NeuroImage 202 (2019) 116085

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

Decoding subcategories of human bodies from both body- and face-responsive cortical regions

Celia Foster [a,b,c,*], Mintao Zhao [a,d], Javier Romero [b], Michael J. Black [b], Betty J. Mohler [a,b], Andreas Bartels [a,c,e,f], Isabelle Bülthoff [a,**]

[a] Max Planck Institute for Biological Cybernetics, Tübingen, Germany
[b] Max Planck Institute for Intelligent Systems, Tübingen, Germany
[c] Centre for Integrative Neuroscience, Tübingen, Germany
[d] School of Psychology, University of East Anglia, UK
[e] Department of Psychology, University of Tübingen, Germany
[f] Bernstein Center for Computational Neuroscience, Tübingen, Germany

ARTICLE INFO

Keywords:
Body perception
Face perception
EBA
FBA
OFA
FFA

ABSTRACT

Our visual system can easily categorize objects (e.g. faces vs. bodies) and further differentiate them into subcategories (e.g. male vs. female). This ability is particularly important for objects of social significance, such as human faces and bodies. While many studies have demonstrated category selectivity to faces and bodies in the brain, how subcategories of faces and bodies are represented remains unclear. Here, we investigated how the brain encodes two prominent subcategories shared by both faces and bodies, sex and weight, and whether neural responses to these subcategories rely on low-level visual, high-level visual or semantic similarity. We recorded brain activity with fMRI while participants viewed faces and bodies that varied in sex, weight, and image size. The results showed that the sex of bodies can be decoded from both body- and face-responsive brain areas, with the former exhibiting more consistent size-invariant decoding than the latter. Body weight could also be decoded in face-responsive areas and in distributed body-responsive areas, and this decoding was also invariant to image size. The weight of faces could be decoded from the fusiform body area (FBA), and weight could be decoded across face and body stimuli in the extrastriate body area (EBA) and a distributed body-responsive area. The sex of well-controlled faces (e.g. excluding hairstyles) could not be decoded from face- or body-responsive regions. These results demonstrate that both face- and body-responsive brain regions encode information that can distinguish the sex and weight of bodies. Moreover, the neural patterns corresponding to sex and weight were invariant to image size and could sometimes generalize across face and body stimuli, suggesting that such subcategorical information is encoded with a high-level visual or semantic code.

1. Introduction

Our visual system makes use of various aspects of shape information to categorize objects (e.g. person vs. house) and to further categorize them into different subcategories (e.g. male vs. female person). This seemingly effortless ability is actually remarkably non-trivial, as objects that fit one subcategory can be of a great variability (e.g. both faces and bodies can belong to the same subcategory male), yet exemplars from different subcategories can look comparably similar (e.g. male vs. female faces). In the brain, both monkey neurophysiology and human neuroimaging studies indicate that object categorization and subcategorization processes are primarily implemented in the ventral temporal cortex (VTC) (Grill-Spector and Weiner, 2014; Gross et al., 1972; Haxby et al., 2001; Kriegeskorte et al., 2008; Logothetis and Sheinberg, 1996; Tanaka, 1996). While the VTC contains high-level category-selective areas for objects (Malach et al., 1995), faces (Gauthier et al., 2000; Kanwisher et al., 1997), bodies (Downing et al., 2001; Peelen and Downing, 2005), scenes (Epstein and Kanwisher, 1998), and visually presented words (Cohen et al., 2000), how, and where in the brain, subcategories are encoded remains to be elucidated. Even less is known about how the brain represents the same semantic categories that are shared by different object categories (e.g. sex of faces and sex of bodies).

* Corresponding author. Max Planck Institute for Biological Cybernetics, Max-Planck-Ring 8, 72076, Tübingen, Germany.
** Corresponding author.
E-mail addresses: celia.foster@tuebingen.mpg.de (C. Foster), isabelle.buelthoff@tuebingen.mpg.de (I. Bülthoff).

https://doi.org/10.1016/j.neuroimage.2019.116085
Received 20 July 2018; Received in revised form 17 July 2019; Accepted 7 August 2019
Available online 8 August 2019

Recent studies on face perception suggest that face-responsive brain areas may contain neural representations of subcategories (e.g. the sex, race, and identity of faces). Different patterns of neural responses to male and female faces have been identified in the fusiform face area (FFA) and other face-responsive regions (Contreras et al., 2013; Freeman et al., 2010; Kaul et al., 2011). Different patterns of neural responses to faces of different races have also been identified in the fusiform gyrus and early visual cortex (Contreras et al., 2013; Ratner et al., 2013). Several studies have investigated how the brain represents face identities (an extreme level of subcategorization) and found stimulus-independent representation of face identities in the anterior temporal face area (ATFA) (Anzellotti et al., 2014; Guntupalli et al., 2016; Kriegeskorte et al., 2007). Other studies found that the FFA (Anzellotti et al., 2014; Axelrod and Yovel, 2015) and the superior intraparietal sulcus (Jeong and Xu, 2016) also encode face identity.

It remains unclear exactly which subcategory features drive distinctive patterns of neural responses to different face subcategories, and whether these subcategories are represented in brain areas beyond those selective for faces (Haxby et al., 2001). The distinction between visual and semantic representation has been observed during general object categorization (Bracci and Op de Beeck, 2016), which may similarly apply to face and body categorization. For instance, different patterns of neural responses to male versus female faces could be driven by differential sensitivity to the *visual feature* of hairstyle, rather than the perceived *semantic category* of biological sex. Similarly, separable neural responses to faces of different races may be caused by the visual feature of skin tone, rather than the semantic category of race. Studies finding different neural responses to bodies of different subcategories in body-responsive areas may have the same visual/semantic concern. There is some evidence that the extrastriate body area (EBA) and fusiform body area (FBA) contain information about the identity of bodies (Ewbank et al., 2011) and the FBA and right middle occipital gyrus contain information about the weight of bodies (Hummel et al., 2013). However, different neural responses to lower vs. higher weight bodies might be supported by different sensitivity to the physical image size of bodies rather than the semantic perception of body weight. In behaviour, perceived body weight has been found to be processed independently of physical image size (Sturman et al., 2017).

In this study, we investigated how information about subcategories shared by faces and bodies is encoded in the face- and body-responsive brain networks. We chose the same two subcategories of faces and bodies: sex (male vs. female) and weight (lower vs. higher). We presented participants with images of faces and bodies varying in sex, weight, and image size (larger vs. smaller) whilst recording their brain activity using functional magnetic resonance imaging (fMRI). Inclusion of both face and body stimuli allowed us to investigate whether the brain encodes shared semantic subcategories (e.g. male/female) in an abstract manner, despite dramatic differences in the visual appearance of stimuli (e.g. faces vs. bodies). We varied image size of stimuli to further test whether the neural coding of subcategories is more abstract (i.e. image size invariant) or more bound to visual features (i.e. image size dependent). Varying image size also helps to clarify whether or not neural processes in high-level visual areas are modulated by low-level visual features (e.g. image size) (Andrews and Ewbank, 2004; Sawamura et al., 2005; Yue et al., 2011).

To investigate how face- and body-responsive brain regions encode the subcategories of sex and weight, we first trained support vector machine (SVM) classifiers to discriminate patterns of blood-oxygen-level dependent (BOLD) activity elicited by each subcategory, and then used them to predict the subcategories within separate test data. We further tested if neural responses to different subcategories generalize across stimuli sizes (e.g. trained with smaller faces, tested on larger faces), and across face and body stimuli (e.g. trained with faces, tested on bodies). We performed these multivoxel pattern analyses (MVPA) for both functionally defined face- and body-responsive areas and in whole-brain searchlight analyses. These analyses allow us to differentiate whether the neural coding of a subcategory is driven by low-level visual features, by high-level visual features, or by semantic processing. Specifically, if neural responses to face and body subcategories were driven by low-level visual features, they would be dependent on image size. If the neural responses to these subcategories were driven by high-level visual features, they would be independent of the image size, but not necessarily able to generalize across face and body stimuli. In contrast, if neural responses were driven by semantic information then they would be able to generalize across image size and across face and body stimuli.

2. Materials and methods

2.1. Participants

Thirteen participants (7 female, 6 male, 22–32 years old) were included in the fMRI experiment analyses presented in the results section, out of fifteen who had originally participated. Due to scanner malfunction the face-stimuli runs could not be completed in one participant. This participant was included in the body-related analyses but not the face ones. Data from two participants were excluded from analyses due to excessive head movement during scanning. All participants provided written informed consent prior to the experiment, and the procedures were approved by the ethics committee of the University Clinic Tübingen.

2.2. Stimuli

The experimental stimuli were grayscale images of faces and bodies that varied in biological sex (i.e. male or female) and weight (i.e. higher weight or lower weight), resulting in four stimulus classes for faces, and four for bodies. Each class was presented in both a smaller and larger image size (larger images were twice the height and width of smaller images) resulting in a total of 16 conditions (Fig. 1A). Each class contained 42 exemplars (e.g. low-weight, male bodies), all of which were different individuals. The perceived sex and weight of these images were validated via ratings from independent observers (see sections 2.2.3 and 3.1.1 below).

2.2.1. Face stimuli

Face stimuli were created using 3D face models of the face database of the Max Planck Institute for Biological Cybernetics (Blanz and Vetter, 1999; Troje and Bülthoff, 1996). To increase the variability of weight of faces, we intensified the perceived weight of the faces using a 3D morphable model (Blanz and Vetter, 1999), using the following procedure. Six observers (5 female, 1 male, 22–27 years old, 1 author), who did not participate in the fMRI experiment, rated the perceived weight of 311 faces from the database. Based on these ratings we selected all faces that were above or below one standard deviation from the average perceived weight, separately for male and female faces. These selected faces were then morphed together to generate average higher and lower weight faces separately for male and female faces. The difference between higher- and lower-weight average morphs was then applied to the original individual face models, so that for each face we were able to create higher and lower weight versions of that face. Any stimuli with artefacts from the morphing procedure were removed.

