Systems/Circuits

Heterogeneous Single-Unit Selectivity in an fMRI-Defined Body-Selective Patch

Ivo D. Popivanov,[1] **Jan Jastorff,**[1] **Wim Vanduffel,**[1,2,3] **and Rufin Vogels**[1]

[1]Laboratorium voor Neuroen Psychofysiologie, KU Leuven, BE-3000 Leuven, Belgium, [2]Massachusetts General Hospital, Martinos Center for Biomedical Imaging, Charlestown, Massachusetts 02129, and [3]Harvard Medical School, Boston, Massachusetts 02115

Although the visual representation of bodies is essential for reproduction, survival, and social communication, little is known about the mechanisms of body recognition at the single neuron level. Imaging studies showed body-category selective regions in the primate occipitotemporal cortex, but it is difficult to infer the stimulus selectivities of the neurons from the population activity measured in these fMRI studies. To overcome this, we recorded single unit activity and local field potentials (LFPs) in the middle superior temporal sulcus body patch, defined by fMRI in the same rhesus monkeys. Both the spiking activity, averaged across single neurons, and LFP gamma power in this body patch was greater for bodies (including monkey bodies, human bodies, mammals, and birds) compared with other objects, which fits the fMRI activation. Single neurons responded to a small proportion of body images. Thus, the category selectivity at the population level resulted from averaging responses of a heterogeneous population of single units. Despite such strong within-category selectivity at the single unit level, two distinct clusters, bodies and nonbodies, were present when analyzing the responses at the population level, and a classifier that was trained using the responses to a subset of images was able to classify novel images of bodies with high accuracy. The body-patch neurons showed strong selectivity for individual body parts at different orientations. Overall, these data suggest that single units in the fMRI-defined body patch are biased to prefer bodies over nonbody objects, including faces, with a strong selectivity for individual body images.

Introduction

Visual representations of bodies of conspecifics and other animals are instrumental for survival. Primates can categorize animals versus nonanimals fast and accurately (Fabre-Thorpe et al., 1998). Headless bodies are detected as fast as faces in scenes, suggesting that not only faces but also body cues contribute to person detection (Bindemann et al., 2010). Nonverbal communication is partially based on the analysis of body shape (de Gelder et al., 2010). Additionally, body posture coding can contribute to action recognition (Giese and Poggio, 2003; Vangeneugden et al., 2011).

Despite this ethological importance of body recognition, little is known about its neural mechanisms. fMRI studies in primates identified occipitotemporal areas that are activated more strongly by images of bodies or body parts compared with other object categories, including faces (Downing et al., 2001; Tsao et al., 2003; Pinsk et al., 2005, 2009; Bell et al., 2009; Popivanov et al.,

2012). However, because fMRI reflects the activity of a large population of neurons, these studies do not inform about the stimulus selectivities of the neurons in these body-selective regions.

Previous fMRI-guided single unit studies were mainly restricted to fMRI-defined face patches, showing that face patches contain a high fraction of face-selective cells (Tsao et al., 2006; Issa and DiCarlo, 2012) and that many face-patch cells respond to a wide variety of face images, including human, macaque, and cartoon faces (Tsao et al., 2006; Freiwald et al., 2009; Freiwald and Tsao, 2010). This raises the question whether the same holds for body patches: do single neurons in a body patch prefer images of bodies compared with other objects, including faces, and do they respond similarly to different body images? Thus far, only one study recorded in a macaque body patch (Bell et al., 2011), reporting that approximately half of the neurons responded stronger to body parts compared with other object classes, which is less than observed for face selectivity in the face patches (Freiwald and Tsao, 2010). No data exist regarding the clustering or selectivity for individual body, or other stimuli in the body patches. Thus, in general little is known about the stimulus and category selectivity of fMRI defined body-patch neurons.

To bridge this gap in our understanding of stimulus processing in the body patches, we recorded single-unit activity and local field potentials (LFPs) within an fMRI defined body selective patch. Previously, we localized two patches inside the superior temporal sulcus (STS) that were activated more strongly by images of monkey bodies compared with control objects, matched in low level image properties, in four monkeys (Popivanov et al., 2012). Here, we recorded single units and LFPs in the posterior,

so called midSTS body patch, in two of these animals, examining their selectivity for animate and inanimate categories and for individual exemplars of these categories. We asked how the neurons that comprise the body patch represent exemplars of the different categories, whether exemplars of the body category cluster together, whether the population responses can classify bodies versus other objects, and whether they show body part selectivity.

Materials and Methods

Subjects

The two male rhesus monkeys (*Macaca mulatta*) were 2 of 4 subjects for our previous fMRI study (Popivanov et al., 2012). They were implanted with a magnetic resonance (MR) compatible headpost and a recording chamber targeting the midSTS. Animal care and experimental procedures complied with the National, European, and National Institute of Health guidelines and were approved by the Ethical Committee of the KU Leuven Medical School.

Stimuli

Main stimulus set. Ten classes of achromatic images, monkey and human bodies (excluding the head), monkey and human faces, four-legged mammals, birds, manmade objects (matched either to the monkey or to the human bodies), fruits/vegetables, and body-like sculptures (by the British artist H. Moore), served as stimuli in the electrophysiological study. Each class consisted of the 10 images which were previously used in the even runs of the fMRI study of Popivanov et al. (2012). Examples of the stimuli are shown on Figure 1A, whereas the full stimulus set together with details about the stimuli can be found in Popivanov et al. (2012). Briefly, the images of monkey bodies depicted headless bodies in different postures and the monkey faces varied in both orientation and viewpoint (profile to frontal views). Most of the images of human bodies were from Downing et al. (2001). The human face stimuli (courtesy of M. J. Tarr, http://www.tarrlab.org/ and the NBU Faces Database, http://nbufaces.yobul.com/ENAboutDatabase.aspx) depicted different individuals and varied in viewpoint. All other stimuli were generated from images downloaded from the public domain.

We made every effort to equate the low-level image characteristics, such as mean luminance, mean contrast, and aspect ratio, across the different stimulus classes. The mean aspect ratio of the monkey and human bodies differed because the upright human bodies tend to be more elongated than the monkey bodies. This was controlled for by using two classes of manmade objects, one matching the aspect ratio of the monkey bodies (objects M) and another one matching the aspect ratio of the human bodies (objects H). The images were resized so that the average area per class was matched across all classes, except for the objects H and human bodies, but still allowing some variation in area (range, 3.7–6.7°; square root of the area) within each class. This variation in size avoided potential clustering of the image classes based on local, pixel-based gray level differences. The mean vertical and horizontal extent of the images was 8.3° and 6.7° of visual angle, respectively. The images were
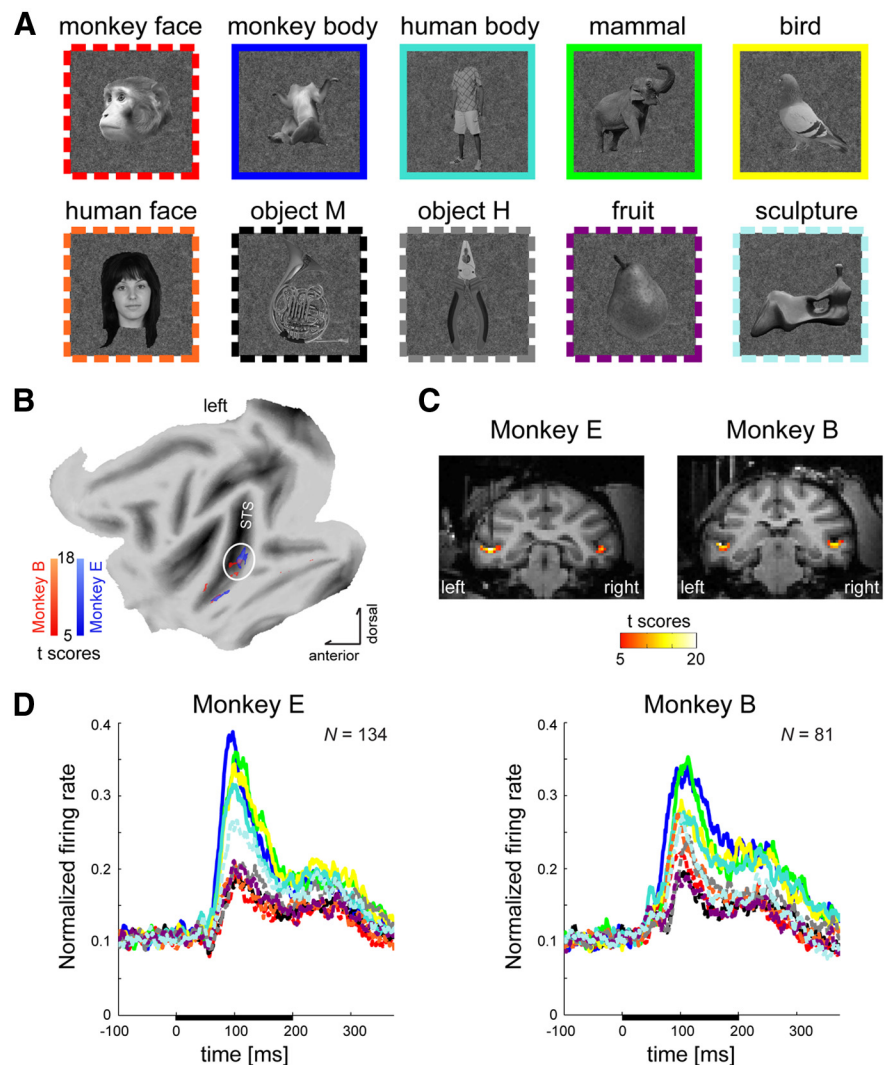
Figure 1. Category selectivity in the fMRI-defined midSTS body patch. **A**, Example image taken from each of the 10 stimulus classes of the main stimulus set, embedded into a pink noise background. **B**, Flattened surface of the left hemisphere of the monkey brain (F99 common space) with the body selective fMRI activations (contrast: monkey bodies − objects M; only activations that passed the FWE corrected level of $p < 0.05$ are shown) of both subjects. The white ellipse indicates the midSTS: the region targeted in the recordings. **C**, MidSTS body patch region targeted in the recordings with activations shown on coronal slices of each monkey (Horsley-Clarke AP, ML coordinates of body patch peak; monkey E, +2, 21; monkey B, +1, 21) An artifact from the electrode in monkey E (left) and tracks from the guide tube in monkey B (right) are clearly visible, targeting the center of the activation. **D**, Normalized population PSTHs (sliding window of 10 ms duration with a step of 1 ms) showing the single-cell responses, averaged across neurons in the midSTS body patch for each of the two monkeys. Before averaging, the PSTH of each neuron was normalized with respect to the maximum firing rate (bin width 10 ms). Line patterns and colors follow the same conventions as the frames in **A**. Stimulus presentation is marked by the thick black line on the abscissa. N indicates number of averaged neurons.

embedded into pink noise backgrounds having the same mean luminance as the images and which filled the entire display (height × width: 30° × 40° of visual angle). Each image was presented on top of nine different backgrounds that varied randomly across stimulus presentations. Although unlikely, the use of different backgrounds may have (slightly) increased response variability. The stimuli were gamma corrected.

