new view of medial temporal lobe function in primates and rodents. *Annual Review of Neuroscience, 30,* 99–122.

Nasr, S., & Tootell, R. B. H. (2012). Role of fusiform and anterior temporal cortical areas in facial recognition. *Neuroimage, 63,* 1743–1753.

Nestor, A., Plaut, D. C., & Behrmann, M. (2011). Unraveling the distributed neural code of facial identity through spatiotemporal pattern analysis. *Proceedings of the National Academy of Sciences, U.S.A., 108,* 9998–10003.

O'Neil, E. B., Cate, A. D., & Köhler, S. (2009). Perirhinal cortex contributes to accuracy in recognition memory and perceptual discriminations. *Journal of Neuroscience, 29,* 8329–8334.

Oram, M. W., & Perrett, D. I. (1994). Responses of anterior superior temporal polysensory (STPa) neurons to "biological motion" stimuli. *Journal of Cognitive Neuroscience, 6,* 99–116.

Peelen, M. V., & Downing, P. E. (2005). Selectivity for the human body in the fusiform gyrus. *Journal of Neurophysiology, 93,* 603–608.

Peelen, M. V., & Downing, P. E. (2007). The neural basis of visual body perception. *Nature Reviews Neuroscience, 8,* 636–648.

Peelen, M. V., Wiggett, A. J., & Downing, P. E. (2006). Patterns of fMRI activity dissociate overlapping functional brain areas that respond to biological motion. *Neuron, 49,* 815–822.

Pinsk, M. A., Arcaro, M., Weiner, K. S., Kalkus, J. F., Inati, S. J., Gross, C. G., et al. (2009). Neural representations of faces and body parts in macaque and human cortex: A comparative fMRI study. *Journal of Neurophysiology, 101,* 2581–2600.

Pitcher, D., Charles, L., Devlin, J. T., Walsh, V., & Duchaine, B. (2009). Triple dissociation of faces, bodies, and objects in exastriate cortex. *Current Biology, 19,* 319–324.

Pitcher, D., Dilks, D. D., Saxe, R. R., Triantafyllou, C., & Kanwisher, N. (2011). Differential selectivity for dynamic versus static information in face-selective cortical regions. *Neuroimage, 56,* 2356–2363.

Puce, A., Allison, T., Bentin, S., Gore, J. C., & McCarthy, G. (1998). Temporal cortex activation in humans viewing eye

and mouth movements. *Journal of Neuroscience, 18,* 2188–2199.

Quiroga, R. Q., Kreiman, G., Koch, C., & Fried, I. (2008). Sparse but not "grandmother-cell" coding in the medial temporal lobe. *Trends in Cognitive Sciences, 12,* 87–91.

Rajimehr, R., Young, J. C., & Tootell, R. B. H. (2009). An anterior temporal face patch in human cortex, predicted by macaque maps. *Proceedings of the National Academy of Sciences, U.S.A., 106,* 1995–2000.

Rice, A., Phillips, P. J., Natu, V., An, X., & O'Toole, A. J. (2013). Unaware person recognition from the body when face identification fails. *Psychological Science, 24,* 2235–2243.

Schiltz, C., Dricot, L., Goebel, R., & Rossion, B. (2010). Holistic perception of individual faces in the right middle fusiform gyrus as evidenced by the composite face illusion. *Journal of Vision, 10,* 25.

Schiltz, C., & Rossion, B. (2006). Faces are represented holistically in the human occipito-temporal cortex. *Neuroimage, 32,* 1385–1394.

Schmalzl, L., Zopf, R., & Williams, M. A. (2012). From head to toe: Evidence for selective brain activation reflecting visual perception of whole individuals. *Frontiers in Human Neuroscience, 6,* 108.

Schwarzlose, R. F., Baker, C. I., & Kanwisher, N. (2005). Separate face and body selectivity on the fusiform gyrus. *Journal of Neuroscience, 25,* 11055–11059.

Song, Y., Luo, Y. L. L., Li, X., Xu, M., & Liu, J. (2013). Representation of contextually related multiple objects in the human ventral visual pathway. *Journal of Cognitive Neuroscience, 25,* 1261–1269.

Suzuki, W. A., & Naya, Y. (2014). The perirhinal cortex. *Annual Review of Neuroscience, 37,* 39–53.

Taylor, J. C., & Downing, P. E. (2011). Division of labor between lateral and ventral extrastriate representations of faces, bodies, and objects. *Journal of Cognitive Neuroscience, 23,* 4122–4137.

Taylor, J. C., Wiggett, A. J., & Downing, P. E. (2007). Functional MRI analysis of body and body part representations in the extrastriate and fusiform body areas. *Journal of Neurophysiology, 98,* 1626–1633.

Thompson, S. A., Graham, K. S., Williams, G., Patterson, K., Kapur, N., & Hodges, J. R. (2004). Dissociating person-specific from general semantic knowledge: Roles of the left and right temporal lobes. *Neuropsychologia, 42,* 359–370.

Tsao, D. Y., & Livingstone, M. S. (2008). Mechanisms of face perception. *Annual Review of Neuroscience, 31,* 411–437.

Tsao, D. Y., Moeller, S., & Freiwald, W. A. (2008). Comparing face patch systems in macaques and humans. *Proceedings of the National Academy of Sciences, U.S.A., 105,* 19514–19519.

Vul, E., Lashkari, D., Hsieh, P.-J., Golland, P., & Kanwisher, N. (2012). Data-driven functional clustering reveals dominance of face, place, and body selectivity in the ventral visual pathway. *Journal of Neurophysiology, 108,* 2306–2322.

Watson, H. C., & Lee, A. C. H. (2013). The perirhinal cortex and recognition memory interference. *Journal of Neuroscience, 33,* 4192–4200.

Weiner, K. S., & Grill-Spector, K. (2010). Sparsely-distributed organization of face and limb activations in human ventral temporal cortex. *Neuroimage, 52,* 1559–1573.

Wiggett, A. J., Pritchard, I. C., & Downing, P. E. (2009). Animate and inanimate objects in human visual cortex: Evidence for task-independent category effects. *Neuropsychologia, 47,* 3111–3117.

Yang, H., Susilo, T., & Duchaine, B. (2016). The anterior temporal face area contains invariant representations of face identity that can persist despite the loss of right FFA and OFA. *Cerebral Cortex, 26,* 1096–1107.

Yovel, G., & Kanwisher, N. (2005). The neural basis of the behavioral face-inversion effect. *Current Biology, 15,* 2256–2262.