2.2.2. Body stimuli

Body stimuli were created based on body scans from the CAESAR dataset (Robinette et al., 2002). We selected bodies with a Body Mass Index (BMI) between one and two standard deviations above the average BMI for our higher weight bodies, and between one and two standard deviations below the average BMI for our lower weight bodies (both calculated separately for male and female bodies). The selected 264 body scans were then registered to a 3D body shape and pose model (Loper et al., 2015). This allowed us to obtain individual body shapes in a standard A-pose (see Fig. 1A). We modified the texture obtained from the

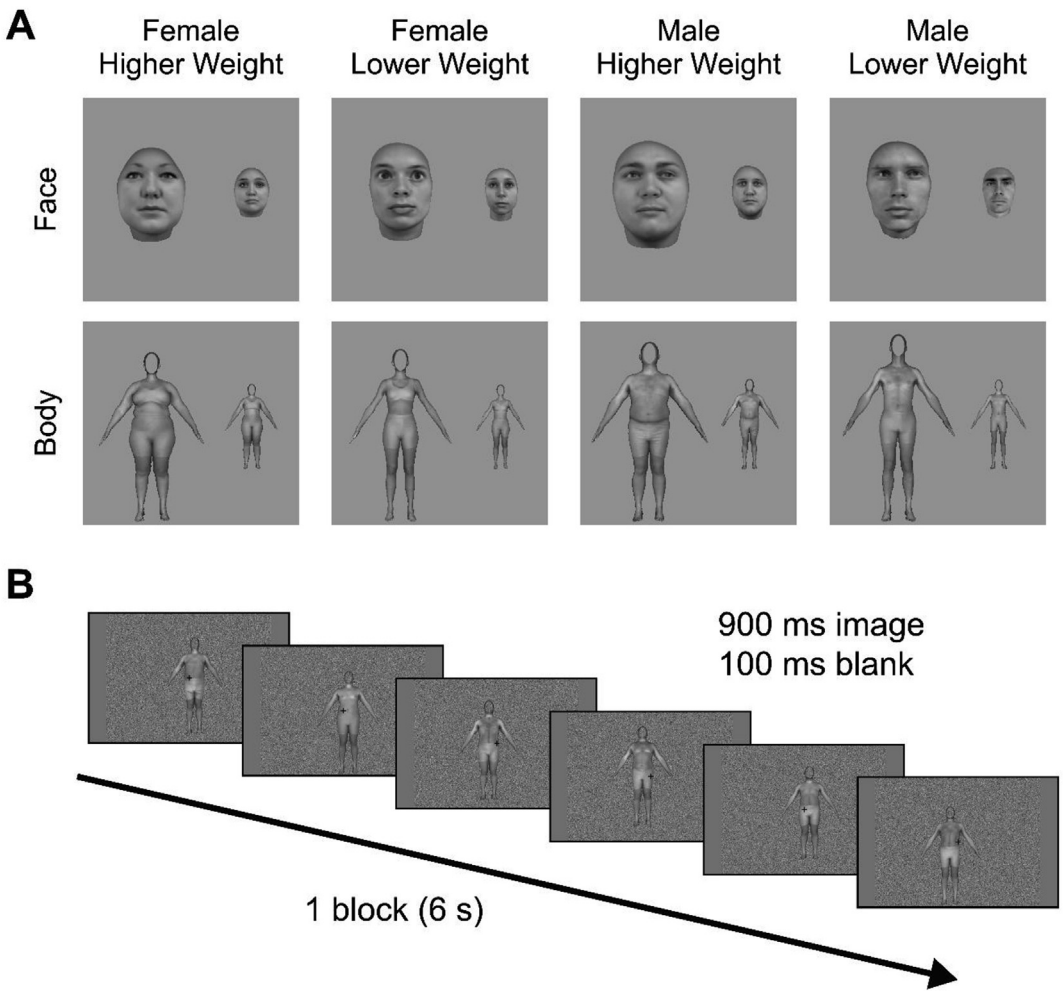

Fig. 1. Experimental stimuli and procedure of the fMRI experiment. (A) Example stimuli for the 16 conditions of the fMRI experiment. Stimuli were shown at a larger and a smaller image size (larger images were twice the height and width of smaller images). (B) An example block of stimuli in the fMRI experiment. Subjects viewed the stimuli in 6 s blocks, where each block contained images from one condition. Each image was presented for 900 m s with a 100 m s blank grayscale screen between images. Two Gaussian noise only images were shown between blocks, each lasted 900 m s followed by a 100 m s blank screen.

original scans in order to remove markers and fill in any missing texture. In the final body images, the faces were covered with an oval in order to exclude any face information from the body images.

2.2.3. Ratings of face and body stimuli

To select stimuli that truly differed in perceived weight and sex, we had six observers (3 female, 3 male, 22–35 years old), who were not participants in the fMRI experiment, rate the perceived sex and weight, for both face and body stimuli, on 7-point Likert scales. Based on these ratings we then selected 42 stimuli for each of the eight conditions (e.g. low-weight, male bodies) that maximised perceived difference in sex and weight.

2.2.4. Background stimuli

During the fMRI experiment, face and body images were shown in front of a randomly generated Gaussian noise background (Fig. 1B), in order to keep the area of retinal stimulation constant for all stimuli despite differences in the foreground image shape.

2.2.5. Localizer stimuli

Stimuli for the localizer experiment consisted of grayscale images of faces, headless bodies, objects and phase-scrambled images. Phase-scrambled images were created by making a collage containing the face and headless body images and then generating Fourier-scrambled images from the collage image.

2.3. fMRI experiment

The study consisted of two fMRI sessions on separate days. On the first day localizer runs and anatomical data were collected. On the second day experimental runs were collected. Stimuli were presented via a projector (resolution 1920×1080) with Matlab 2013b using the Psychophysics Toolbox extensions (Brainard, 1997; Kleiner et al., 2007) on a Windows PC. Participants lay supine in the scanner and viewed a screen positioned behind their head via a mirror attached to the head coil. The screen was positioned at a distance of 82 cm, and spanned $28^\circ \times 16^\circ$ of visual angle in horizontal and vertical directions respectively.

2.3.1. Experimental runs

Participants completed 8 experimental runs where 4 runs contained face stimuli and 4 runs contained body stimuli. Each run contained 8 conditions of a 2 (Sex: male vs. female) x 2 (Weight: lower vs. higher) x 2 (Image Size: larger vs. smaller) factorial design (Fig. 1A). Conditions were presented in a carryover counterbalanced blocked design, such that each condition block was preceded by each condition block once in a run (Brooks, 2012). This was to avoid biases due to remaining BOLD activation from a previous condition block (Aguirre, 2007). Stimuli were presented in front of a centred Gaussian noise background (width 5.1°,

height 7.9°). Average stimulus sizes were of the following widths and heights; larger faces 2.4° × 3.7°; smaller faces 1.2° × 1.8°; larger bodies 4.0° × 6.5°; smaller bodies 2.0° × 3.3°. In each block 6 images were shown, where each image was shown for 0.9 s followed by a 0.1 s blank grayscale screen. Two Gaussian noise only images were shown between blocks, each for 0.9 s followed by a 0.1 s blank grayscale screen. Images in each block appeared in 6 different positions (horizontally either 0.2° left or right of the screen centre; vertically either centred, or 0.2° above or below the screen centre) in a random order.

2.3.1.1. fMRI task. Participants fixated a central fixation cross at all times and pressed a button whenever a red dot appeared on the fixation cross. This occurred twice during each block at randomly determined times at least 2 s apart.

2.3.2. Localizer runs

Participants completed 5 localizer runs, which were collected on a separate day prior to the experimental runs. In each run participants viewed blocks of faces, bodies, objects and phase-scrambled images. Faces, bodies and objects were shown in front of the phase-scrambled images to match the amount of retinal stimulation in all blocks. In each block 8 images were shown, with each image being presented for 1.8 s, and a 0.2 s blank grayscale screen between images. Images were presented in a carryover counterbalanced sequence (Brooks, 2012). Participants performed a one-back matching task on the images, to ensure balanced attention across conditions. Image repetitions occurred on average once every 9 s.

2.4. fMRI scan parameters

Images were acquired using a 3T Siemens Prisma scanner with a 64-channel head coil (Siemens, Erlangen, Germany). Functional T2* echo-planar images (EPI) were acquired using a sequence with the following parameters; multiband acceleration factor 2, TR 1.2 s, TE 30 m s, flip angle 68°, FOV 192 × 192 mm. Volumes consisted of 36 slices, with an isotropic voxel size of 3 × 3 × 3 mm. The first 8 vol of each run were discarded to allow for equilibration of the T1 signal. During each session a gradient echo field map was recorded so that magnetic field inhomogeneity could be corrected during preprocessing. For each participant a high-resolution T1-weighted anatomical scan was acquired with the following parameters; TR 2 s, TE 3.06 m s, FOV 232 × 256 mm, 192 slices, isotropic voxel size of 1 × 1 × 1 mm.

2.5. fMRI data preprocessing

Data was preprocessed using SPM12 (http://www.fil.ion.ucl.ac.uk/spm/). All functional data was realigned, unwarped to correct for inhomogenities in the magnetic field and coregistered to the anatomical data. Localizer data was spatially smoothed with a 6 mm Gaussian kernel. For the whole-brain univariate analyses the data was normalized to MNI (Montreal Neurological Institute) space and spatially smoothed with a 6 mm Gaussian kernel. Multivariate data analyses were performed in individual subject space on unsmoothed data. For searchlight analyses the resulting single-subject maps of classification accuracies were normalized to MNI space, and spatially smoothed with a 6 mm Gaussian kernel.

2.6. Definition of regions of interest

We defined separate face- and body-responsive regions of interest (ROIs) using data from the localizer runs. The contrast faces > objects was used to identify the OFA and FFA and the contrast bodies > objects was used to identify the EBA and FBA (Table 1). As ROI size has been shown to affect decoding accuracy (Gardumi et al., 2016) we kept ROI sizes constant by selecting the 100 most active voxels in each region bilaterally to form the ROI. To achieve this, we initially attempted to identify each ROI using a threshold of $p < 0.05$ (FWE corrected). If the ROI was not identifiable in 100 voxels we attempted to define the ROI using a lower threshold of $t = 2$. ROIs were defined using localizer data only, therefore their definition was independent to the main experiment analyses. The FFA and the FBA are known to partially overlap (Schwarzlose et al., 2005). We initially performed analyses in both ROIs without removing any overlapping voxels (this overlap had a mean of 32% of FFA voxels and 38% of FBA voxels). In any instances where an analysis was significant for both FFA and FBA (and therefore significant decoding could be caused be overlapping voxels) we planned to run additional analyses removing the overlapping voxels to investigate this possibility.

In addition to the separate ROIs we defined distributed face- and body-responsive ROIs that contained voxels from several of the isolated ROIs, as described previously (Hahn and O'Toole, 2017). This allowed us to investigate whether neural information from distributed brain regions improves categorical classification of faces and bodies. Distributed face-responsive ROIs were defined using voxels from OFA, FFA, STS and ATFA. Distributed body-responsive ROIs were defined using voxels from EBA and FBA. The 300 and 500 most responsive voxels were selected to create two sizes of distributed ROIs. We defined 300 voxel ROIs as every participant had at least 300 face- and body-responsive voxels, thus we were able to define a distributed face- and body-responsive ROI of this size in every participant. Most participants also had many more face-and body-responsive voxels, thus we additionally defined 500 voxel face- and body-responsive ROIs (in N = 10 and N = 11 participants respectively) to see if classifier performance would benefit from information in these additional voxels.

We used V1 as a control ROI, which was bilaterally localized using anatomical labels generated using the Freesurfer software package (Hinds et al., 2009) (https://surfer.nmr.mgh.harvard.edu/). This method generates V1 ROIs based on the anatomy of the participant, and the method has been validated to show that there is close agreement between these anatomical V1 ROIs and V1 defined functionally using retinotopic mapping. For each participant, we selected the 50 most posterior voxels (corresponding to the foveal section of V1) from each hemisphere and combined them to create a 100-voxel V1 ROI.

Table 1
ROI coordinates (in MNI space), ROI volume and number of subjects each ROI was identified in (N). All ROI analyses were done in subject space. ROIs were subsequently normalized to MNI space in order to show the mean ROI locations. Coordinates show mean x, y and z locations and volume, ± standard deviations.