Body part stimulus set. A second stimulus set consisted of seven male monkey body part classes, i.e., arm, foot, genitals, hand, leg, tail, and torso. We presented three exemplars of each body part class (Fig. 2A) and each exemplar was shown at five orientations (rotations in the image plane with step size of 45°) including the 180° inversion (Fig. 2B; illustrated for one body part exemplar). Thus, the body part stimulus set included 105 images (3 exemplars × 5 orientations × 7 body part classes). The stimuli were taken from snapshots of movies depicting male
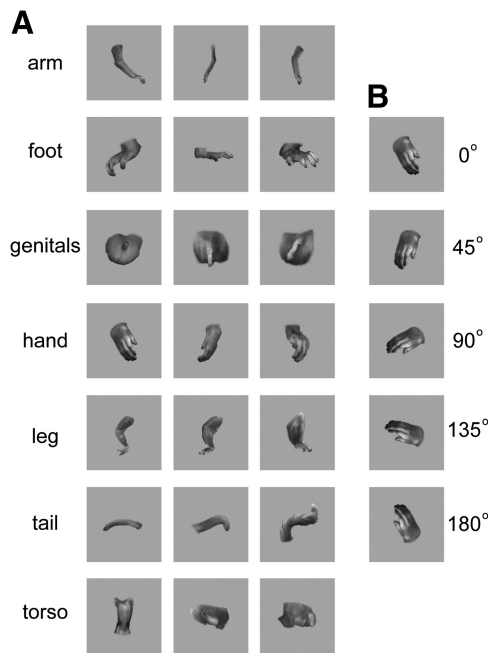
Figure 2. Body part stimulus set. **A**, The three exemplars of each of the seven body-part classes are shown in rows per class. **B**, Five in-plane orientations of one particular exemplar (step size 45°). The body parts measured 4° on a side and were shown on uniform gray background.

monkeys in our colony. The body parts were first resized so that the maximum of their vertical and horizontal extent was 4° at an orientation of 0° (Fig. 2A). Then these images were rotated around their center of mass to obtain the five different orientations (Fig. 2B). The mean luminance of all images was equated. The body parts were presented on top of a uniform gray background, having the same grayscale value as the mean luminance of the body parts. The artificial edges where the body part was dismembered from the rest of the body were blurred and faded into the gray background using Adobe Photoshop CS3. All images were gamma corrected.

fMRI
Details of the fMRI procedure, data analysis and results are provided by Popivanov et al. (2012) and will only be summarized here. The monkeys were scanned while fixating a small red target (0.2° of visual angle) superimposed on the stimuli. During scanning, the monkeys sat in a sphinx position with their heads fixed in a MR compatible plastic monkey chair. Eye position was continuously monitored (120 Hz; Iscan) during scanning. The monkey received a juice reward when maintaining fixation within a square window of 2° × 2° of visual angle. Immediately before scanning, a contrast agent, Monocrystalline Iron Oxide Nanoparticle (MION; Feraheme, AMAG Pharmaceuticals; 8–11 mg/kg) was injected into the monkey femoral/saphenous vein.

In a block design experiment, monkey bodies, monkey faces, objects M, mammals, birds, and fruits/vegetables classes were presented in six discrete blocks of 20 s each. Each class consisted of 20 images of which 10 were identical to those used in the subsequent recordings (main stimulus set). Stimuli were presented for 500 ms each without interstimulus interval (ISI). Each run contained 21 blocks in total: the six classes plus a "fixation" block (fixation target superimposed on the pink noise background) were repeated three times. The monkeys were scanned on a 3T Siemens Trio scanner following standard procedures (Vanduffel et al., 2001). Functional MR images were acquired using a custom-made 8-channel monkey coil (Ekstrom et al., 2008) and a gradient-echo single-shot echo planar imaging sequence (repetition time = 2 s, echo time = 17 ms, flip angle = 75°; 80 × 80 matrix, 40 slices, no gap, 1.25 mm isotropic voxel size). The functional images were coregistered with a

high-resolution (0.4 mm isotropic) anatomical image of each monkey's individual brain, serving as a template.

Only runs in which the animals were fixating the target for at least 90% of the time were included in the analysis. The functional data were resampled to 1 mm isotropic voxel size. All analyses were performed in each monkey's native space without smoothing the functional data. All valid runs (24 and 28 in monkeys E and B, respectively) were combined in a fixed effects model for each monkey separately in native space. They were analyzed using a general linear model with seven regressors, one for each of the six stimulus classes and the fixation condition, plus six additional head-motion regressors (translation and rotation in three dimensions) per run. The resulting t maps were thresholded at $p < 0.05$, family-wise error (FWE) rate, corresponding to $t > 4.9$.

Electrophysiological recordings
Standard single-unit and LFP recordings were performed with epoxylite-insulated tungsten microelectrodes (FHC; *in situ* measured impedance between 1.3 and 1.6 MΩ) using techniques as described previously (Sawamura et al., 2006). Briefly, the electrode was lowered with a Narishige microdrive into the brain using a stainless steel or an MR-compatible (when a position verification scan was performed after recording) guide tube that was fixed in a standard Crist grid positioned within the recording chamber. After amplification and filtering between 540 and 6 kHz, spikes of a single unit were isolated online using a custom amplitude- and time-based discriminator. The simultaneously measured LFPs were filtered on-line using a 1–300 Hz bandpass filter and saved for off-line analysis.

The position of one eye was continuously tracked by means of an infrared video-based tracking system (SR Research EyeLink; sampling rate 1 kHz). Stimuli were displayed on a CRT display (Philips Brilliance 202 P4; 1024 × 768 screen resolution; 75 Hz vertical refresh rate) at a distance of 57 cm from the monkey's eyes. As in all our previous studies, the onset and offset of the stimulus was signaled by means of a photodiode detecting luminance changes in a small square in the corner of the display (but invisible to the animal), placed in the same frame as the stimulus events. A digital signal processing-based computer system developed in-house controlled stimulus presentation, event timing, and juice delivery while sampling the photodiode signal, vertical, and horizontal eye positions, spikes, LFP signals, and behavioral events. Time stamps of the recorded spikes, eye positions, continuous filtered LFP signals (sampling rate 1 kHz), stimulus, and behavioral events were stored for off-line analyses.

The recording grid locations were defined so that the electrode targeted the left midSTS body patch in both animals. Before the recordings started, we performed a structural MRI in each monkey (3T Siemens Trio; magnetization-prepared rapid acquisition with gradient echo sequence; 0.6 mm resolution) and visualized long glass capillaries filled with the MRI opaque copper sulfate ($CuSO_4$) that were inserted into the recording chamber grid (until the dura) at predetermined positions. Then, the functional images (the contrast between the monkey bodies and objects M) of each monkey were coregistered with its anatomical MRI using the coregistration toolbox of SPM8 (Wellcome Department of Cognitive Neurology, London, UK) and the registration was verified by visual examination. Grid positions were selected for body patch recordings if the electrode would end in a voxel that was activated significantly more by monkey bodies than objects M ($p < 0.05$ FWE corrected) and was not significantly activated by monkey faces compared with objects M ($p > 0.05$ FWE corrected). These neighboring voxels included the most significant activation when monkey bodies were contrasted with objects M in the midSTS body patch. During the course of the recordings, we verified the recording locations with 10 and four additional anatomical MRI scans in monkeys E and B, respectively. Four of these scans in monkey E were performed immediately after recording sessions that targeted the body patch, using an MR compatible (fused silica; Plastics One) guide tube with the electrode left in the cortex during the MRI scan (for an example MR image with an electrode *in situ*; Fig. 1C, left). In all other scans we visualized long glass capillaries filled with copper sulfate that were inserted into the grid at recorded grid positions. The recording locations along the medial–lateral and anterior–posterior

dimensions were extrapolated from the trajectories of the imaged capillaries. The validity of the latter method to verify recording locations is supported by the four MRI scans in monkey E in which the electrode was imaged directly and was indeed shown to be present at the predicted location in the anterior–posterior and medial–lateral dimensions. In addition to the imaged capillaries, tracks of the repeated guide penetrations were clearly visible in the MR images of monkey B above the targeted body patch (Fig. 1C, right), providing further evidence that the recordings were at the targeted region. The ventral–dorsal location of the electrode tip was verified in each recording session using the transitions of white and gray matter and the silence marking the sulcus between the banks of the STS.

In addition to the body patch recording locations described in Results, we recorded at 10 and nine neighboring grid positions (1 mm spacing) in the STS of monkeys B and E, respectively, to ensure that we did not miss a body patch containing a very high proportion of body selective neurons. We recorded more extensively lateral to the midSTS body patch location in monkey E only.

Electrophysiology: tasks
Neurons were searched while presenting the 100 images of the main stimulus set in a pseudo-random order. Stimuli were presented for 200 ms each with an ISI of ~400 ms during passive fixation (fixation window size 2° × 2°). The pink noise background was present throughout the task, but refreshed together with the stimulus onset, as in previous studies (Tsao et al., 2006; Issa and DiCarlo, 2012). Fixation was required in a period from 100 ms prestimulus to 200 ms poststimulus. A trial was aborted when the monkey interrupted fixation in this interval. In the pseudo-randomization procedure, all 100 stimuli were presented randomly interleaved in blocks of 100 unaborted trials. Aborted stimulus presentations were repeated within the same block in a subsequent randomly chosen trial. The number of unaborted presentations per stimulus could differ by 1 at most. ISIs within and between successive blocks were the same. Aborted trials were not analyzed further. Juice rewards were given with decreasing intervals (2000–1350 ms) as long as the monkeys maintained fixation. All neurons (N = 185 and 114 for monkeys E and B, respectively) were tested using this procedure and testing was continued when a response was notable in the on-line peristimulus time histograms (PSTHs) for at least one of the stimuli.

Stimuli during the initial search for responsive neurons were presented foveally. When responses to the foveal stimuli were present but weak (as judged by visual inspection of the online PSTHs), the stimulus producing the largest estimated response was selected for receptive field (RF) mapping. For the RF mapping, a scaled version of the selected image (the maximum horizontal or vertical extent was 4°) was presented for 200 ms at 35 positions ranging from 3° ipsilateral to 9° contralateral and from 9° below to 9° above the horizontal meridian. Adjacent positions differed by 3°, horizontally or vertically. The different stimulus positions were presented interleaved. The mean number of unaborted presentations per position was six and five, averaged across the mapped neurons for monkey E and monkey B, respectively. Based on the PSTHs of the RF mapping test, the optimal stimulus location was determined and then the main test was rerun by presenting the stimuli at this location. When two main tests, using different stimulus locations, were available, the one producing the largest response was included in the further analysis. Most responsive neurons searched for with the main stimulus set were also tested with other tests, which are part of another study and will not be reported in this paper.

We recorded the responses of body-patch neurons to the body part stimulus set in a second series of recording sessions that took place after the conclusion of data collection using the main stimulus set. For both monkeys, these neurons were recorded using the grid position that yielded the majority of neurons recorded for the main stimulus set. Procedures were identical to those described above for the main stimulus set; the only exception was that responsive neurons were searched for using the body part stimuli.

Single-unit data analysis
Firing rate was computed for each unaborted stimulus presentation in two analysis windows: a baseline window ranging from 100 to 0 ms before stimulus onset and a response window ranging from 50 to 250 ms after stimulus onset. Responsiveness of each recorded neuron was tested offline by a split-plot ANOVA with repeated measure factor baseline versus response window and between-trial factor stimulus. Only neurons for which either the main effect of the repeated factor or the interaction of the two factors was significant and were recorded for at least five trials per stimulus were analyzed further. Using these criteria for the main stimulus test, 134 of 185 neurons and 81 of 114 neurons were defined as responsive for monkeys E and B, respectively. For this test, the mean number of unaborted presentations per stimulus was 9 for both animals, averaged across responsive neurons. For the body part stimulus set, the mean number of unaborted presentations per body part stimulus was 9.2 pooled across animals (N = 52 neurons; 26 for each monkey). Because our implementation of the split-plot ANOVA required an equal number of observations per cell, we equated the number of unaborted stimulus presentations for that analysis. This was done by removing the last unaborted presentation of the stimulus that was presented by one trial more than the rest. All other analyses included the responses to all unaborted stimulus presentations.