ROI	hemisphere	x	y	z	Volume (mm^3)	N
OFA	left	−34 ± 6.4	−83 ± 7.1	−12 ± 3.4	520 ± 296.9	13
	right	37 ± 5.9	−80 ± 7.4	−13 ± 3.5	945 ± 356.4	13
FFA	left	−41 ± 2.8	−56 ± 8.8	−19 ± 3.7	519 ± 253.4	12
	right	43 ± 3.1	−53 ± 6.0	−18 ± 2.6	979 ± 274.5	13
EBA	left	−47 ± 2.8	−77 ± 3.7	4 ± 5.0	385 ± 153.0	12
	right	48 ± 2.6	−70 ± 3.3	−1 ± 5.4	1106 ± 252.1	13
FBA	left	−42 ± 2.7	−48 ± 5.8	−21 ± 5.0	374 ± 277.2	11
	right	44 ± 2.9	−49 ± 4.0	−18 ± 3.5	994 ± 272.2	11
V1	left	−9 ± 1.4	−100 ± 1.8	−3 ± 4.3	390 ± 70.8	13
	right	12 ± 3.9	−98 ± 1.8	0 ± 3.9	468 ± 80.2	13

To investigate if our choice to fix the ROI size at a constant number of voxels, rather than fixing ROI size at a constant t-contrast threshold had any impact on our results, we additionally conducted supplemental analyses where ROIs were defined using this alternative method (Supplementary Table S1). Specifically, we defined supplemental versions of our face- and body-responsive ROIs by selecting all active voxels at a constant threshold of $t = 3$ rather than a constant number of voxels. A supplemental V1 ROI was defined using the same method as the original V1 ROI, except that the 510 most posterior voxels were selected in order to match the size of this V1 ROI to the average size of the largest separate ROI (EBA). All results of these supplemental analyses can be found in the supplemental materials.

2.7. Multivoxel pattern analyses (MVPA)

Data from the main experiment was modelled with General Linear Models (GLMs) using SPM12. The responses to each block (i.e. trial) were modelled as separate regressors in the GLM. Classification analyses were performed using The Decoding Toolbox (Hebart et al., 2015) from the resulting beta weight images. Input data was feature-scaled using z-score normalization and outlier values (greater than 2 standard deviations) were set to 2 or -2. Mean and standard deviation for z-scoring were estimated using training data and then these values were applied to testing data to ensure independence of training and testing data. Classification was performed using a linear support vector machine classifier (LIBSVM).

We performed three different MVPA analyses (one combining both image sizes, one across image size, and one across face and body stimuli), and each was done for separate ROIs, distributed ROIs and in whole-brain searchlight analyses. Statistical significance for ROI analyses was assessed using permutation tests with the following procedure. For each ROI classification analysis the entire analysis was repeated 10,000 times with condition labels randomly assigned. Thus, we generated a null distribution of classification accuracies expected by chance and specific to each analysis. We assessed the significance of our ROI classification results by testing how often in the null distribution we obtained a mean decoding performance greater than or equal to the actual mean decoding performance found in our analysis. We first tested whether this was significant at the $p < 0.05$ threshold using a one-sided test (mean decoding performance in the null distribution must be greater than or equal to the actual mean decoding performance in less than 500 out of 10,000 tests). We then additionally applied a Bonferroni-correction to correct *p*-values for the number of ROIs tested. *P*-values following Bonferroni correction were limited to a maximum value of one (i.e. if Bonferroni correction caused a *p*-value to be greater than one, we set its value to one).

We additionally performed whole-brain searchlight analyses to investigate if any regions outside of our defined ROIs would be able to decode sex or weight. Searchlight analyses involved performing each classification analysis in 4-voxel radius spheres, where each sphere was centred around each voxel in the brain once, thus producing whole-brain maps of classification accuracies. To assess the significance of searchlight data we performed one-sided *t*-tests in SPM12 using a False Discovery Rate (FDR) correction to adjust for multiple comparisons.

2.7.1. MVPA analysis 1: classification of sex and weight

Our first set of analyses aimed to determine which brain regions encode the subcategories sex and weight, separately based on neural activity induced by viewing faces and bodies of these subcategories. For classification of sex, the classifier was trained and tested on male versus female stimuli, regardless of image size and weight. For classification of weight, the classifier was trained and tested on higher weight versus lower weight stimuli, regardless of image size and sex. Thus, we maximised the amount of data for the classifier to make its decision, and pooled across irrelevant subcategories. This meant also that the classifier was trained to be invariant to one of the features when distinguishing the other. We used 3 out of 4 runs as training data for the SVM classifier. We then used the trained classifier to predict the subcategories of the blocks in the independent test data from the 4th withheld run. A four-fold cross-validation procedure was used, such that each run was used once as the held out test dataset. Final decoding accuracy was determined by averaging over the four cross-validation iterations.

2.7.2. MVPA analysis 2: size-invariant classification of sex and weight

Our second set of analyses aimed to determine which brain regions encode the subcategories sex and weight in an image-size invariant manner. To this end, we trained the classifier to decode the subcategory from neural data evoked by one image size and then tested its ability to decode this subcategory from neural activity evoked by the other image size. As previously, 3 out of 4 runs were used as training data for the SVM classifier and the 4th withheld run was used as test data. However, in this case the training data only used neural activity data evoked by one image size, whereas the test data only used neural activity data evoked by subjects viewing the other image size. A four-fold cross-validation procedure was again used, such that each run was used once as the held out test dataset. This was repeated two times, once using the smaller image size data as training data and the larger image size data as test data, and vice-versa. The final decoding accuracy was determined by averaging over the four cross-validation iterations and both image size training and test set combinations.

2.7.3. MVPA analysis 3: classification of sex and weight across face and body stimuli

Our third set of analyses investigated which brain regions encode the subcategories sex and weight across face and body stimuli. To this end, we trained classifiers to decode each subcategory from neural data evoked by faces and tested them on neural activity evoked by bodies, and vice-versa. As face and body stimuli were shown in separate runs (4 runs each) we trained classifiers on all 4 runs of data evoked by one category (i.e. face or body stimuli) and tested them on all 4 runs of data evoked by the other category. The final decoding accuracy was determined by averaging over the two training and test set combinations.

2.8. Univariate analyses

We also performed univariate analyses to identify brain regions that are sensitive to the sex, weight, and image size of faces and bodies. In both separate and distributed ROIs we used *t*-tests to look for significant differences in the average BOLD response. *P*-values were Bonferroni-corrected for the number of ROIs tested. In addition, whole-brain analyses were conducted to look for additional brain regions with univariate activation differences, using a FDR correction to adjust for multiple comparisons.

2.9. Data and code availability statement

Data cannot be shared as participants were informed that their data would be stored confidentially, in accordance with the rules of the local ethics committee. Code is available on request.

3. Results

3.1. Behavioural results

3.1.1. Perceptual differences of face and body stimuli

Our selected face and body stimuli were perceptually different in terms of perceived sex and weight (Fig. 2). Participants gave significantly different ratings of sex to male and female faces and bodies (Fig. 2A; faces: $t_5 = 15.2$, $p = 2.3 \times 10^{-5}$; bodies: $t_5 = 20.0$, $p = 5.8 \times 10^{-6}$) and gave a significantly higher weight score for higher-weight stimuli than lower-weight stimuli (Fig. 2B; faces: $t_5 = 7.1$, $p = 8.8 \times 10^{-4}$; bodies: $t_5 = 9.3$, $p = 2.5 \times 10^{-4}$). There were also differences in the rating

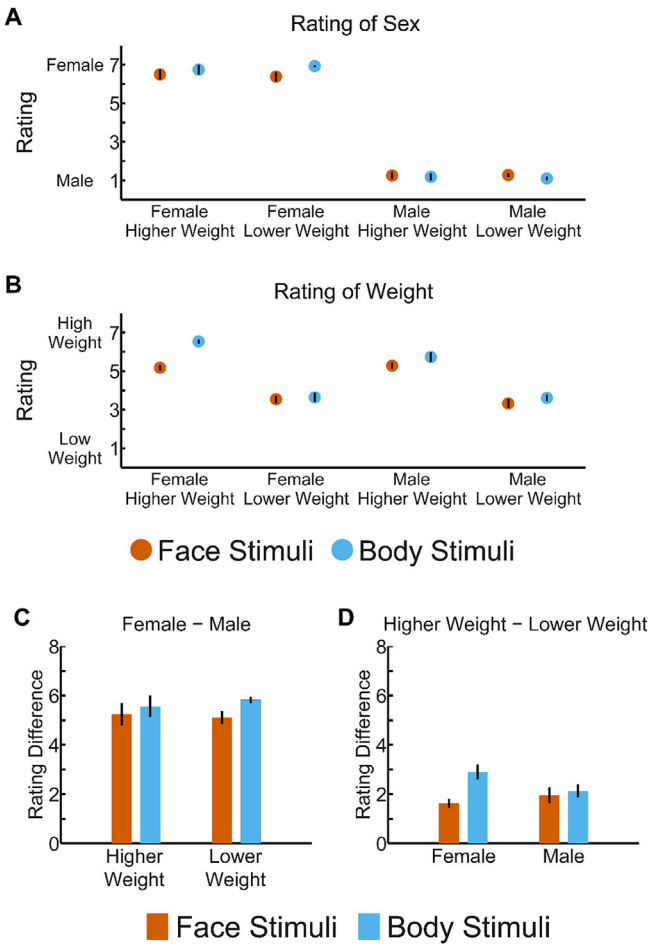

Fig. 2. Perceptual ratings of stimuli in an independent behavioural experiment. Participants rated how they perceived the biological sex (A) and weight (B) of face and body stimuli on 7-point scales. Circles indicate the mean rating scores and error bars show the standard error of the mean. (C) and (D) show the difference between the mean ratings of sex for the male and female stimuli (C), and between the mean ratings of weight for the higher and lower weight stimuli (D). Error bars show the standard error of the mean.

between face and body stimuli. The difference between male and female ratings was greater for body stimuli as compared to face stimuli ($t_5 = 5.3$, $p = 0.0031$; Fig. 2C) and the difference between higher and lower weight ratings was greater for body stimuli as compared to face stimuli ($t_5 = 2.8$, $p = 0.04$; Fig. 2D). Note that during the fMRI experiment, participants performed a simple task on the fixation cross that was unrelated to the stimuli.

3.1.2. Attention task during scanning

Participants showed high performance on the attention task during scanning (mean detection rate of 92% across all conditions). Detection performance showed no difference between face and body conditions ($t_{11} = 0.42$, $p = 0.68$), between male and female conditions ($t_{11} = -0.34$, $p = 0.74$), between higher and lower weight conditions ($t_{11} = 0.02$, $p = 0.98$) or between larger and smaller image size conditions ($t_{11} = -0.33$, $p = 0.74$). There was also no difference in reaction times between face and body conditions ($t_{11} = -0.36$, $p = 0.71$), between male and female conditions ($t_{11} = 0.06$, $p = 0.95$), between higher and lower weight conditions ($t_{11} = -0.03$, $p = 0.98$) or between larger and smaller image size conditions ($t_{11} = -0.21$, $p = 0.84$). Therefore, any observed decoding difference between experimental conditions (i.e. faces or bodies, sex, weight or image size) cannot be attributed to different levels of attention during scanning.

3.2. Classification of sex and weight

We first tested which brain regions show separable patterns of neural activity evoked by male and female stimuli or by higher- and lower-weight stimuli, regardless of image size. We used a leave-one-run-out cross-validation method, so that for each iteration we trained a linear support vector machine (SVM) classifier using data from three fMRI runs and tested the classifier with the data from the one remaining run. We performed these classification analyses in three types of brain regions; (1) separate face- and body-responsive regions, (2) distributed face- and body-responsive regions, and (3) in whole-brain searchlight analyses. The results are shown in Figs. 3 and 4.

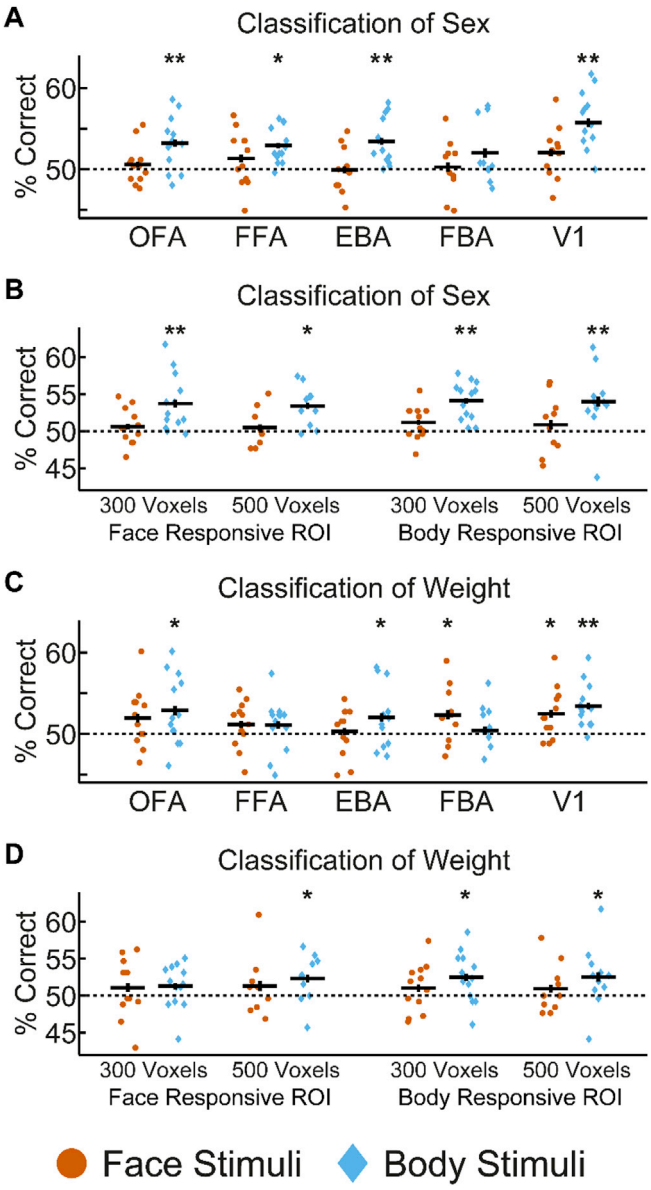

Fig. 3. Classification of sex and weight. Classification results for sex in separate ROIs (A) and distributed ROIs (B). Classification results for weight in separate ROIs (C) and distributed ROIs (D). Scatter plots show decoding accuracy for individual participants, horizontal black lines show group mean decoding accuracies and vertical error bars show the standard error of the mean. * indicates $p < 0.05$, ** indicates $p < 0.001$, Bonferroni-corrected for the number of ROIs (separate ROIs: N = 5; distributed ROIs: N = 4). Dotted lines indicate chance-level decoding performance, 50%.

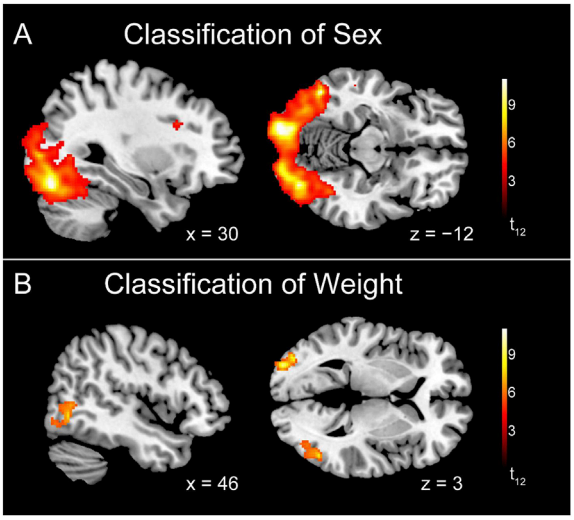

Fig. 4. Results of searchlight analyses for the classification of the sex (A) and weight (B) of bodies. (A) Regions able to classify the sex of bodies in the searchlight analysis overlapped with the mean coordinates of the rEBA, OFA, FFA, rFBA and V1. (B) Regions able to classify the weight of bodies in the searchlight analysis overlapped with the mean coordinates of the lOFA and rEBA.

3.2.1. Classification using separate ROIs

For body stimuli, classification of sex (Fig. 3A) was significantly above chance in EBA (53.4%, $p < 0.0005$), but not FBA (52.0%, $p = 0.078$). In addition, face-responsive areas OFA (53.2%, $p = 0.001$) and FFA (52.9%, $p = 0.0025$), as well as V1 (55.7%, $p < 0.0005$), also showed higher-than-chance classification of the sex of bodies. Classification of body weight (Fig. 3C) was significantly above chance in the EBA (52.0%, $p = 0.044$), but not the FBA (50.4%, $p = 1$). Again, we found higher-than chance level classification of body weight in OFA (52.9%, $p = 0.0015$) and V1 (53.4%, $p < 0.0005$). These results indicate that the encoding of body subcategories is not unique to body-responsive regions as both body- and face-responsive areas encode information about the sex and weight of bodies. Higher-than-chance decoding performance observed in V1 suggests that low-level visual features could be used to distinguish body subcategories.

For face stimuli, none of the separate ROIs was able to classify the sex of faces at a higher-than-chance level (Fig. 3A): OFA (50.6%, $p = 1$), FFA (51.3%, $p = 0.36$), EBA (49.9%, $p = 1$), FBA (50.3%, $p = 1$), V1 (52.1%, $p = 0.073$). Classification of the weight of faces was above chance in FBA (52.3%, $p = 0.046$) and V1 (52.5%, $p = 0.017$), but not in OFA (52.0%, $p = 0.072$), FFA (51.2%, $p = 0.48$), or EBA (50.3%, $p = 1$).

To examine whether our results were influenced by the way we defined ROIs (i.e. using a fixed number of voxels), we performed additional analyses using ROIs defined at a constant t-contrast threshold rather than a constant number of voxels. These additional analyses revealed the same results as reported above (see Supplementary Fig. S1), with the exception that the classification of the weight of faces was not significant in the FBA (50.9%, $p = 0.80$).

3.2.2. Classification using distributed ROIs

We further investigated classification of sex and weight in larger, distributed face- and body-responsive ROIs. These ROIs may show improved classification performance compared to smaller separated ROIs, due to increased information available for classification and known effects of ROI size on classification accuracy (Cox and Savoy, 2003; Gardumi et al., 2016). We defined ROIs of two different sizes: 300 voxels and 500 voxels (see Section 2.6 for details). For body stimuli, classification of sex (Fig. 3B) was higher than chance in both sizes of body-responsive ROI (300 voxels: 54.1%, $p < 0.0004$; 500 voxels: 54.0%, $p < 0.0004$), as well as both sizes of face-responsive ROI (300 voxels: 53.7%, $p < 0.0004$; 500 voxels: 53.4%, $p = 0.0024$). Classification of body weight (Fig. 3D) was also higher than chance in both sizes of body-responsive ROI (300 voxels: 52.5%, $p = 0.012$; 500 voxels: 52.5%, $p = 0.015$), and in the larger face-responsive ROI (500 voxels: 52.3%, $p = 0.03$).

For face stimuli, no distributed face or body ROIs were able to decode the sex or weight of faces, even using the large 500-voxel ROIs (which had the most information available for classification). Specifically, for the sex of faces the decoding performance was 50.5% ($p = 1$) in the larger face-responsive ROI and 50.9% ($p = 0.75$) in the larger body-responsive ROI. For the weight of faces the decoding performance was 51.3% ($p = 0.40$) in the larger face-responsive ROI and 50.9% ($p = 0.68$) in the larger body-responsive ROI. It is worth noting that when the distributed face-responsive ROI was defined at a constant t-contrast threshold rather than constant number of voxels, we were able to decode the weight of faces (52.2%, $p = 0.015$; see Supplementary Fig. S1D).

3.2.3. Classification using whole-brain searchlight analysis

To investigate if any other brain regions were sensitive to sex or weight information of faces and bodies, we performed the same classification analyses at the whole-brain level using a searchlight analysis. Fig. 4 shows the brain regions identified in the searchlight analyses with body stimuli. Higher-than-chance decoding of the sex of body stimuli was observed in much of occipitotemporal cortex (Fig. 4A) and consistent with the ROI classification analysis, the regions sensitive to the sex of bodies overlapped with the peak coordinates of body-responsive rEBA and rFBA, as well as face-responsive OFA and FFA and V1. Additional clusters sensitive to the sex of bodies were revealed in parietal (MNI coordinates: −2, −66, 56) and frontal regions (MNI coordinates: 40, 60, 6; 20, 32, 42; 28, 16, 28). These results show that many regions contain low-level visual, high-level visual or semantic information that allows a classifier to distinguish between neural activity evoked by bodies of different sexes.

Classification of the weight of body stimuli was significantly above chance in two regions that overlapped with the mean coordinates of lOFA and rEBA, as well as a small cluster in the right anterior temporal lobe (MNI coordinates: 40, −6, 32). For the face stimuli, classification of weight revealed a cluster in the left early visual cortex (MNI coordinates: −14, −94, −12). No regions were identified that could decode the sex of face stimuli.

3.3. Size-invariant classification of sex and weight

To investigate whether neural responses to sex and weight were invariant to image size, we tested whether patterns of neural activity evoked by one size of stimuli could be generalized to decode the sex or weight when participants view stimuli of a different size. Thus, we trained SVM classifiers to distinguish sex and weight using data obtained with only one of the two image sizes. We then tested the classifier using only the data obtained with the other image size. Significant decoding would suggest that the pattern of neural response to sex or weight was invariant to image size. Again, we performed the size invariant classification analysis using separate ROIs, distributed ROIs, and in searchlight analyses across the whole brain. The results are shown in Fig. 5.

For classification analysis using the separate ROIs, both body-responsive ROIs showed higher-than-chance performance when decoding the sex of body stimuli (EBA, 52.7%, $p = 0.002$; FBA, 52.8%, $p = 0.003$), but face-responsive ROIs OFA (50.9%, $p = 0.67$) and FFA (51.1%, $p = 0.47$), as well as V1 (50.2%, $p = 1$), did not. Paired-sample t-tests showed that decoding performance was significantly higher than V1 in EBA ($t_{12} = 3.2$, $p = 0.0077$) but not in FBA ($t_{10} = 1.6$, $p = 0.15$). Furthermore, classification results using ROIs defined at a constant t-contrast threshold (Supplementary Fig. S2A) showed significant decoding of the sex of bodies in EBA (53.0%, $p = 0.0005$), OFA (52.2%, $p = 0.016$) and V1 (52.3%, $p = 0.0070$) but not in FBA (51.1%, $p = 0.43$).