All analyses were based on baseline subtracted, average net firing rate, except stated otherwise. In most analyses, the net firing rates of each neuron to the stimuli were normalized by dividing the firing rate for a particular stimulus by the maximum firing rate of the neuron (the response to the "best" stimulus).

For each neuron we computed several indices. The body selectivity index (BSI) was computed as follows:

$$BSI = \frac{\overline{R}_{body} - \overline{R}_{non\text{-}body}}{|\overline{R}_{body}| + |\overline{R}_{non\text{-}body}|},$$

where $\overline{R}_{body}$ and $\overline{R}_{non\text{-}body}$ are the mean net firing rates to bodies and nonbodies of the main stimulus set, respectively. To compare our results with previous studies in the face patches, we computed the BSI on net firing rates. However, we also computed BSI using raw responses, without baseline subtraction (see Results). In addition, we computed BSIs for which the nonbody category did not include the ambiguous category of the body-like Moore sculptures and BSIs that included the sculptures as nonbodies. The face selectivity index (FSI) was computed, likewise, as the difference between the mean net firing rate to faces and nonfaces divided by the sum of the absolute mean net firing rate to the faces and nonface stimuli. The nonface stimuli included all images without head, i.e., excluding the mammals and birds.

We also computed d' indices (Afraz et al., 2006; Ohayon et al., 2012) which take into account differences in mean responses to the stimulus categories as well as the variability of the responses to the different stimuli within a category. The d' indices were computed for bodies versus nonbodies [d' (body)] and faces versus nonfaces [d' (faces)] using both net and raw responses. Thus:

$$d'(\text{body}) = \frac{\overline{R}_{body} - \overline{R}_{non\text{-}body}}{\sqrt{\dfrac{SD_{body}^2 + SD_{non\text{-}body}^2}{2}}},$$

where $\overline{R}_{body}$ and $\overline{R}_{non\text{-}body}$ are the mean firing rates and SD_{body} and $SD_{non\text{-}body}$ are the SDs of the firing rates for the bodies and nonbodies, respectively. The d' (faces) were defined contrasting the responses to faces and nonface images, excluding the mammals and birds because these had "faces." We tested whether the d' value for each neuron was significantly different from zero by comparing it to the null distribution of d' ($p \leq 0.025$). This null distribution was obtained by computing the d' 1000 times with different permutations of the body and nonbody labels.

Hierarchical cluster analysis with Ward's method was performed on a dissimilarity matrix of pairwise Euclidean distances between the responses to the individual images. The Euclidean distance d_{1-2} for a pair of images 1 and 2 was defined as follows:

$$d_{1-2} = \sqrt{\sum_{i=1}^{n}(R_{1,i} - R_{2,i})^2},$$
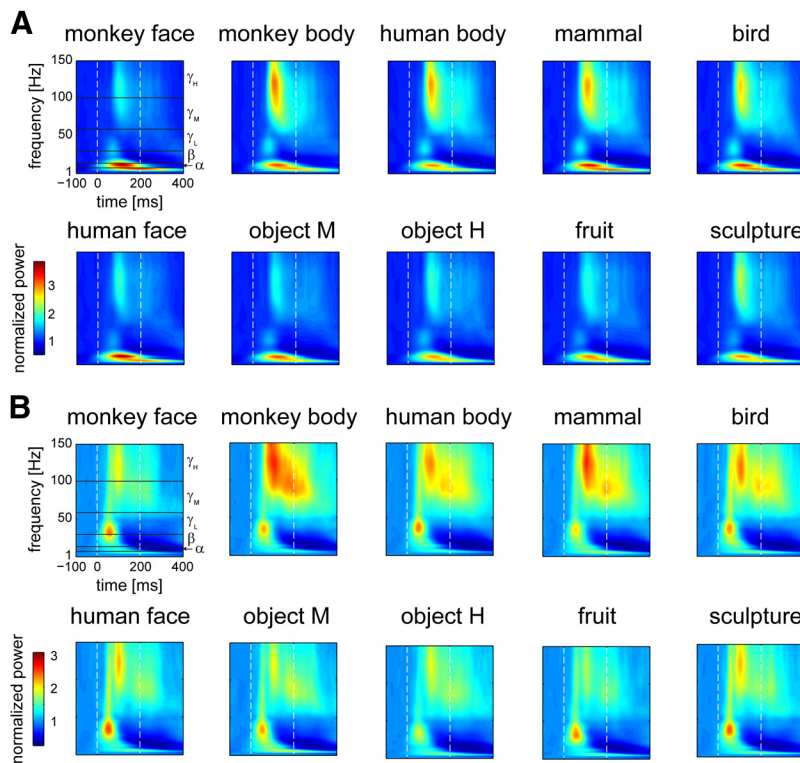
Figure 3. Category selectivity of the LFP power. Time-frequency plots representing the power change normalized to baseline for monkey E (**A**) and monkey B (**B**), averaged across sites (monkey E, $N = 133$; monkey B, $N = 66$) for each of the 10 stimulus classes. The boundaries of the frequency bands (alpha, α; beta, β; low gamma, γ_L; middle gamma, γ_M; high gamma, γ_H) are shown on the first plot in both panels. Stimulus onset and offset are marked by white vertical dotted lines.

LFP data analysis

Previously published procedures were used (De Baene and Vogels, 2010) to analyze the LFPs for the main stimulus set. First, we applied a digital 50 Hz notch filter (fourth-order Butterworth FIR filter; Fieldtrip Toolbox) to remove 50 Hz line contamination. Trials in which the signal exceeded the 0.05–99.95% window of the total amplitude input range (clipping) were excluded from the analyses. Although we recorded LFPs and spikes simultaneously using the same electrode, the number of LFP sites ($N = 133$ and 66 for monkeys E and B, respectively; Fig. 3) is less than those for spiking (Fig. 1D), because we did not have a valid LFP signal during all recording sessions (i.e., the signal was clipped in too many trials). By convolving single-trial data using complex Morlet wavelets and taking the square of the convolution between wavelet and signal, we obtained the time-varying power of the signal for every frequency (Tallon-Baudry and Bertrand, 1999) The complex Morlet wavelets had a constant center frequency–spectral bandwidth ratio (f_0/σ_f) of 7, with f_0 ranging from 1 to 150 Hz in steps of 1 Hz. We took the mean power across trials per spectral frequency and site. The power was normalized by dividing it by the average power in a baseline window that ranged from 100 to 0 ms before stimulus onset. The normalized power was averaged across sites and stimuli per class to generate the time frequency plots of Figure 3. The LFP power response per frequency band was computed by taking the averaged normalized power at each frequency in a 50–250 ms window relative to stimulus onset followed by an averaging across the frequencies of the frequency band of interest. The frequency bands were defined as follows: alpha, 8–12 Hz; beta, 13–29 Hz; low gamma, 30–59 Hz; middle gamma, 60–99 Hz; high gamma, 100–150 Hz (Fig. 3). For quantitative analyses of the mean power for each frequency band across sites (Fig. 4), we equated the contribution of each site to the population response, by dividing the power by the maximum power across the 100 stimuli for each site. Dissimilarity matrices were obtained for each frequency band by computing pairwise Euclidean distances on the percentage change in power from the baseline, normalized by the maximum percentage difference across the stimuli, for each site.

where $R_{1,i}$ and $R_{2,i}$ are the normalized net responses of neuron i, averaged over trials, to stimulus 1 and 2, respectively, and n is the number of neurons tested with that pair of stimuli. Unlike correlation as similarity metric, the Euclidean distance metric takes into account differences in the response patterns of the population of neurons between the images as well as differences in overall response level between the images and hence makes fewer assumptions about the (unknown) metric used by the brain.

To test the extent to which the neural distances reflect the pure image, pixel-based dissimilarities, we computed the pairwise Euclidean distances d_{1-2} between the gray levels of the corresponding pixels of all possible image pairs. This was achieved using the formula above where $R_{1,i}$ and $R_{2,i}$ were the gray values for pixel i in image 1 and 2, respectively. The neural- and pixel-based dissimilarities were compared by correlating the dissimilarity matrices, using the Spearman rank coefficient. To assess whether the obtained value for the coefficient was significant we compared it to a distribution of 10,000 coefficients computed after permuting one of the dissimilarity matrices (threshold $p = 0.02$; two-tailed).

Linear support vector machines (SVMs) were used to classify bodies versus nonbodies, faces versus nonfaces or the 10 image classes of the main stimulus set with the responses of the population of neurons of a single monkey to the individual stimuli as input. In each of the three cases, SVMs were trained using the average net firing rates to seven randomly chosen images of each class and tested using the remaining images. SVMs were trained with cross-validation and a grid search for the regularization parameter to reduce overfitting. The SVM analyses were run using the Weka library (Hall et al., 2009). We trained and tested 100 SVMs, each with a different random sampling of training and test images. The classification rates are averages across the 100 SVMs and test stimuli per SVM. Chance classification rates were determined empirically by running 100 SVMs on the same neural responses but with shuffled stimulus labels. In the case of the 10 class SVMs, the chance classification rates had a mean of 9% (range, 7–11%) and 10% (range, 7–15%) in monkeys E and B, respectively.

Results

The midSTS body patch was defined by contrasting images of headless monkey bodies with control objects (Popivanov et al. (2012); Fig. 1A,B). The recording locations in each monkey were guided by their individual fMRI data. We recorded at the location showing the most significant activation (Fig. 1C) and at neighboring locations. In monkey E, four neighboring grid positions (1 mm spacing; along the posterior-anterior dimension) coincided with the portion of the midSTS body patch that was activated strongest by bodies compared with the control objects. There was no significant fMRI activation to faces (contrast monkey faces −object M) at these locations. In monkey B, three neighboring grid locations were probed that corresponded to the most significant activations of his midSTS body patch. As in the other monkey, there was no significant activation to faces at these voxels. In both monkeys, we recorded from their left hemisphere only.

Recording locations were verified using anatomical MRI scans between recording sessions in both animals (see Materials and Methods) and by direct visualization of the electrode *in situ* after
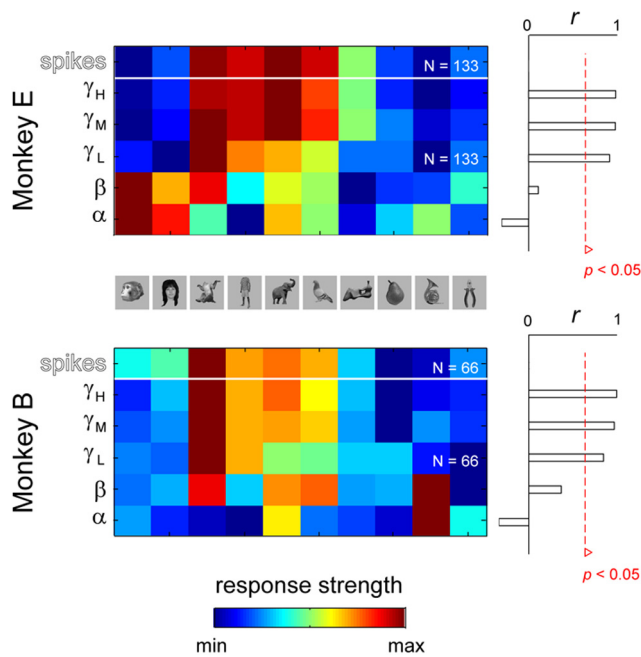
Figure 4. Correlation between the category selectivity of spiking activity and LFP power in different frequency bands. Left, Normalized spiking activity (spikes) averaged per class of those units for which valid LFP measurements were obtained and normalized LFP power averaged for the same sites. The mean LFP power is averaged within each of five frequency bands. Colors indicate the normalized response strength (see bottom color bar). Each class is indicated by an example image. Right, Pearson correlation coefficient r between the mean spiking response and the mean power for each frequency band. The red line depicts the statistical threshold above which the correlations are significantly different from 0 ($p < 0.05$; $N = 10$ classes).

actual recordings in monkey E (four scans; Fig. 1C, left, example; see Materials and Methods). In addition, guide tube tracks were clearly visible on the MRI of monkey B at positions consistent with the targeted recording location (Fig. 1C, right). These MRI scans indicate that we recorded at the targeted location in the medial–lateral and anterior–posterior dimensions.