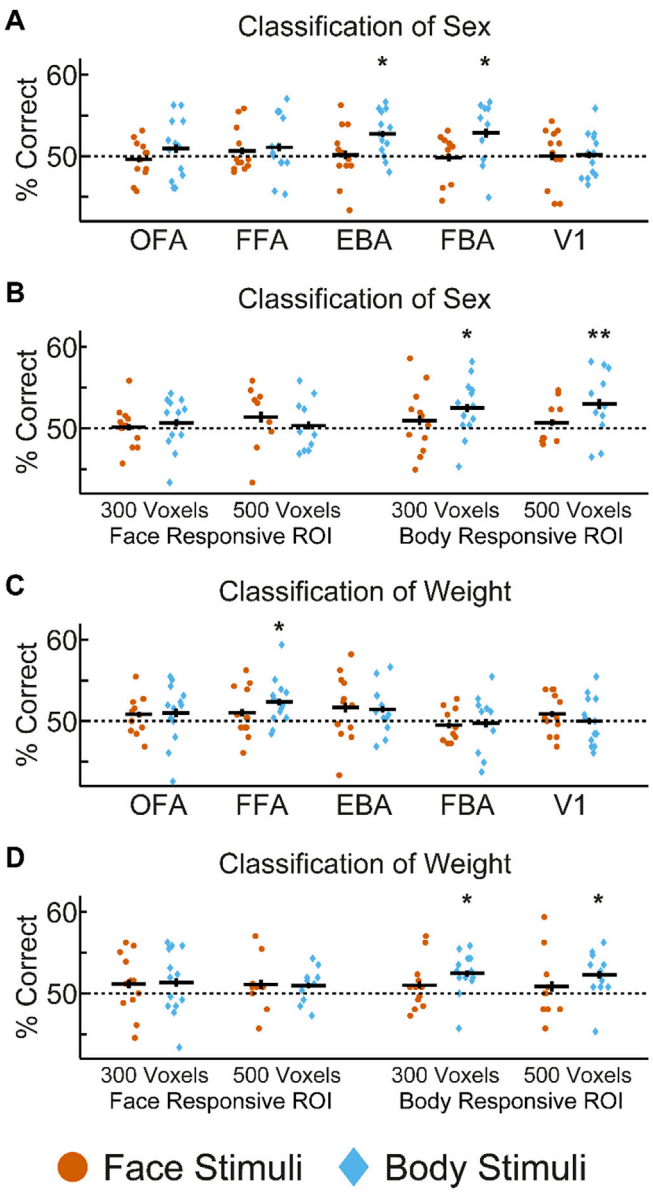

Fig. 5. Size-invariant classification of sex and weight. In this cross-classification analysis, a SVM classifier was trained on neural activity from subjects viewing stimuli of one image size and tested on its decoding performance on neural activity from the subjects viewing a different image size. (A) and (B) show classification results for sex in separate ROIs (A) and distributed ROIs (B). (C) and (D) show classification results for weight in separate ROIs (C) and distributed ROIs (D). Scatter plots show decoding accuracy for individual participants, horizontal black lines show group mean decoding accuracies and vertical error bars show the standard error of the mean.* indicates $p < 0.05$, ** indicates $p < 0.001$, Bonferroni-corrected for the number of ROIs (separate ROIs: N = 5; distributed ROIs: N = 4). Dotted lines indicate chance-level decoding performance, 50%.

These results suggest that decoding the sex of body stimuli across image size is most robust in the EBA.

Size-invariant classification of the weight of bodies was significant in FFA (52.3%, $p = 0.013$) but not in the body-related ROIs (EBA: 51.4%, $p = 0.20$; FBA: 49.8%, $p = 1$). A paired *t*-test showed that the classification of body weight in FFA was also significantly higher than in V1 ($t_{12} = 2.3$, $p = 0.042$). Classification results using ROIs defined at a constant t-contrast threshold (Supplementary Fig. S2C) showed significant decoding of body weight in OFA (52.3%, $p = 0.013$) but not in FFA (51.2%, $p = 0.40$). This difference may be due to an increase in information to the larger OFA ROI defined at a contrast t-contrast threshold, and an increase in noisy information hindering classification in the FFA ROI. No regions contained size-invariant information about the sex or weight of face stimuli.

For classification analyses using distributed ROIs, size-invariant classification of the sex of body stimuli was observed in both sizes of body-responsive ROI (300 voxels: 52.5%, $p = 0.0048$; 500 voxels: 53.0%, $p = 0.0008$) but not in face-responsive ROIs (300 voxels: 50.7%, $p = 0.85$; 500 voxels: 50.4%, $p = 1$). Similarly, size-invariant classification of the weight of bodies was found in both sizes of body-responsive ROI (300 voxels: 52.5%, $p = 0.0028$; 500 voxels: 52.3%, $p = 0.022$), but not in face-responsive ROIs (300 voxels: 51.4%, $p = 0.20$; 500 voxels: 51.0%, $p = 0.62$). Neither face- nor body-responsive ROIs could decode the sex or weight of face stimuli across image size (sex classification, face-responsive ROI, 500 voxels: 51.4%, $p = 0.32$; sex classification, body-responsive ROI, 500 voxels: 50.7%, $p = 0.94$; weight classification, face-responsive ROI, 500 voxels: 51.1%, $p = 0.57$; weight classification, body-responsive ROI, 500 voxels: 50.9%, $p = 0.73$).

The whole-brain searchlight analyses revealed no significant regions that could reliably decode the sex or weight of stimuli. This null-result, despite significant ROI decoding, might be due to the individual variance in ROI location that would weaken the group searchlight result in normalized space, combined with the overall weaker signal in the cross-decoding approach.

3.4. Classification of sex and weight across face and body stimuli

We investigated whether patterns of neural response to sex and weight would be able to generalize from patterns of activity evoked by faces to those evoked by bodies and vice-versa. To do this we trained SVM classifiers to distinguish sex and weight using neural activity data when participants viewed faces, and then tested the classifier on neural activity data when participants viewed bodies, and vice-versa. Significant decoding across face and body stimuli would suggest a semantic representation of the subcategory. We performed classification analyses with separate ROIs, distributed ROIs, and in searchlight analyses across the whole brain. The results are shown in Fig. 6.

We found significant decoding of weight across face and body stimuli in the EBA (51.5%, $p = 0.021$) and the 500 voxel body-responsive ROI (51.7%, $p = 0.016$), but not in any other ROIs (FBA: 50.6%, $p = 0.89$; OFA: 51.1%, $p = 0.14$; FFA: 49.8%, $p = 1$; V1: 49.4%, $p = 1$; face-responsive ROI, 500 voxels: 50.6%, $p = 0.84$). Paired-sample t-tests showed that decoding performance was significantly higher than V1 in EBA ($t_{12} = 3.3$, $p = 0.0068$) and the 500 voxel body-responsive ROI ($t_9 = 3.1$, $p = 0.012$). We additionally compared the classification performance of weight across face and body stimuli using ROIs defined at a constant t-contrast threshold (Supplementary Fig. S3C). We did not find significant decoding of weight in EBA (50.4%, $p = 1$) or the distributed body-responsive ROI (51.1%, $p = 0.063$) as defined using this method.

We were not able to decode sex across face and body stimuli in any of our separate or distributed ROIs (EBA: 50.0%, $p = 1$; FBA: 50.6%, $p = 0.9$; OFA: 50.4%, $p = 0.39$; FFA: 49.2%, $p = 1$; V1: 50.0%, $p = 1$; body-responsive ROI, 500 voxels: 50.1%, $p = 1$; face-responsive ROI, 500 voxels: 49.4%, $p = 1$). This non-significant decoding of sex across face and body stimuli is probably due to the fact that we were unable to identify any regions that could decode the sex of faces.

We additionally performed searchlight analyses to investigate if any other brain regions would be able to decode sex or weight across face and body stimuli. We did not identify any regions in these analyses.

3.5. Univariate analyses

To investigate whether the sex, weight, and image size of faces or bodies elicited different overall levels of neural activity, we conducted both ROI and whole-brain univariate analyses. The results are shown in Figs. 7–9.

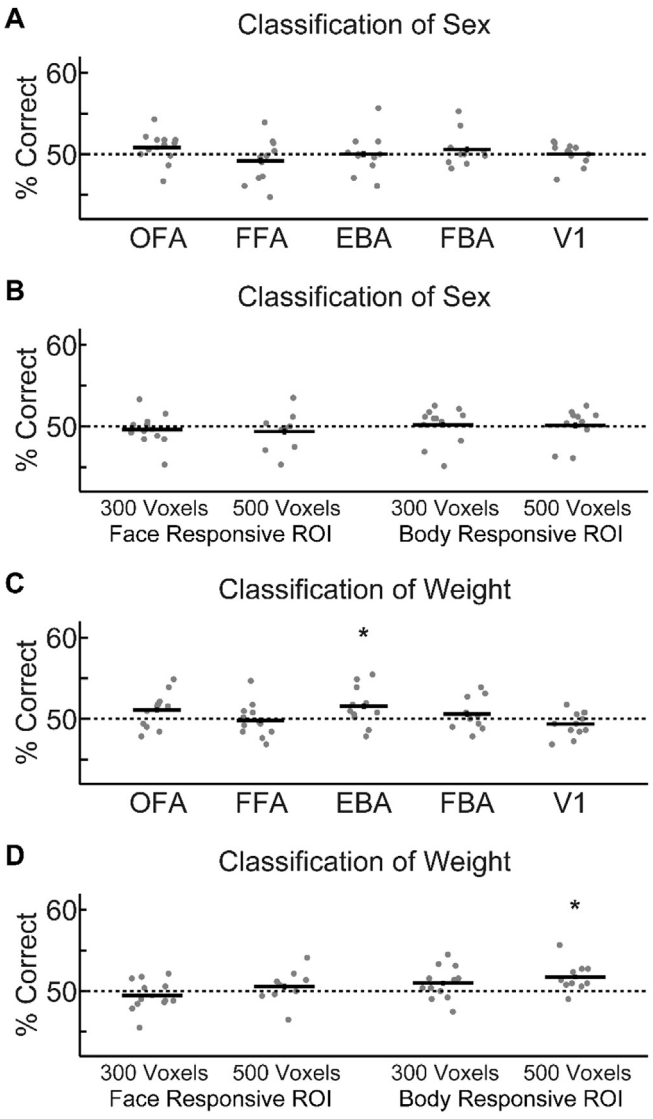

Fig. 6. Classification of sex and weight across face and body stimuli. In these cross-classification analyses, SVM classifiers were trained to distinguish sex and weight from neural activity when participants viewed faces and then subsequently tested on neural activity when participants viewed bodies, and vice-versa. (A) and (B) show classification results for sex in separate ROIs (A) and distributed ROIs (B). (C) and (D) show classification results for weight in separate ROIs (C) and distributed ROIs (D). Scatter plots show decoding accuracy for individual participants, horizontal black lines show group mean decoding accuracies and vertical error bars show the standard error of the mean.* indicates $p < 0.05$ Bonferroni-corrected for the number of ROIs (separate ROIs: N = 5; distributed ROIs: N = 4). Dotted lines indicate chance-level decoding performance, 50%.

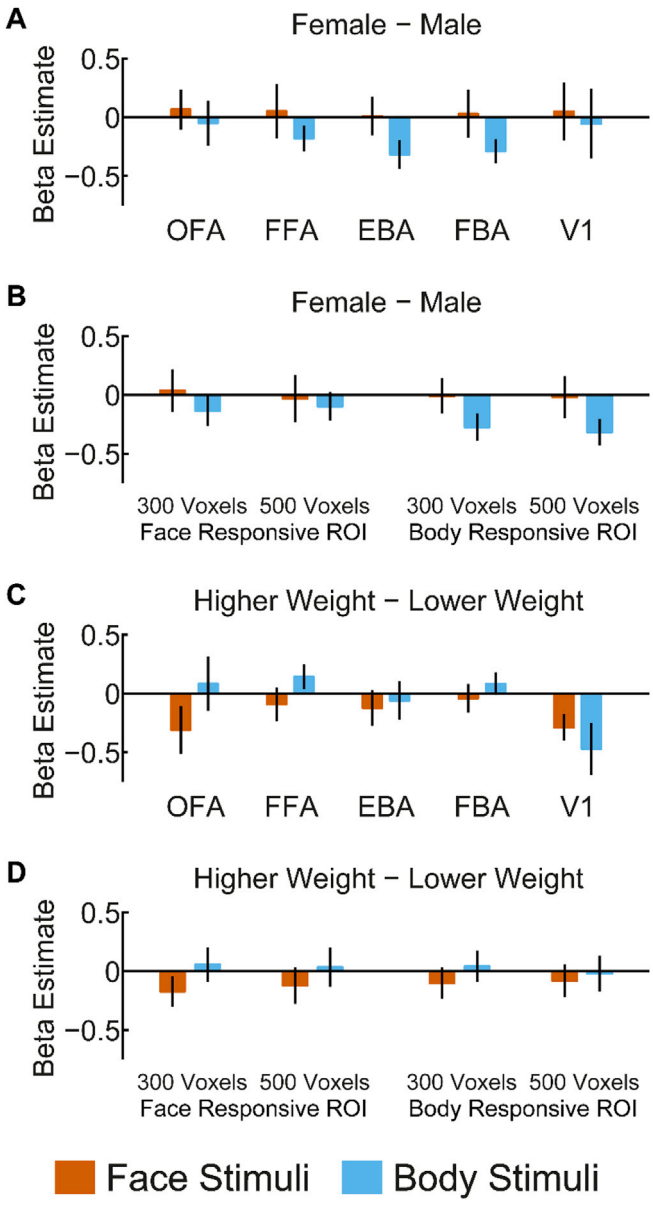

Fig. 7. Mean BOLD differences between the categories sex and weight. (A) and (B) illustrate mean differences between male and female stimuli for separate (A) and distributed (B) ROIs. Positive values indicate higher activation by female stimuli, negative values higher activation by male stimuli. (C) and (D) illustrate mean differences between higher and lower weight stimuli for separate (C) and distributed (D) ROIs. Positive values indicate higher activation by higher weight stimuli, negative values higher activation by lower weight stimuli. None of the differences were significant in any of the ROIs. Error bars show the standard error of the mean.

3.5.1. ROI analyses

Firstly, to investigate if there were any differences in mean response between faces or bodies of different sexes, we performed the contrast female stimuli minus male stimuli. None of the face- or body-responsive ROIs showed significant differences between male and female stimuli (Fig. 7A–B). There was a slight trend for higher activity to male bodies compared to female bodies in the body responsive regions, EBA ($t_{12} = -2.6$, $p = 0.12$), FBA ($t_{10} = -2.8$, $p = 0.098$), 300 voxel body-responsive ROI ($t_{12} = -2.4$, $p = 0.14$) and 500 voxel body-responsive ROI ($t_{10} = -2.9$, $p = 0.069$).

Secondly, we investigated if there were any differences in the mean response to faces or bodies of different weights, using the contrast higher weight stimuli minus lower weight stimuli. No significant differences between higher and lower weight stimuli were identified in any ROI (Fig. 7C–D). V1 showed a slight trend to higher activity to lower weight bodies compared to higher weight bodies ($t_{12} = -2.1$, $p = 0.27$) and to lower weight faces compared to higher weight faces ($t_{12} = -2.6$, $p = 0.12$).

Finally, we investigated if there were any differences in the mean response to faces or bodies of different image sizes (Fig. 8). For face stimuli, we found significantly higher BOLD activation to larger faces compared to smaller faces in FFA ($t_{11} = 8.2$, $p = 2.6 \times 10^{-5}$), FBA ($t_9 = 5.9$, $p = 0.0011$), both distributed face-responsive ROIs (300 voxels: $t_{11} = 3.8$, $p = 0.012$; 500 voxels: $t_8 = 3.7$, $p = 0.024$) and both distributed body-responsive ROIs (300 voxels: $t_{11} = 3.8$, $p = 0.013$; 500 voxels:

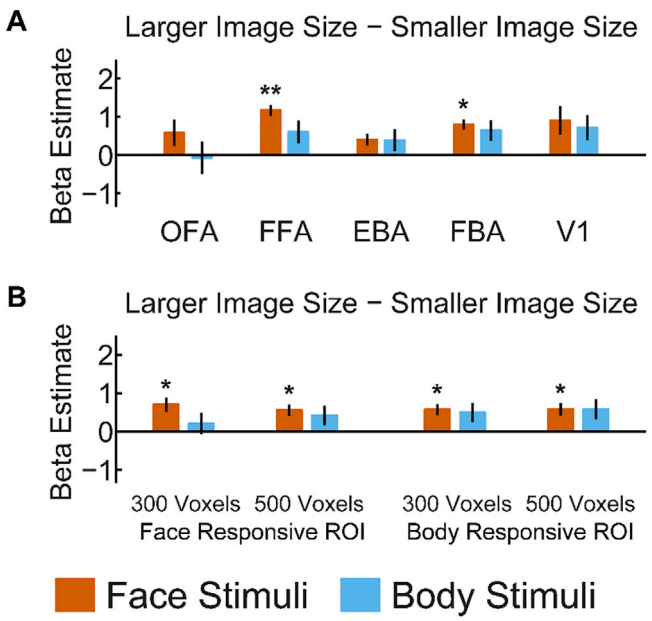

Fig. 8. Mean BOLD differences between larger and smaller images. (A) and (B) illustrate mean differences between larger and smaller image size stimuli for separate (A) and distributed (B) ROIs. Positive values indicate higher activation by larger size stimuli, negative values higher activation by smaller size stimuli. * indicates $p < 0.05$, ** indicates $p < 0.001$, Bonferroni-corrected for the number of ROIs (separate ROIs: N = 5; distributed ROIs: N = 4). Error bars show the standard error of the mean.

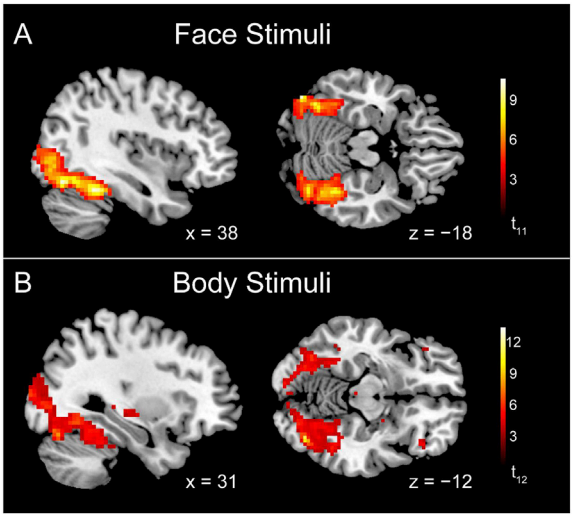

Fig. 9. Whole-brain results showing univariate activation differences between larger and smaller image size stimuli. (A) Brain regions showing higher activation to larger faces compared to smaller faces. These regions overlapped with the mean coordinates of the OFA, FFA and FBA bilaterally. (B) Brain regions showing higher activation to larger bodies compared to smaller bodies. These regions overlapped with the mean coordinates of the FBA, rEBA and rFFA.

$t_9 = 3.5$, $p = 0.027$). Additionally, a slight trend to higher activity to larger compared to smaller faces was seen in EBA ($t_{11} = 2.7$, $p = 0.10$) and V1 ($t_{11} = 2.4$, $p = 0.17$). For body stimuli, none of ROIs showed significant differences between larger and smaller bodies. A trend for higher activity to larger than smaller bodies was seen in FBA ($t_{10} = 2.4$, $p = 0.19$), V1 ($t_{12} = 2.2$, $p = 0.25$) and the 500-voxel body-responsive ROI ($t_{10} = 2.2$, $p = 0.20$).

3.5.2. Whole-brain analyses

To investigate if any other brain regions were sensitive to the sex, weight, and image size of faces and bodies, we performed the same three contrasts in whole-brain analyses. We found that the occipitotemporal cortex showed higher brain activity for larger stimuli than smaller stimuli, for both faces (Fig. 9A) and bodies (Fig. 9B). For face stimuli the significant area overlapped with the mean coordinates of the OFA, FFA and FBA bilaterally. For body stimuli the significant area overlapped with the mean coordinates of the FBA, rEBA and rFFA. As for the sex and weight of stimuli, no brain regions showed significant differences between male and female stimuli or between higher- and lower-weight stimuli.

4. Discussion

In this study, we investigated how face- and body-responsive brain regions encode information about the subcategories sex and weight. We show, for the first time to our knowledge, that subcategorical information about bodies is encoded in both body- and face-responsive areas in the brain, and that this information is encoded in a size-invariant manner, more so for the body-than face-related brain network. Furthermore, we find evidence that the FBA responds to the weight of faces, and that weight is encoded in an abstract manner in the EBA and distributed body-responsive network. These results indicate that subcategories shared by faces and bodies (e.g. sex) are encoded by different patterns of neural responses in the brain network related to person perception.

4.1. Neural coding of body subcategories

We found that both face-responsive and body-responsive regions encoded information about the sex and weight of bodies. For the sex of bodies, higher-than-chance decoding was observed not only in all body-responsive ROIs but also in all face-responsive ROIs. For the weight of bodies, both the distributed body-responsive areas and the largest-size distributed face-responsive area showed higher-than-chance classification of body weight. For the separate ROI analyses, higher-than-chance decoding of body weight was observed in the EBA as well as the face-responsive ROIs OFA and FFA. In contrast to the differential multivariate activity patterns, no ROI showed stronger net activity for one subcategory than another.

Our searchlight results showed that brain regions responding to the weight of bodies overlapped with those responding to the sex of bodies, though the latter is more widespread. This result shows similarity to a recent behavioural finding that judgements of the sex of bodies are independent of judgements of the weight of bodies but not vice-versa (Johnstone and Downing, 2017). This coincidence suggests that the above behavioural difference might be related to how the brain regions processing these two subcategories overlap. The encoding of body subcategory information in face-responsive areas indicates that these face areas are not exclusively involved in face processing (cf. Kanwisher and Yovel, 2006). These results are in line with previous findings showing that categorical (e.g. faces and bodies) and subcategorical information is distributed across occipitotemporal cortex, rather than selective to specific sub-regions of cortex (Haxby et al., 2001; Huth et al., 2012; Op de Beeck et al., 2010).

Although both face- and body-responsive brain areas encoded the sex and weight of bodies, size-invariant body information was more prominent and consistent in the body-responsive areas. When we used data obtained from one image size to train the classifier and then tested it with the data obtained from a different image size, we were able to decode the sex of bodies from EBA, FBA and both distributed body-responsive regions, but not from any face-responsive region. Similarly, the two distributed body-responsive ROIs, but not face-responsive ROIs, showed higher-than-chance decoding of body weight in a size-invariant manner. Although the FFA showed size-invariant decoding of the weight of bodies, in total less face-responsive ROIs showed size-invariant decoding

of body subcategories than body-responsive ROIs.

In our first classification analysis we found that V1 was able to decode the sex of bodies and the weight of bodies and faces, suggesting there may be some low-level visual information that could be used for classification in these analyses (where both image-sizes were present in both the training and test set). In contrast, size-invariant decoding of sex and weight was not possible in our V1 ROI.