Category selectivity of the midSTS body patch
We measured single units and LFPs, simultaneously, for the 100 images of the main stimulus set that were presented randomly interleaved for 200 ms each during passive fixation (see Materials and Methods). These images were half of the stimuli used in the fMRI study of Popivanov et al. (2012). There were 10 images in each of the 10 stimulus classes: monkey faces, human faces, headless monkey bodies, headless human bodies, mammals, birds, body-like sculptures, fruits/vegetables, and two sets of control objects matched in low-level stimulus properties to the monkey bodies (objects M) and the human bodies (objects H), respectively (Fig. 1A, examples). Responsive neurons were searched when presenting the 100 images, centered at the fovea. In those cases in which responses were weak, we mapped the receptive field with the image that elicited the strongest response foveally, and then retested the neuron by presenting the 100 images at the center of the receptive field. Of the 134 responsive neurons (for definition, see Materials and Methods) recorded at the category-selective body patch of monkey E, 35 (26%) neurons were tested at peripheral locations (average eccentricity, 5.5°). In monkey B, only 10% (8/81) were tested at peripheral locations (average eccentricity, 3.7°). Below, we will report only data on responsive neurons for the optimal location or for the foveal location when no mapping was obtained, ensuring that each neuron is contributing only once to the

sample. Results were similar when restricting the sample only to the foveal presentations. All responsive neurons showed an excitatory response for at least one image and often showed inhibitory responses to some images, which is typical for inferior temporal cortex.

The single-unit responses, averaged across the images of a class, differed significantly across classes (repeated-measures ANOVA on normalized responses per neuron; $p < 0.0001$ in each animal). For each monkey, the average response was greater for the four classes that contained bodies (monkey and human bodies, mammals, and birds) compared with the other classes, including the body like sculptures (Fig. 1D). Thus, we defined the body category as consisting of images of these four classes. The nonbody category included all the other classes: the monkey and human faces, objects M, objects H, fruits/vegetables, and sculptures. In some of the analyses, explicitly mentioned below, we excluded the sculptures from the nonbody category, because of their body-like appearance. In both monkeys, the average normalized response to the body category was significantly larger than the average response to the nonbody images (paired t test; $p < 0.00001$ in each animal). This preference for the body category was present early on in the response of monkey E but it was more pronounced in the later phase of the response in monkey B. Responses were stronger for bodies compared with either human or monkey faces. This difference was highly significant for each of the four body classes in monkey E (*post hoc* Bonferroni t tests; each body class, each face class; $p < 0.00001$); however, it reached significance only for the monkey bodies (*post hoc* Bonferroni t tests; $p < 0.01$) but not for the other body classes (all $p > 0.48$) in monkey B. Interestingly, the monkey bodies elicited a larger response than the human bodies in each animal, but this difference failed to reach significance. The monkey bodies produced a significantly larger response compared with the objects M class (*post hoc* Bonferroni t tests; $p < 0.00001$ in each animal), in agreement with the fMRI contrast that was used to define the recording location.

The single-unit data represent a relatively small sample of the population of neurons in the targeted body-patch regions. Therefore, we also simultaneously measured LFPs (using the same electrode) and computed the power as a function of peristimulus time and spectral frequency. It has been suggested that the power for frequencies >50 Hz can be used as a proxy for the spiking activity of the population of neurons close to the electrode (Ray and Maunsell, 2011). As shown in Figure 3, the LFP power for those frequencies was strongly selective for stimulus class in both monkeys, with greater power for the four body classes, which aligns perfectly with the single-unit data. We quantified the mean body category selectivity of the LFP signal by comparing the average normalized power for each of five spectral frequency bands (see Materials and Methods and Fig. 3 for definitions of frequency bands) for bodies and nonbodies (excluding sculptures). In each animal, the mean normalized power was significantly larger (paired t test) for bodies compared with nonbodies for each of the gamma bands (Table 1). The same trend was present for the beta band, but the body category selectivity became much weaker than for the gamma bands and reached significance in one animal only. The alpha bands showed a stronger mean response to nonbodies compared with bodies, but as for the beta bands, the difference between the two categories was relatively small (Table 1). For comparison, Table 1 also shows the mean normalized spiking activity of single units recorded at the same sites as the LFPs.

Table 1. Mean normalized responses to bodies and nonbodies for spiking activity and LFP power

	Monkey E (N = 133)			Monkey B (N = 66)		
	Body	Nonbody	p	Body	Nonbody	p
Spiking	0.14	0.06	0.0001	0.14	0.08	0.0001
High gamma	0.50	0.40	0.0001	0.49	0.39	0.0001
Middle gamma	0.53	0.42	0.0001	0.47	0.40	0.0001
Low gamma	0.56	0.50	0.0001	0.49	0.46	0.0001
Beta	0.51	0.50	0.09	0.51	0.49	0.003
Alpha	0.52	0.53	0.002	0.49	0.51	0.03

p denotes the significance value of the paired t test between the responses to bodies and nonbodies; N represents the number of analyzed neurons and sites.

Figure 4 directly compares the normalized single unit spiking activity with the LFP power in different spectral bands for each of the 10 stimulus classes for the population of the same recording sites. In both monkeys, the Pearson correlations between the mean spiking activity and mean power in the gamma bands were all >0.92 ($p < 0.0002$; $N = 10$ classes), except for a correlation of 0.82 ($p < 0.005$) between the low-gamma band power and spiking activity in monkey B. However, no significant correlations between spiking activity and power were present for the beta and alpha bands (Fig. 4, right), despite a significant, yet completely different, class-specific modulation of the LFP power in these lower frequency bands (ANOVA; $p < 0.001$ for each animal and band). This pattern of the correlations between spiking activity and the LFP power across frequency bands (Fig. 4) fits the presence of significant body category selectivity for both spiking activity and gamma power and the weaker or even reversed selectivity for the lower-frequency bands (Table 1).

Category selectivity of single units

The above data indicate that the mean neuronal activity in the fMRI defined body patch is greater for bodies compared with other stimulus classes, including faces. However, based on these neuronal population analyses one cannot conclude that there is body category selectivity at the single unit level. In other words, does each of the neurons within the body patch prefer body images above images of other classes? Or is there a small pool of highly selective body cells embedded within a pool of noncategory selective cells? Alternatively, are there many weakly selective body cells that drive the population response?

To assess this, we computed for each single neuron a BSI that contrasts the mean net responses to body and nonbody images (Materials and Methods; Fig. 5A). A BSI larger than zero shows a preference for bodies with a BSI of 0.33 corresponding to a twofold greater net response to bodies compared nonbodies. In a first conservative analysis, we excluded the body-like sculptures from the nonbody category. The median BSI with only monkey and human faces, fruits/vegetables, objects M, and objects H as nonbody classes was 0.47 and 0.25 for monkeys E and B, respectively; values significantly >0 (Wilcoxon test; $p < 0.00005$ in each animal; median across animals, 0.38; mean, 0.33). However, as is clear from the distribution of the BSI (Fig. 5A; Table 2), both monkeys showed a considerable variation in the magnitude of the BSI. Previous studies on face selectivity, using the same sort of index computed on net responses, used a criterion of 0.33 to define a face category selective cell (Tsao et al., 2006; Issa and DiCarlo, 2012). Employing the same criterion, 61 and 48% of the neurons can be classified as body-selective in monkeys E and B, respectively (53% across both animals). When the BSI was recomputed with the sculptures included as nonbodies the median
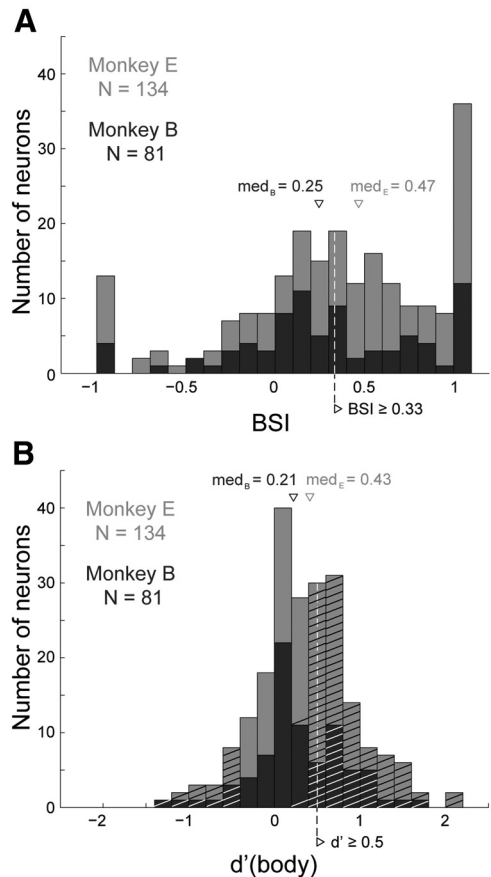
Figure 5. Body category selectivity in the midSTS body patch. ***A***, Distribution of the BSI of body-patch single neurons. The values for the two animals are indicated by different gray levels (darker corresponds to monkey B). The triangles indicate the median BSI of each monkey. The dotted line marks a BSI of 0.33, corresponding to a twofold greater net response to bodies with respect to nonbodies. BSI was computed using net responses. ***B***, Distribution of d′ (body) of the same neurons. The same conventions as in ***A***, except that the dotted line represents a threshold of $d' = 0.5$. The neurons with d′ values significantly different from 0 ($p ≤ 0.025$, Permutation test) are hatched.

Table 2. BSI and d′ (body) values computed using net and raw firing rates

	BSI (net)	BSI (raw)	d′ (net)	d′ (raw)
Monkey E (N = 134)				
Median	0.47	0.25	0.43	0.44
P$_{25}$	0.08	0.05	0.10	0.11
P$_{75}$	0.86	0.42	0.69	0.72
% ≥ Thr	60%	35%	43%	46%
Monkey B (N = 81)				
Median	0.25	0.14	0.21	0.21
P$_{25}$	0.03	0.01	0.03	0.02
P$_{75}$	0.72	0.38	0.68	0.74
% ≥ Thr	43%	30%	36%	35%

For each index the median and the 25 th and the 75 th percentiles (P$_{25}$ and P$_{75}$) are shown. The last row (% ≥ Thr) represents the percentage of neurons having an index greater than or equal to the respective threshold (0.33 for the BSI and 0.5 for d′). N represents the number of analyzed neurons.

BSIs were similar (0.41 and 0.26 in monkeys E and B, respectively) with 56 and 47% of the neurons classified as body cells.

The body patch was defined by comparing fMRI activations for the monkey bodies and objects M. Computing a BSI index with only the net responses to these two classes yielded median BSIs of 0.45 and 0.42 in monkeys E and B, respectively. Based on these BSIs, 55 and 54% of the neurons were "monkey body" selective in E and B, respectively. Thus, using the same contrasts

for single units and fMRI, approximately half of the neurons found in the body patch could be classified as body cells using the conventional criterion and category index.

To compare with previous fMRI-guided studies on face selectivity (Tsao et al., 2006; Freiwald et al., 2009; Freiwald and Tsao, 2010; Issa and DiCarlo, 2012), we computed the BSIs on net responses. Because such BSIs can be affected (in both directions) by strong inhibitions to a few stimuli of a class, we recomputed BSIs using raw responses, i.e., including baseline and ignoring the distinction between inhibitory and excitatory responses. As shown in Table 2, median BSIs computed on raw responses were, as expected, smaller than those computed on net responses but were still significantly larger than zero in each animal (Wilcoxon test; $p < 0.00005$ in each animal; median across animals, 0.19) with 35% and 30% of the neurons having a BSI larger than 0.33 in monkeys E and B, respectively.