4.2. Neural coding of face subcategories

We were able to decode the weight of faces from the FBA, and classify weight across face and body stimuli in the EBA and distributed body-responsive network. The decoding of weight across face and body stimuli in the EBA and distributed body-responsive network suggests that these regions may contain a semantic encoding of weight (i.e. not dependent upon low-level visual features). Although the sex of faces can be clearly differentiated in perception (as shown by perceptual ratings), we were unable to decode the sex of faces based on the pattern of neural activity recorded with fMRI. Previous studies have found mixed results about whether the sex of faces can be decoded based on the patterns of neural activity in face-responsive brain areas. Some studies were not able to decode the sex of faces from the occipitotemporal cortex (Kriegeskorte et al., 2007) and the FFA (Kanwisher, 2017), whereas others showed that the sex of faces can be successfully decoded from the FFA or extended face network (Contreras et al., 2013; Kaul et al., 2011). One factor that may cause such discrepant results is the stimuli. Our face stimuli were carefully controlled and lacked external face features (such as hair or make-up) whereas Contreras et al. (2013) and Kaul et al. (2011) used photographs of faces varying in hairstyle, make-up, and beards. These external cues are often diagnostic for the sex of faces and facilitate sex categorization (Rossion, 2002). Thus, despite clear perceptual differences shown in the rating task, the sex differences in our face stimuli might be insufficient to elicit distinct patterns of brain activity that can be detected and decoded using MVPA. Note that a lack of ability to decode face subcategories does not necessarily reflect a lack of information about the subcategories in these regions, it may simply be beyond the resolution that fMRI MVPA can detect. For instance (Dubois et al., 2015), compared identity decoding from fMRI recordings with that from single cell recordings in the anterior temporal cortex of macaques: they found that identity information decodable from single cells was not decodable from the fMRI data.

Another factor that might have hindered our ability to decode the sex of faces was the task. Our attention task (i.e. detecting a red dot) did not involve any effortful processing of faces, whereas both Contreras et al. (2013) and Kaul et al. (2011) employed tasks encouraging effortful face processing (i.e. judgments about the gender or attractiveness of faces). Given that task-related attention has been shown to affect the neural representation of faces and semantic categories in natural vision (Çukur et al., 2013; Dobs et al., 2018; Kaiser et al., 2016), the automatic processing of faces in our study may lead to weak responses to the sex of faces, thereby reducing the possibility of higher-than-chance decoding. In addition, while sex and weight are two of the most salient dimensions that differentiate the shape of bodies (Hill et al., 2016), they are probably not the most important ones for faces (e.g. in comparison with the identity or emotional expression of faces (Burton et al., 1999; O'Toole et al., 2011)). In line with this, we found greater differences in perceptual ratings between male and female stimuli and between higher and lower weight stimuli for body stimuli than for face stimuli. We note however that lower rating differences cannot be the only reason for the differences we see in decoding performance as rating differences of the sex of faces (for which we do not find significant decoding in any region) are considerably larger than for the weight of bodies and faces (for which we find significant decoding in a number of brain regions). Lastly, our relatively small sample size ($N = 12$ for face stimuli) might also have impacted our ability to decode the sex of faces. A larger sample size might be able to improve the decoding of the sex of faces.

4.3. Neural coding of subcategories shared by faces and bodies

Previous studies investigating neural processing of face and body information have suggested that such information is supported by largely separated neural networks in occipitotemporal cortex (Peelen and Downing, 2007; Pitcher et al., 2009). This segregated neural processing of faces and bodies has been demonstrated with both human and nonhuman primates (Premereur et al., 2016). For example, neuroimaging studies in macaques and humans have shown that responses to whole-agents (i.e. where the image contains both the face and body) can be best modelled by a linear combination of responses to faces and bodies alone, suggesting that there are separate neural populations responsive to face and body information respectively (Fisher and Freiwald, 2015; Kaiser et al., 2014). In contrast to this parallel processing hypothesis, we found that the EBA and distributed body-responsive network could classify weight across face and body stimuli, suggesting that these regions encode some shared subcategorical information from faces and bodies in an abstract manner. This abstract coding may be semantic or based on high-level visual features (for example concavity vs convexity). Although sex can also be perceived from both faces and bodies, we were unable to identify brain regions showing stimuli-independent encoding of sex. In behaviour, psychological adaptation studies have demonstrated that perception of gender (Ghuman et al., 2010; Palumbo et al., 2014) can adapt across face and body stimuli, suggesting an overlapping representation.

Recent studies have suggested that the anterior temporal lobes may contain an integrated representation of body and face information. For example, the right ATFA showed an equivalent level of brain activity to faces and headless bodies and exhibited significant correlation between face- and body-elicited neural responses (Harry et al., 2016). Similarly, the anterior temporal face patch (ATFP) in macaques has been shown to respond more strongly to whole-agents than the addition of the responses to the face and body shown separately, suggesting an integration of face and body information in this region (Fisher and Freiwald, 2015). Such integration may help differentiate individuals' identity, as this area has been linked to identity representation in humans (Anzellotti et al., 2014; Guntupalli et al., 2016; Kriegeskorte et al., 2007). However, our results (from searchlight analyses across image size or across face and body stimuli) suggest that this identity-sensitive area does not automatically encode the sex or weight of a person in an abstract manner.

4.4. Effect of low-level stimulus features

We found consistently enhanced neural responses to faces and bodies when the size of images increased. For both faces and bodies, significantly higher brain activity for larger than smaller stimuli was observed in distributed areas across occipitotemporal cortex, covering face- and body-responsive ROIs, respectively. For face stimuli, the active area overlapped with the mean coordinates of face-responsive OFA and FFA, as well as body-responsive FBA. For body stimuli the overlap was with the mean coordinates of body-responsive FBA and rEBA, as well as face-responsive rFFA. Neural responses to faces appeared to be more sensitive to the stimulus size than those to bodies (see Figs. 8 and 9). The influence of stimulus size on neural response of high-level visual areas has been shown for faces (Yue et al., 2011) and general everyday objects (Konkle and Oliva, 2012), but not, to our knowledge, for human bodies. These results demonstrate that neural responses to faces and bodies in high-level visual areas are modulated by low-level stimulus properties, rather than being size-invariant (cf. Andrews and Ewbank, 2004; Sawamura et al., 2005).

5. Conclusion

Our study provides the first evidence, to our knowledge, that the sex and weight of human bodies can be decoded from neural activity in the person-related brain network using MVPA. By demonstrating size-

invariant decoding of body subcategories in both the body- and face-responsive brain network, as well as cross-classification of weight across face and body stimuli, we show that the neural responses to these subcategories are largely driven by high-level visual or semantic features rather than by merely low-level visual features. The present study also offers an alternative approach to repetition suppression for investigating neural responses to subcategorical body information using fMRI. Methods like repetition suppression may be biased by top-down effects (Summerfield et al., 2008). Together, these findings not only offer new insights into how the brain encodes person-related visual information like faces and bodies, but also shed light on how shared subcategories from visually distinctive object categories may be encoded in the human brain.

Acknowledgements

This research was supported by the Max Planck Society, Germany.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.neuroimage.2019.116085.

References

Aguirre, G.K., 2007. Continuous carry-over designs for fMRI. Neuroimage 35, 1480–1494. https://doi.org/10.1016/j.neuroimage.2007.02.005.

Andrews, T.J., Ewbank, M.P., 2004. Distinct representations for facial identity and changeable aspects of faces in the human temporal lobe. Neuroimage 23, 905–913. https://doi.org/10.1016/j.neuroimage.2004.07.060.

Anzellotti, S., Fairhall, S.L., Caramazza, A., 2014. Decoding representations of face identity that are tolerant to rotation. Cerebr. Cortex 24, 1988–1995. https://doi.org/10.1093/cercor/bht046.

Axelrod, V., Yovel, G., 2015. Successful decoding of famous faces in the fusiform face area. PLoS One 10, 1–20. https://doi.org/10.1371/journal.pone.0117126.

Blanz, V., Vetter, T., 1999. A morphable model for the synthesis of 3D faces. SIGGRAPH'99 Conf. Proc. 187–194. https://doi.org/10.1145/311535.311556.

Bracci, S., Op de Beeck, H., 2016. Dissociations and associations between shape and category representations in the two visual pathways. J. Neurosci. 36, 432–444. https://doi.org/10.1523/JNEUROSCI.2314-15.2016.

Brainard, D.H., 1997. The Psychophysics toolbox. Spat. Vis. 10, 433–436. https://doi.org/10.1163/156856897X00357.

Brooks, J.L., 2012. Counterbalancing for serial order carryover effects in experimental condition orders. Psychol. Methods 17, 600–614. https://doi.org/10.1037/a0029310.

Burton, A.M., Wilson, S., Cowan, M., Bruce, V., 1999. Face recognition in poor quality video. Psychol. Sci. 10, 243–248. https://doi.org/10.1111/1467-9280.00144.

Cohen, L., Dehaene, S., Naccache, L., Lehéricy, S., Dehaene-Lambertz, G., Hénaff, M., Michel, F., 2000. The visual word form area: spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. Brain 123, 291–307.

Contreras, J.M., Banaji, M.R., Mitchell, J.P., 2013. Multivoxel patterns in fusiform face area differentiate faces by sex and race. PLoS One 8, 1–6. https://doi.org/10.1371/journal.pone.0069684.

Cox, D.D., Savoy, R.L., 2003. Functional magnetic resonance imaging (fMRI) "brain reading": detecting and classifying distributed patterns of fMRI activity in human visual cortex. Neuroimage 19, 261–270. https://doi.org/10.1016/S1053-8119(03)00049-1.

Çukur, T., Nishimoto, S., Huth, A.G., Gallant, J.L., 2013. Attention during natural vision warps semantic representation across the human brain. Nat. Neurosci. 16, 763–770. https://doi.org/10.1038/nn.3381.

Dobs, K., Schultz, J., Bülthoff, I., Gardner, J.L., 2018. Task-dependent enhancement of facial expression and identity representations in human cortex. Neuroimage 172, 689–702. https://doi.org/10.1016/j.neuroimage.2018.02.013.

Downing, P.E., Jiang, Y., Shuman, M., Kanwisher, N., 2001. A cortical area selective for visual processing of the human body. Science 293, 2470–2473. https://doi.org/10.1126/science.1063414.

Dubois, X.J., de Berker, A.O., Tsao, D.Y., 2015. Single-unit recordings in the macaque face patch system reveal limitations of fMRI MVPA. J. Neurosci. 35, 2791–2802. https://doi.org/10.1523/JNEUROSCI.4037-14.2015.

Epstein, R., Kanwisher, N., 1998. A cortical representation of the local visual environment. Nature 392, 598–601. https://doi.org/10.1038/33402.

Ewbank, M.P., Lawson, R.P., Henson, R.N., Rowe, J.B., Passamonti, L., Calder, A.J., 2011. Changes in "Top-Down" connectivity underlie repetition suppression in the ventral visual pathway. J. Neurosci. 31, 5635–5642. https://doi.org/10.1523/JNEUROSCI.5013-10.2011.

Fisher, C., Freiwald, W.A., 2015. Whole-agent selectivity within the macaque face-processing system. Proc. Natl. Acad. Sci. 112, 14717–14722. https://doi.org/10.1073/pnas.1512378112.

Freeman, J.B., Rule, N.O., Adams, R.B., Ambady, N., 2010. The neural basis of categorical face perception: graded representations of face gender in fusiform and orbitofrontal cortices. Cerebr. Cortex 20, 1314–1322. https://doi.org/10.1093/cercor/bhp195.

Gardumi, A., Ivanov, D., Hausfeld, L., Valente, G., Formisano, E., Uludağ, K., 2016. The effect of spatial resolution on decoding accuracy in fMRI multivariate pattern analysis. Neuroimage 132, 32–42. https://doi.org/10.1016/j.neuroimage.2016.02.033.

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. J. Cogn. Neurosci. 12, 495–504. https://doi.org/10.1162/089892900562165.

Ghuman, A.S., McDaniel, J.R., Martin, A., 2010. Face adaptation without a face. Curr. Biol. 20, 32–36. https://doi.org/10.1016/j.cub.2009.10.077.

Grill-Spector, K., Weiner, K.S., 2014. The functional architecture of the ventral temporal cortex and its role in categorization. Nat. Rev. Neurosci. 15, 536–548. https://doi.org/10.1038/nrn3747.

Gross, C.G., Roch-Miranda, C.E., Bender, D.B., 1972. Visual properties of neurons in inferotemporal of the macaque. J. Neurophysiol. 35, 96–111.

Guntupalli, J.S., Wheeler, K.G., Gobbini, M.I., 2016. Disentangling the representation of identity from head view along the human face processing pathway. Cerebr. Cortex 27, 46–53. https://doi.org/10.1093/cercor/bhw344.

Hahn, C.A., O'Toole, A.J., 2017. Recognizing approaching walkers: neural decoding of person familiarity in cortical areas responsive to faces, bodies, and biological motion. Neuroimage 146, 859–868. https://doi.org/10.1016/j.neuroimage.2016.10.042.

Harry, B., Umla-Runge, K., Lawrence, A., Graham, K., Downing, P., 2016. Evidence for integrated visual face and body representations in the anterior temporal lobes. J. Cogn. Neurosci. 28, 1178–1193. https://doi.org/10.1162/jocn_a_00966.

Haxby, J.V., Gobbini, M.I., Furey, M.L., Ishai, A., Schouten, J.L., Pietrini, P., 2001. Distributed and overlapping representations of faces and objects in ventral temporal cortex. Science 293, 2425–2430. https://doi.org/10.1126/science.1063736.

Hebart, M.N., Görgen, K., Haynes, J.-D., 2015. The Decoding Toolbox (TDT): a versatile software package for multivariate analyses of functional imaging data. Front. Neuroinf. 8, 1–18. https://doi.org/10.3389/fninf.2014.00088.

Hill, M.Q., Streuber, S., Hahn, C.A., Black, M.J., O'Toole, A.J., 2016. Creating body shapes from verbal descriptions by linking similarity spaces. Psychol. Sci. 27, 1486–1497. https://doi.org/10.1177/0956797616663878.

Hinds, O., Polimeni, J.R., Rajendran, N., Balasubramanian, M., Amunts, K., Zilles, K., Schwartz, E.L., Fischl, B., Triantafyllou, C., 2009. Locating the functional and anatomical boundaries of human primary visual cortex. Neuroimage 46, 915–922. https://doi.org/10.1016/j.neuroimage.2009.03.036.

Hummel, D., Rudolf, A.K., Brandi, M.L., Untch, K.H., Grabhorn, R., Hampel, H., Mohr, H.M., 2013. Neural adaptation to thin and fat bodies in the fusiform body area and middle occipital gyrus: an fMRI adaptation study. Hum. Brain Mapp. 34, 3233–3246. https://doi.org/10.1002/hbm.22135.

Huth, A.G., Nishimoto, S., Vu, A.T., Gallant, J.L., 2012. A continuous semantic space describes the representation of thousands of object and action categories across the human brain. Neuron 76, 1210–1224. https://doi.org/10.1016/j.neuron.2012.10.014.

Jeong, S.K., Xu, Y., 2016. Behaviorally relevant abstract object identity representation in the human parietal cortex. J. Neurosci. 36, 1607–1619. https://doi.org/10.1523/JNEUROSCI.1016-15.2016.

Johnstone, L.T., Downing, P.E., 2017. Dissecting the visual perception of body shape with the Garner selective attention paradigm. Vis. Cogn. 25, 507–523. https://doi.org/10.1080/13506285.2017.1334733.

Kaiser, D., Oosterhof, N.N., Peelen, M.V., 2016. The neural dynamics of attentional selection in natural scenes. J. Neurosci. 36, 10522–10528. https://doi.org/10.1523/JNEUROSCI.1385-16.2016.

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. J. Neurophysiol. 111, 82–90. https://doi.org/10.1152/jn.00371.2013.

Kanwisher, N., 2017. The quest for the FFA and where it led. J. Neurosci. 37, 1056–1061. https://doi.org/10.1523/JNEUROSCI.1706-16.2016.

Kanwisher, N., McDermott, J., Chun, M.M., 1997. The fusiform face area: a module in human extrastriate cortex specialized for face perception. J. Neurosci. 17, 4302–4311.

Kanwisher, N., Yovel, G., 2006. The fusiform face area: a cortical region specialized for the perception of faces. Philos. Trans. R. Soc. Biol. Sci. 361, 2109–2128. https://doi.org/10.1098/rstb.2006.1934.

Kaul, C., Rees, G., Ishai, A., 2011. The gender of face stimuli is represented in multiple regions in the human brain. Front. Hum. Neurosci. 4, 1–12. https://doi.org/10.3389/fnhum.2010.00238.

Kleiner, M., Brainard, D., Pelli, D., 2007. What's new in Psychtoolbox-3?. In: Perception 36 ECVP Abstract Supplement.

Konkle, T., Oliva, A., 2012. A real-world size organization of object responses in occipitotemporal cortex. Neuron 74, 1114–1124. https://doi.org/10.1016/j.neuron.2012.04.036.

Kriegeskorte, N., Formisano, E., Sorger, B., Goebel, R., 2007. Individual faces elicit distinct response patterns in human anterior temporal cortex. Proc. Natl. Acad. Sci. 104, 20600–20605. https://doi.org/10.1073/pnas.0705654104.

Kriegeskorte, N., Mur, M., Ruff, D.A., Kiani, R., Bodurka, J., Esteky, H., Tanaka, K., Bandettini, P.A., 2008. Matching categorical object representations in inferior temporal cortex of man and monkey. Neuron 60, 1126–1141. https://doi.org/10.1016/j.neuron.2008.10.043.

Logothetis, N.K., Sheinberg, D.L., 1996. Visual object recognition. Annu. Rev. Neurosci. 19, 577–621.

Loper, M., Mahmood, N., Romero, J., Pons-moll, G., Black, M.J., 2015. SMPL : a skinned multi-person linear model. ACM Trans. Graph. (Proc. SIGGRAPH Asia) 34 (248), 1–248, 16. https://doi.org/10.1145/2816795.2818013.

Malach, R., Reppas, J.B., Benson, R.R., Kwong, K.K., Jiang, H., Kennedy, W.A., Ledden, P.J., Brady, T.J., Rosen, B.R., Tootell, R.B.H., 1995. Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. Proc. Natl. Acad. Sci. U.S.A. 92, 8135–8139. https://doi.org/10.1073/pnas.92.18.8135.

O'Toole, A.J., Phillips, P.J., Weimer, S., Roark, D.A., Ayyad, J., Barwick, R., Dunlop, J., 2011. Recognizing people from dynamic and static faces and bodies: dissecting identity with a fusion approach. Vis. Res. 51, 74–83. https://doi.org/10.1016/j.visres.2010.09.035.

Op de Beeck, H.P., Brants, M., Baeck, A., Wagemans, J., 2010. Distributed subordinate specificity for bodies, faces, and buildings in human ventral visual cortex. Neuroimage 49, 3414–3425. https://doi.org/10.1016/j.neuroimage.2009.11.022.

Palumbo, R., D'Ascenzo, S., Tommasi, L., 2014. Cross-category adaptation: exposure to faces produces gender aftereffects in body perception. Psychol. Res. 79, 380–388. https://doi.org/10.1007/s00426-014-0576-2.

Peelen, M.V., Downing, P.E., 2007. The neural basis of visual body perception. Nat. Rev. Neurosci. 8, 636–648. https://doi.org/10.1038/nrn2195.

Peelen, M.V., Downing, P.E., 2005. Selectivity for the human body in the fusiform gyrus. J. Neurophysiol. 93, 603–608. https://doi.org/10.1152/jn.00513.2004.

Pitcher, D., Charles, L., Devlin, J.T., Walsh, V., Duchaine, B., 2009. Triple dissociation of faces, bodies, and objects in extrastriate cortex. Curr. Biol. 19, 319–324. https://doi.org/10.1016/j.cub.2009.01.007.

Premereur, E., Taubert, J., Janssen, P., Vogels, R., Vanduffel, W., 2016. Effective connectivity reveals largely independent parallel networks of face and body patches. Curr. Biol. 26, 3269–3279. https://doi.org/10.1016/j.cub.2016.09.059.

Ratner, K.G., Kaul, C., Van Bavel, J.J., 2013. Is race erased? decoding race from patterns of neural activity when skin color is not diagnostic of group boundaries. Soc. Cogn. Affect. Neurosci. 8, 750–755. https://doi.org/10.1093/scan/nss063.

Robinette, K.M., Blackwell, S., Daanen, H., Boehmer, M., Fleming, S., Brill, T., Hoeferlin, D., Burnsides, D., 2002. Civilian American and European Surface Anthropometry Resource (CAESAR), Final Report. US Air Force Research Laboratory. Tech. Rep. AFRL-HE- WP-TR-2002-0169.

Rossion, B., 2002. Is sex categorization from faces really parallel to face recognition? Vis. Cogn. 9, 1003–1020. https://doi.org/10.1080/13506280143000485.

Sawamura, H., Georgieva, S., Vogels, R., Vanduffel, W., Orban, G.A., 2005. Using functional magnetic resonance imaging to assess adaptation and size invariance of shape processing by humans and monkeys. J. Neurosci. 25, 4294–4306. https://doi.org/10.1523/JNEUROSCI.0377-05.2005.

Schwarzlose, R.F., Baker, C.I., Kanwisher, N., 2005. Separate face and body selectivity on the fusiform gyrus. J. Neurosci. 25, 11055–11059. https://doi.org/10.1523/JNEUROSCI.2621-05.2005.

Sturman, D., Stephen, I.D., Mond, J., Stevenson, R.J., Brooks, K.R., 2017. Independent Aftereffects of Fat and Muscle: implications for neural encoding, body space representation, and body image disturbance. Sci. Rep. 7, 1–8. https://doi.org/10.1038/srep40392.

Summerfield, C., Trittschuh, E.H., Monti, J.M., Mesulam, M.M., Egner, T., 2008. Neural repetition suppression reflects fulfilled perceptual expectations. Nat. Neurosci. 11, 1004–1006. https://doi.org/10.1038/nn.2163.

Tanaka, K., 1996. Inferotemporal cortex and object vision. Annu. Rev. Neurosci. 19, 109–139. https://doi.org/10.1146/annurev.neuro.19.1.109.

Troje, N.F., Bülthoff, H.H., 1996. Face recognition under varying poses: the role of texture and shape. Vis. Res. 36, 1761–1771. https://doi.org/10.1016/0042-6989(95)00230-8.

Yue, X., Cassidy, B.S., Devaney, K.J., Holt, D.J., Tootell, R.B.H., 2011. Lower-level stimulus features strongly influence responses in the fusiform face area. Cerebr. Cortex 21, 35–47. https://doi.org/10.1093/cercor/bhq050.