In addition to BSI we also computed another category selectivity index, d' (body), which takes into account the mean responses to the contrasting categories as well as the variability of the responses to the individual images of a category (see Materials and Methods). The median d' (body), computed using net responses, was 0.43 and 0.21 in monkeys E and B, respectively (Fig. 5C; median d' across animals, 0.35; d's computed on raw response produced similar results; Table 2). As expected, the distribution of the d' (body) was significantly biased toward positive values (Wilcoxon test; $p < 0.00005$ in each animal). Assessing the statistical significance of the d' (body) for each neuron by a permutation test showed that d's larger than 0.5 (or smaller than −0.5) were statistically significant. Taking the criterion of 0.5 (which happens to be the same one used by Ohayon et al. (2012) who also used d' in their face selectivity study) to define body category selectivity, between 35 and 46% of the neurons (depending on the animal and on whether one computes d' on raw or net responses; Table 2) showed a significant body selectivity. Also, pooled across monkeys, 8% of the body-patch neurons showed a d' (body) significantly smaller than −0.5 (Fig. 5B), indicating a significant category selectivity for nonbodies.

Thus, using several category selectivity metrics, we can conclude that although the midSTS body patch shows body category selectivity at the population level, the single neurons that comprise this population differ greatly in their degree of body category selectivity. Also, independent of the used category selectivity metric, body category selectivity and the percentage of body category selective neurons are lower than that reported for face category selectivity in the face patches (Tsao et al., 2006; Issa and DiCarlo, 2012; Ohayon et al., 2012).
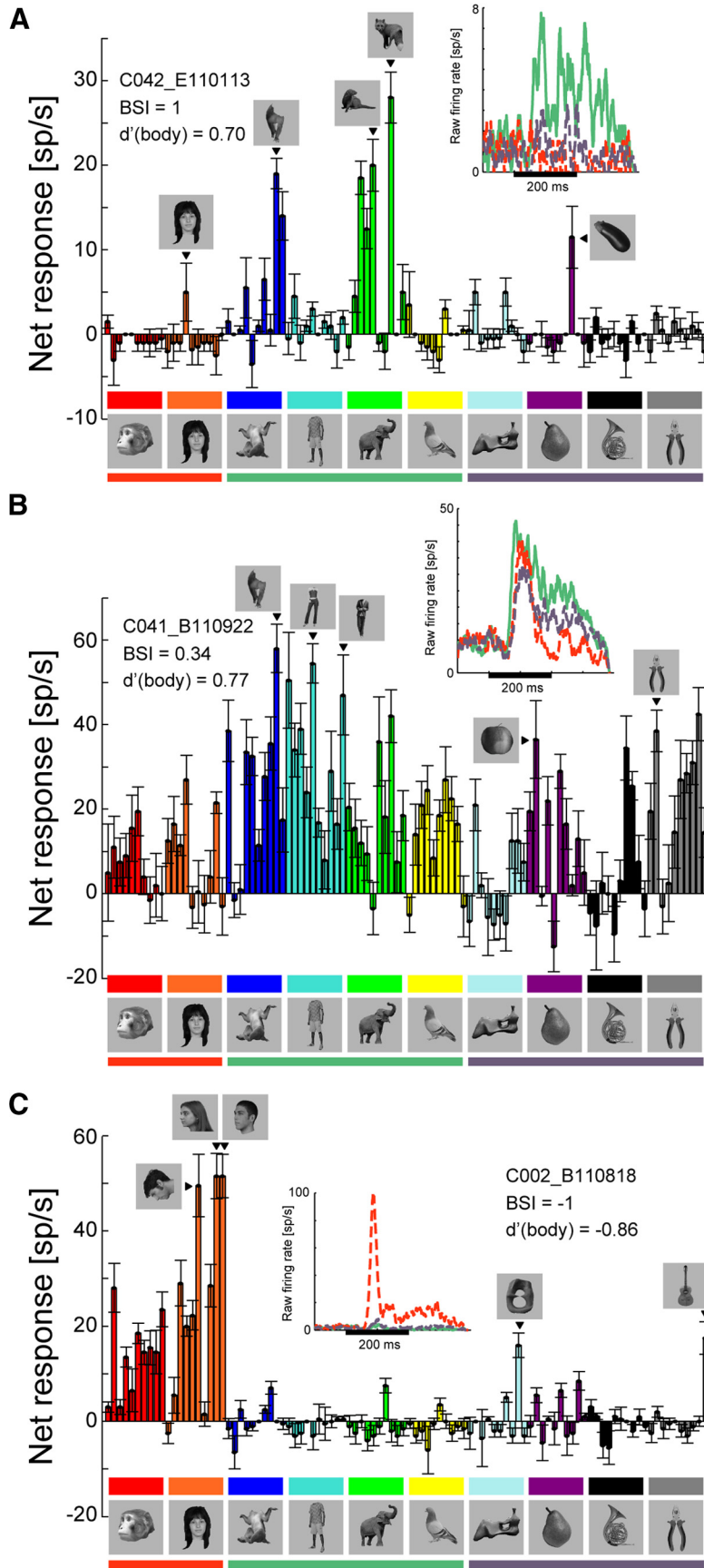
Stimulus selectivity of single units

Figure 6 shows three single neuron examples, whereas the stimulus selectivity of all recorded body-patch neurons is shown in Figure 7A. Both figures illustrate the variation in category and stimulus selectivity that was manifest in our sample of body-patch neurons. Most neurons responded to many stimuli of different classes, including nonbodies (Figs. 6B, 7A). In fact, some neurons showed on average stronger responses to faces compared with nonfaces, when computing a conventional FSI (Tsao et al., 2006) or a d' (face), which contrasts the mean responses to faces versus the other images (except for animals and birds because these images depicted heads as well). The FSI was >0.33 for 16 and 21% of the body-patch neurons recorded in monkeys E and B, respectively (a twofold greater average response to faces compared with the other stimuli) and 8 and 19% of the neurons in monkeys E and B, respectively, had a d' (face) larger than 0.5 (the

same criterion as Ohayon et al., 2012). A cell with a FSI of 0.98 and a d' (face) of 1.66 is illustrated in Figure 6C. These face category selective neurons were intermingled with body-selective neurons within single penetrations. To demonstrate this, we selected neurons that had a FSI >0.33, a d' (face) > 0.5 and a twofold stronger response to faces compared with bodies. In the nine penetrations in which there was at least one recorded neuron below the selected face-selective neuron, the median FSI and d' (face) for the face selective neurons was 0.72 and 1.18, respectively, and reversed to a median FSI of −0.44 and a d' (face) of −0.33 for the neighboring neuron. The median BSI and d' (body) of the face selective neurons was −0.50 and −0.57, respectively, which increased significantly for the neighboring neurons (median BSI, 0.73; median d' (body), 0.31; Mann–Whitney U test; $p < 0.05$). The same reversals of the FSI, d' (face), BSI, and d' (body) were present for the 10 penetrations for which there was a neuron recorded above the face selective one (median FSI, 0.86 vs −0.06; median d' (face), 1.20 vs −0.04; median BSI, −0.48 vs 0.23; median d' (body), −0.57 vs 0.14; Mann–Whitney U test; $p < 0.05$). Importantly, this also held for the five penetrations in which a face selective neuron was recorded in between two recorded neurons (median test on BSI and d' (body), $p < 0.05$), showing that face category selective neurons were mixed with neurons with other stimulus preferences in this body patch.

Figures 6 and 7A illustrate that the neurons in the body patch responded to only some exemplars of a class. For example, the neuron shown in Figure 6A responded to a minority of the body images. This explains its relatively low d', despite the BSI of 1 (this neuron also showed excitatory responses to a couple of nonbodies, but these were compensated by the negative net responses for many nonbody images; its BSI computed on raw responses was 0.63). The marked within-class selectivity was examined for our population of neurons by ranking the images of a class according to the elicited net response of each neuron that responded significantly to at least one of the images of that class. The statistical significance was assessed by a split-plot ANOVA (stimulus as between-trial factor and baseline versus stimulus period as repeated, within trial factor) for each of the 10 classes and an excitatory net response to at least one image of the class was required. The image ranking was performed with the mean responses obtained in 50% of the trials and the responses of the other 50% of the trials were then plotted as a function of the image rank. This avoided an erroneous induction of stimulus selectivity by the ranking procedure. This ranking analysis yielded evidence of strong within-class selectivity for all classes in both monkeys (illustrated for six classes in Fig. 7B). In fact, the net normalized response to the "worst" image of each class (Fig. 7B, rank 10) was not significantly larger than zero for each of the 10 stimulus classes in each animal (Bonferroni corrected Wilcoxon signed rank test; $p > 0.05$). Even for this relatively small number of images (10) per class, the single-unit responses varied within a large range, being absent for some images of the class. A highly similar within-class selectivity was also observed when ranking the stimuli only for those neurons and classes that demonstrated class-selectivity (a twofold stronger response to the class compared with controls) or only for the class that included the preferred stimulus (among the 100 stimuli tested) of a neuron. The strong within-class selectivity was not due to differences in stimulus area, contrast, or aspect ratio. Indeed, the mean normalized responses did not depend on the differences in these stimulus parameters between the preferred image of a neuron and the other images of a class (data not shown).
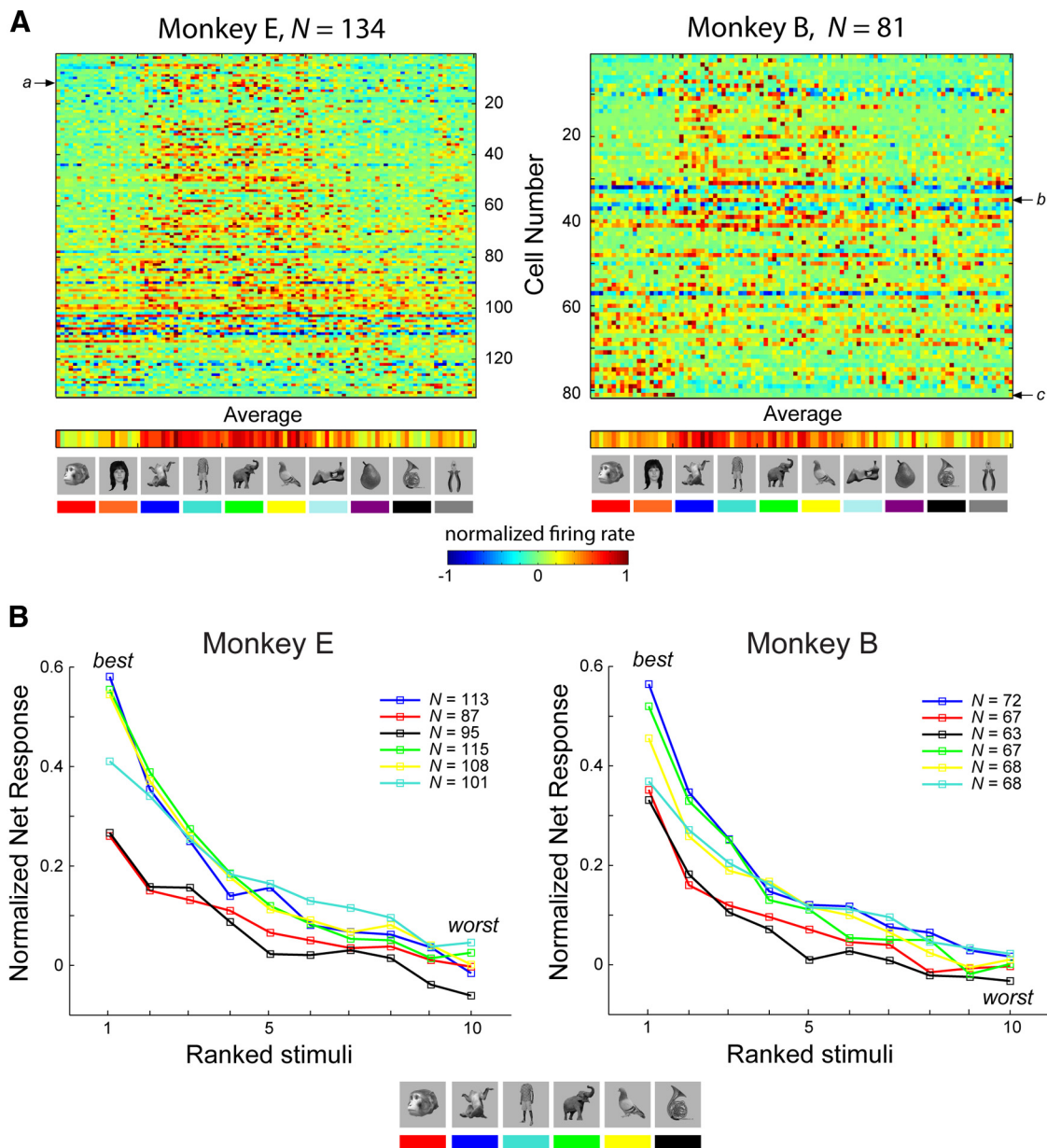
Figure 7. Stimulus selectivity of midSTS body-patch neurons. ***A***, Spiking activity matrices where each row represents the normalized responses of a neuron to each of the 100 stimuli. Each column corresponds to a stimulus. The stimuli are grouped per class as indicated by the example images. Normalized response strength is indicated by a color code (see colored bar). Cells are ordered by their BSI, i.e., the cells on top of the matrix are the most body-selective. The mean normalized responses, averaged across neurons, for each stimulus are shown below the matrices ("average"). The horizontal arrows denote the example neurons and the letters corresponding to the panels of Figure 6. ***B***, Within-category selectivity. Mean normalized responses to the 10 stimuli within a class, ranked by response strength from best (rank 1) to worst (rank 10). For each neuron, half of the trials were used for ranking the stimuli and the other half of the trials were used to average the responses. Normalization was performed with respect to the maximum responses across all 100 stimuli tested. Only neurons for which there was a significant response to at least one of the stimuli in a given class were included for that class, explaining the different number of neurons (N) for the classes. The ranking was performed for the six classes indicated by color coded legend.

neural response-based dissimilarities between all possible image pairs. As a metric of neural-based stimulus dissimilarity we used the Euclidean distance between the images in a multidimensional space where the responses of the single neurons defined the dimensions (Op de Beeck et al., 2001;De Baene et al., 2007;Kayaert et al., 2005; see Materials and Methods). Figure 8B shows the Euclidean distance for all possible stimulus pairs for the neurons of both monkeys combined. It is obvious that the dissimilarities are large for pairs of body images (mean Euclidian distance, 5.32; SEM = 0.02). In particular, this was the case for monkey bodies (mean distance for pairs of monkey bodies, 5.48; SEM = 0.09),

mammals (5.32; SEM = 0.06) and birds (5.07; SEM = 0.09), but pairs of human bodies showed lower dissimilarities (mean distance, 4.54; SEM = 0.06). This may reflect the fact that all human body images, except one, depicted an upright standing person, and thus showed less variation in posture than the other body classes. The mean dissimilarities were the smallest for pairs of faces (mean dissimilarity, 4.13; SEM = 0.03; human and monkey faces combined) followed by pairs of inanimate objects (4.42; SEM = 0.02). The dissimilarities for face versus bodies (5.36; SEM = 0.01) or inanimate objects versus bodies (5.19; SEM = 0.01) were large but comparable to those for the body pairs.
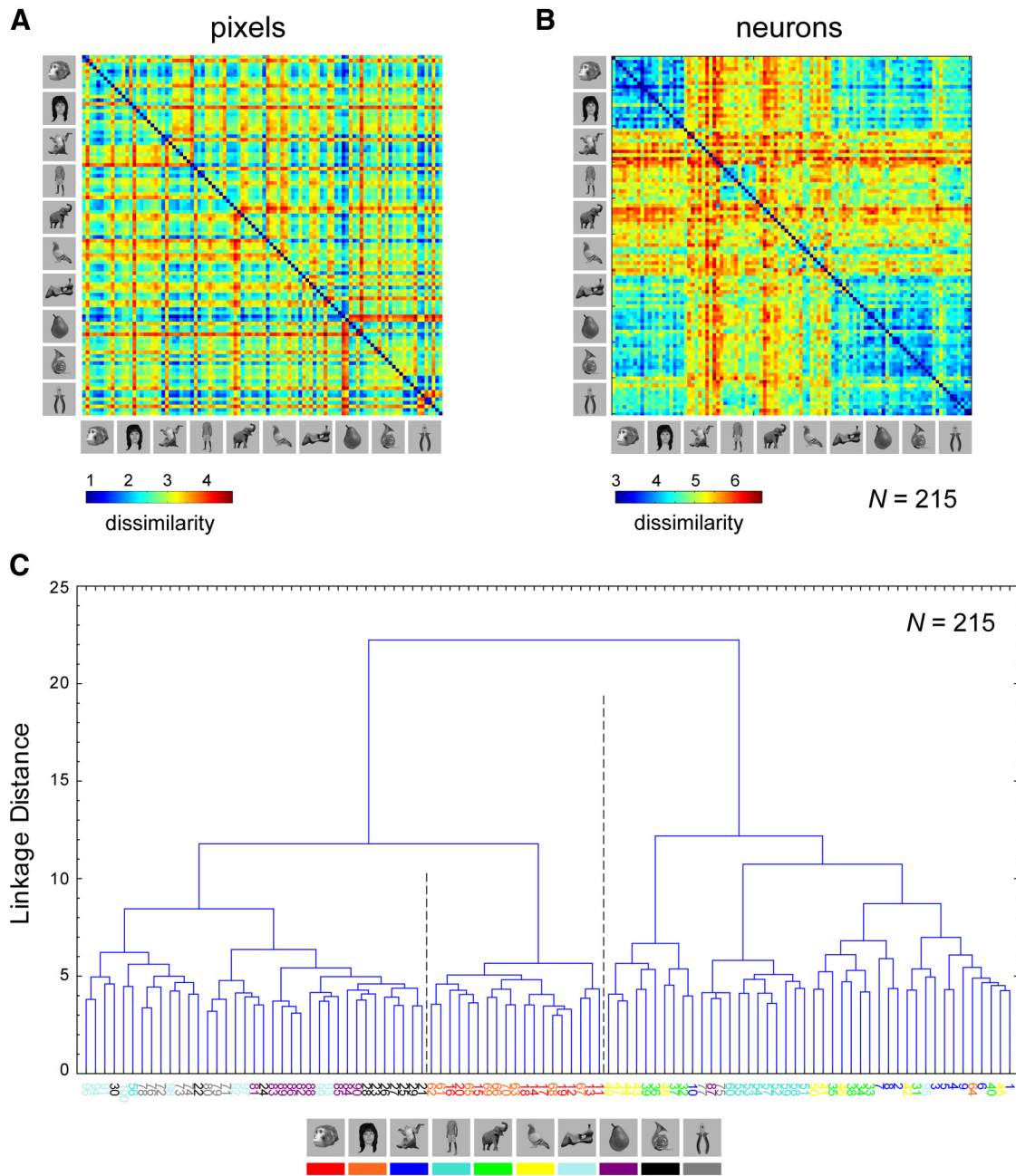
Figure 8. Representation of animate and inanimate object exemplars in the midSTS body patch. *A*, Dissimilarity matrix of pairwise Euclidean distances for all stimulus pairs based on the gray level difference in the corresponding pixels. The stimuli are grouped per class as indicated by the example images. The dissimilarity is indicated by the color scale with red indicating high dissimilarities. Note that the matrix is symmetric with respect to the zero distance diagonal. *B*, Dissimilarity matrix of pairwise Euclidean distances for all stimulus pairs based on the spiking responses of 215 body-patch neurons. The same conventions as in *A*. *C*, Hierarchical clustering of the stimuli based on the spiking responses of both monkeys ($N = 215$ neurons). Vertical stippled lines indicate the boundaries of the major clusters. The stimulus classes are color-coded according to the legend (bottom).

To determine whether these neural dissimilarities merely reflect physical image dissimilarities, we computed the Euclidean distances between the images in the multidimensional space defined by the pixel gray levels, i.e., the input to the visual system (Fig. 8A). Comparing the two dissimilarity matrices (Fig. 8A,B), it is clear that the pixel-based dissimilarities are quite different from the neural dissimilarities. Indeed, the Spearman rank correlation between the two matrices was very small, $r_S = 0.03$ (n.s., permutation test), indicating the neural dissimilarities do not simply reflect physical image similarities.

We examined the stimulus representation of the body-patch neurons further by performing a hierarchical cluster analysis of

the dissimilarity matrix of Figure 8B. The advantage of this technique, compared with direct testing of the similarities in response patterns between bodies and the other classes, is that it provides an unbiased description of the similarities among the responses to the stimuli, irrespectively of their class. The cluster analysis of the spiking data when both animals were combined showed two main clusters (Fig. 8C). One cluster contained 44 images of which 39 (97.5%) were bodies (10 monkey bodies, 9 human bodies, 10 mammals, and 10 birds). The percentage of bodies in this cluster (97.5%) was significantly higher than the 40% expected if bodies were randomly distributed between the two clusters (binomial test; $p < 0.01$). This "body" cluster contained a subcluster con-

sisting of the nine human bodies, two elongated objects (objects H), and one vegetable with the same aspect ratio (a vertically oriented corn). The body cluster also contained one human face, which is shown in Figure 1A. Interestingly, unlike the other human face stimuli, this person had long hair, which might appear as two limbs below the neck. The other, nonbody, cluster diverged into two distinct clusters with one consisting entirely of faces. Note that human and monkey faces, despite the morphological differences between species, were dispersed within this face cluster. A similar clustering of bodies versus other images was also present in the individual data of each animal, but noisier than when pooling the data across both animals (percentage of bodies in body cluster, monkey E: 71% with 34/40 bodies in body cluster; monkey B: 65% with 40/40 bodies in body cluster).

In summary, the pairs of body images showed high within-class dissimilarities based on spiking responses of single units (Fig. 8B), which is consistent with the high within-class selectivity of the single neurons (Fig. 7B). Despite this high selectivity among the different body exemplars, the population of body-patch neurons clustered bodies versus nonbodies. This resulted from a combination of the greater responses for bodies compared with nonbodies (Fig. 1D), the relatively low dissimilarities for the nonbody image pairs, which evoked a weaker response on average, and the relatively high dissimilarities for the body–nonbody image pairs.

Because LFPs sum the activity of a population of neurons and assuming that neighboring neurons can have different preferences within the body class but still tend to prefer bodies over other stimulus classes, one would expect that the mean dissimilarities for pairs of body images would be smaller than the dissimilarities for body–nonbody pairs for the LFP power. This was indeed the case for the high and middle gamma power (Fig. 9): the mean dissimilarity for the body pairs was 4.53 (SEM = 0.01) and 4.52 (SEM = 0.01) for the high and middle gamma power, respectively, which was lower than the mean dissimilarities for the face-body (4.86; SEM = 0.02 and 4.97; SEM = 0.02) and for body-inanimate object pairs (4.68; SEM = 0.01 and 4.70; SEM = 0.01). As expected, cluster analysis showed for both these gamma bands a cluster predominantly containing bodies (39/47, 97.5%; $p < 0.01$ and 37/42, 92.5%; $p < 0.01$ for the high and middle gamma, respectively). For the low gamma power, the distinction between bodies and nonbodies was less (mean dissimilarity for body, body-face, and body-inanimate object pairs was 5.36, 5.47, and 5.36, respectively; Fig. 9) and the cluster analysis showed a cluster containing bodies (22/22), but only 55% of the bodies were represented in that cluster. The clustering of bodies versus nonbodies was weak (29/46, 72.5%; $p < 0.01$) and absent (21/44, 52.5%; n.s.) for the alpha and beta bands, respectively.

Classification of bodies versus nonbodies

Because the cluster analysis of the spiking activity showed distinct clusters of the bodies versus other stimulus classes, it was expected that one could determine whether a stimulus is a body from the population response of the midSTS body-patch neurons. This prediction was tested by having a classifier decide whether an image contains a body, or not, given only the population response vector for that stimulus. This population response vector consisted of a concatenation of the mean responses (averaged across trials) of the neurons to that stimulus. For both monkeys, separately, we trained SVMs on 70% of the stimuli of each category and tested classification performance on the remaining 30%. Hence, we tested explicitly for generalization, a hallmark of categorization. In both monkeys, the proportion of

correct classifications for bodies versus nonbodies was high, and well above chance level (50%; confirmed by permuting the stimulus labels): 90 and 89% correct in monkeys E and B, respectively. Interestingly, when having the classifier deciding whether a face or nonface (excluding mammals and birds, which had heads) was present, the classification scores were also high: 92 and 97% correct in E and B, respectively. Thus, the population of body-patch neurons can be used to classify bodies versus nonbodies and, also, faces versus nonfaces. This led to the question of whether other stimulus classes can also be classified with the body-patch population responses. To answer this question we trained SVMs to classify an image as belonging to one of the 10 stimulus classes. The confusion matrices (Fig. 10) show that correct classification scores for all of the 10 classes were well above chance (Fig. 10, diagonal; chance level = 10% correct). Interestingly, objects M and objects H are both object sets, mainly differing in aspect ratio, and the body-patch neurons could classify these rather well with little confusion between these two classes (Fig. 10). The different classes of the body category (monkey bodies, human bodies, mammals, and birds) could also be distinguished reliably, except for the confusion of animals and birds in monkey E (Fig. 10).

Location specificity of stimulus representation in STS

The body-selective location was bordered laterally and medially by a region in which the activation for monkey bodies was still stronger compared with the objects M class, but that also showed stronger activation for monkey faces compared with objects M (Fig. 11A). It was interesting to assess whether the single unit and LFP selectivity would change away from the body patch. Thus, in monkey E, we recorded also single units and LFPs 1 and 3 mm lateral to the primary target location (Fig. 11A, position 1). The class selectivity of the mean single unit and high gamma LFP power changed when moving more lateral: mean responses to the headless monkey and human bodies became weaker and the responses to faces increased (Fig. 11B). This difference among recording locations in class selectivity was highly significant (ANOVA; interaction stimulus class and recording position, $p < 0.0005$ for spikes and high gamma power). A cluster analysis of the most lateral position (Fig. 11A, position 3) drew a distinction between a face cluster, consisting of 17 of the 20 faces, and all other images, including bodies (Fig. 11C). This contrasts with position 1, which showed a body cluster that included 34/40 bodies (percentage of body images in cluster, 71%) and was separated from a cluster of faces and the other objects (Fig. 11C). Thus, moving laterally away from the body patch, there is a gradual transition from a representation of mainly bodies to one of faces.

Selectivity for body parts in the midSTS body patch

One could argue that the relatively low BSI and d' (body) in the midSTS body patch and the strong within-body class selectivity results from a tuning to individual body parts rather than to the whole body. Indeed, it is possible that different body-patch neurons are selective for different body parts, i.e., some neurons preferring a hand, other neurons a leg and still others a torso, etc. Because some body parts were partially occluded in some of the body images of our main stimulus set this could have contributed to the strong within-body category selectivity that we observed with this stimulus set. To examine this question, we measured the responses of neurons in this body patch to segmented body parts in a control experiment. The body part stimulus set (Fig. 2) consisted of seven classes of male monkey body parts from which three exemplars were presented at five orientations each. The midSTS body-patch neurons responded well to these body part
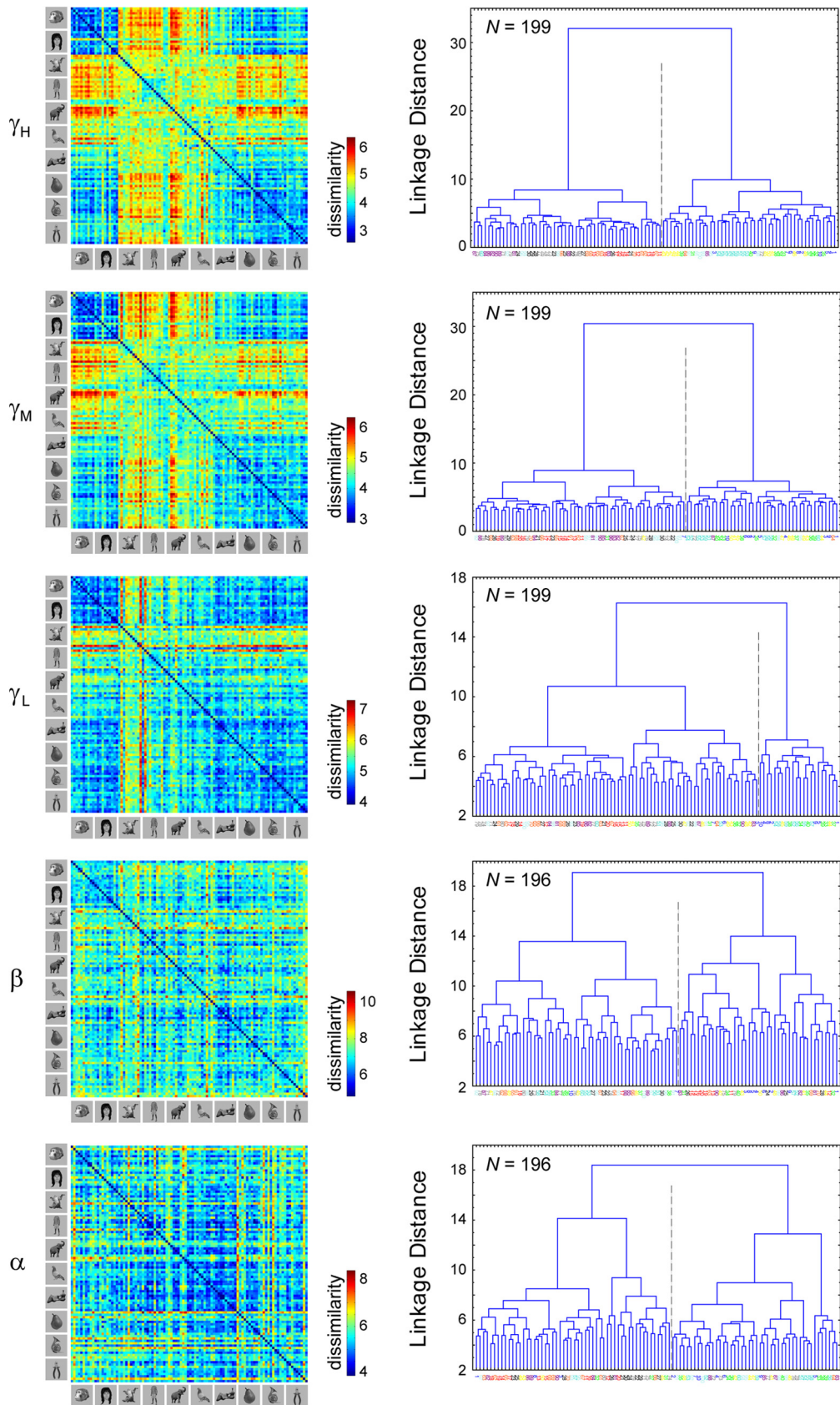
Figure 9. Representation of animate and inanimate object exemplars based on the LFP power in different frequency bands. Left, The dissimilarity matrices of pairwise Euclidean distances for all stimuli pairs, based on the LFP power in each frequency band (same conventions as in Fig. 8B). Right, The hierarchical clustering of all stimuli (same conventions as in Fig. 8C). The sites from both monkeys are pooled together resulting in $N = 199$ sites for the gamma bands and $N = 196$ sites for the beta and alpha bands. Note that three sites from monkey B had to be excluded from the analysis of the alpha and beta LFP power due to a close to zero maximal percentage power relative to baseline, which after normalization resulted in very large values, distorting the mean distances.

images (mean net response to preferred image (of the 105 body part images) was 55 spikes/s (SEM = 6; N = 52 neurons)), indicating that isolated body parts are sufficient to elicit sizeable responses from the midSTS body patch. Because only a small number of these neurons were also tested with the whole-body images and these only with a small number of trials, a proper within-neuron comparison between the strengths of the responses for whole body and body parts stimuli could not be performed. However, one can compare the strength of the response to the body parts in this sample of body-patch neurons to the strength of the response to whole-body images for the body-patch neurons that were recorded with the main stimulus set (same neuronal sample as in Fig. 7A). This showed that the mean net response of the latter neuronal sample to the preferred whole-body image (44 spikes/s; SEM = 3; N = 215 cells) was comparable to that obtained for the body parts (n.s., Mann–Whitney U test). All 52 responsive neurons (assessed with a split-plot ANOVA; see Materials and Methods) showed highly selective responses to the body part stimulus set (Fig. 12A), with a profound selectivity for body part orientation. We quantified the orientation selectivity of each neuron for the body part exemplar producing the greatest response by computing a best-worst index:

$$BWI = \frac{R_{best} - R_{worst}}{R_{best}},$$

where R_{best} and R_{worst} are the net responses to the best and worst orientations for a particular exemplar, respectively. Note that an index of 1 means no response to the worst orientation. The median best-worst index for the body part eliciting the best response was 0.97 (25th percentile = 0.86; 75th percentile = 1.08; N = 52), demonstrating the strong dependence of the response on the orientation of a body part.

Figure 12B (top) also shows the Euclidean distances between all the body part stimulus pairs, based on the responses of the 52 body-patch neurons. First, note that there is no evidence of any clustering of the stimuli according to body part class, e.g., a clustering of all the images depicting a hand (also supported by hierarchical cluster analysis). Thus, these midSTS body-patch neurons do not appear to signal body part class per se, but instead show a pronounced selectivity for body part exemplars, viewed at specific orientations. Second, inspection of the dissimilarity matrix reveals that two images show a marked increase in pairwise dissimilarity relative to many other images (Fig. 12B,
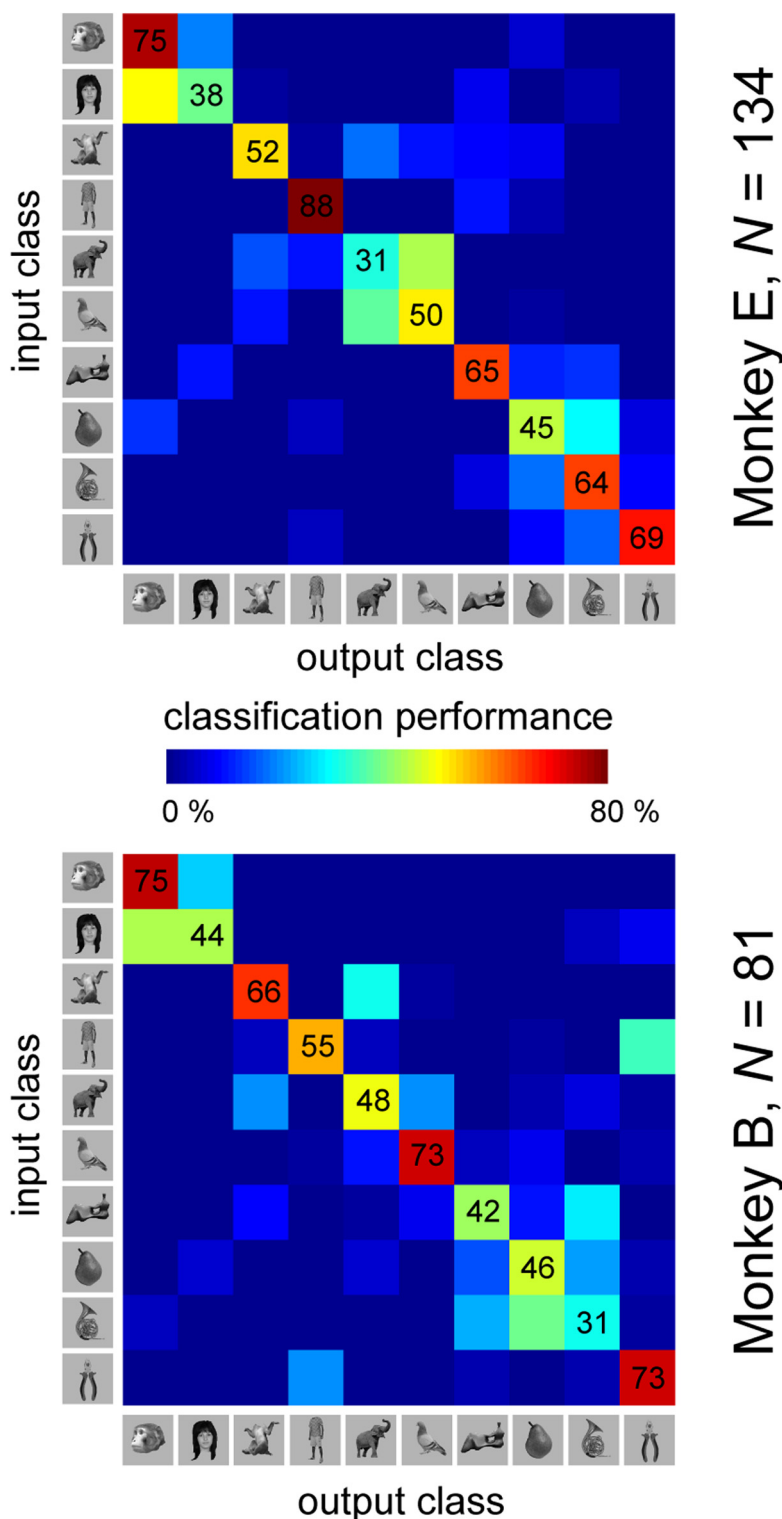
Figure 10. Confusion matrices displaying the performance of the linear SVM classifier for the two subjects (top and bottom). In each panel, the rows indicate the presented stimulus of a particular class (input class) and the columns the classifications made by the classifier (output class). The classification scores are color coded according to the color scale shown between the panels. Numbers indicate correct classification scores as a percentage. Perfect classification corresponds to values of 100% along the diagonal. The proportions of a row add up to 100%. Each class is indicated by an example image.

top, arrows): a penis and a leg in a grasping pose. This greater dissimilarity was not due to a greater overall response to these images (Fig. 12A, arrows): the mean net response, averaged across neurons, was 15 and 17 spikes/s for the penis and leg image, respectively, which compares well to the mean net re-
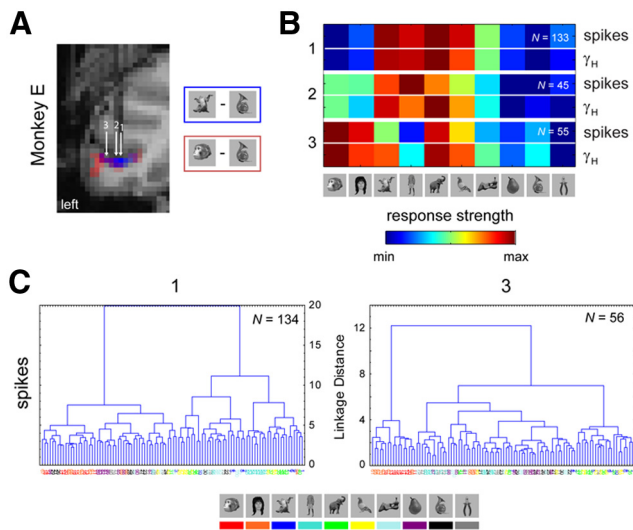
Figure 11. Comparison of the category selectivity between different medial-lateral locations within the midSTS of monkey E. ***A***, Coronal slice of monkey E's brain, showing the activation for bodies (blue, contrast; monkey bodies, objects M) and faces (red, contrast; monkey faces, objects M). The overlapping activations are in purple. Only activations that passed the FWE corrected level of *p* < 0.05 are shown. The white arrows show recording locations (1, center of midSTS body patch; 2, 1 mm more lateral than 1; 3, 3 mm more lateral than 1). ***B***, Mean normalized response per class for the spiking activity and high gamma (γ$_H$) band for each recording location as indicated in ***A***. Only neurons for which valid LFPs were obtained were included. ***C***, Hierarchical clustering of the stimuli based on the spiking responses of monkey E for the population of neurons recorded at locations 1 and 3 (***A***). The stimulus classes are color-coded according to the legend shown at the bottom. *N* indicates number of recorded neurons within each of the populations.

sponse for all other images (15 spikes/s; SD = 3). Thus, the higher dissimilarity for these images reflects greater differences in response between these images and the other images within the neurons. The higher average dissimilarity was strongly orientation selective, being present only for one of the five orientations of these body part images (Fig. 12*B*, top). Interestingly, the most marked dissimilarity was demonstrated by an upright, vertically oriented, erect penis that has obvious ethological significance. Note that this increased dissimilarity was not present for the other orientations [compare the dissimilarities for the vertical penis (Fig. 12*B*, left arrow) with the dissimilarities depicted for the next vertical line in the dissimilarity matrix of Fig. 12*B* (top), which indicates the data for the same image but rotated by 45°]. Finally, note that the pixel-based dissimilarities for the body part images (Fig. 12*B*, bottom) reveal a different pattern compared with the neural-based ones (Fig. 12*B*, top). Although the Spearman rank correlation between the two dissimilarity matrices was significantly different from 0 (*p* < 0.0001, Permutation test), it was small: r_s = 0.17.
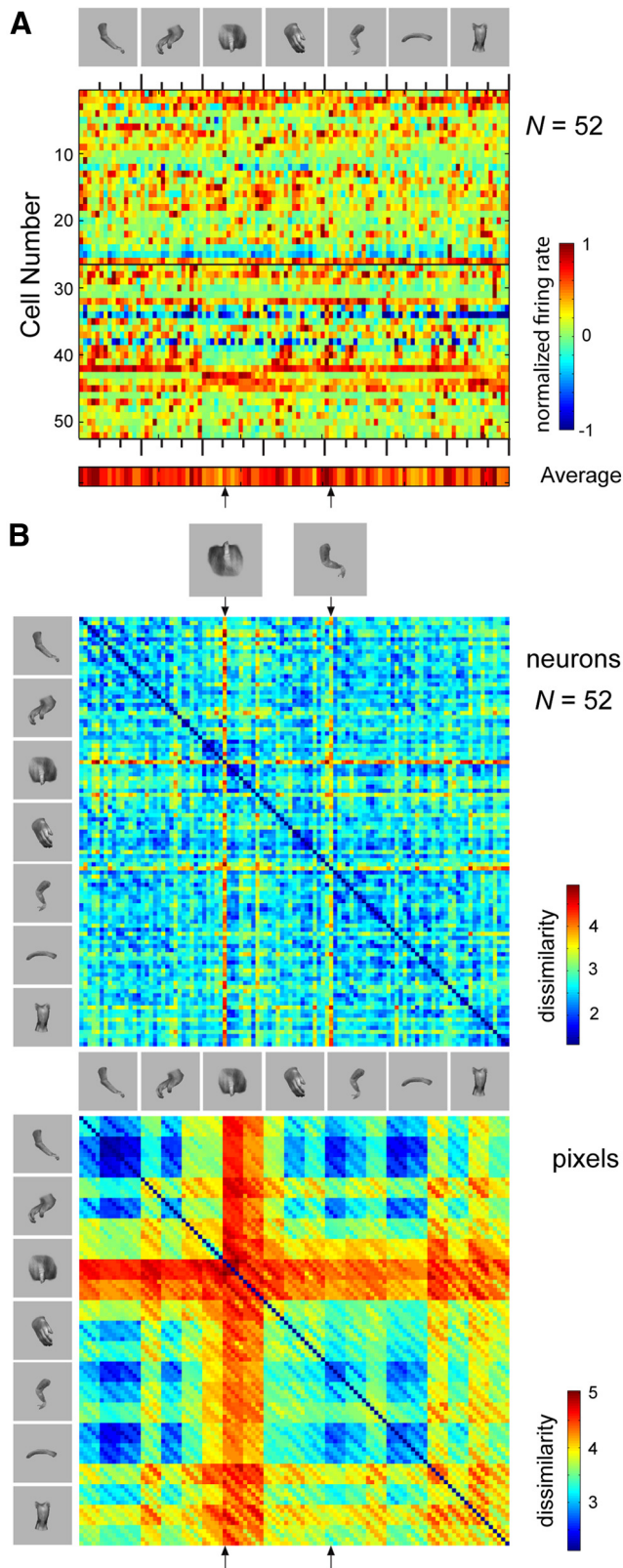
Discussion

Both the population spiking activity and LFP gamma power in the fMRI-defined midSTS body patch was greater for bodies (including monkey bodies, human bodies, mammals, and birds) compared with other objects, which fits the fMRI activation. This stronger response for bodies was absent in subgamma frequencies, despite the category selective responses for those frequencies. Importantly, the category selectivity at the population level resulted from averaging responses of a heterogeneous population of single units. The neurons showed a strong within-category selectivity, responding to only a small proportion of bodies. Despite such strong within-category selectivity at the single unit level, two distinct clusters, bodies versus nonbodies, were present

when analyzing the responses at the population level. A classifier that was trained using the responses to a subset of images was able to classify untrained images of bodies with high accuracy. Furthermore, the heterogeneous response properties of the neurons within the body patch allowed accurate classifications of all other classes, including faces and even artificial objects. In line with the fMRI data, the category selectivity depended on the location in the STS. The body-patch neurons showed strong selectivity for individual body parts of different orientations. Overall, these data suggest that single units in this fMRI defined midSTS body patch show a strong selectivity for individual body as well as nonbody images but with an overall bias toward a stronger response to bodies.

The proportion of body-category selective neurons depended on the metric used to define category selectivity, ranging from 33 to 53% (data pooled across both monkeys). These proportions were smaller than those observed for face selective cells in the neighboring face patches (Tsao et al., 2006, ML/MF 97% based on FSI; Issa and DiCarlo, 2012, PL 83%, ML 75% based on FSI; Ohayon et al., 2012, ML 82% based on *d'*). A low body-category selectivity can result from neurons responding to other stimuli than bodies and/or a high within-category selectivity. Indeed, both these factors contributed to the low body category selectivity in the body patch. First, only for 67% of the body patch neurons a body image produced the largest response of the neuron. Second, neurons showed a strong within-category selectivity, which reduces the overall mean response to bodies decreasing the category selectivity index. An often neglected issue when assessing category selectivity is the homogeneity of the stimuli within a category: the more homogeneous the stimuli within a class are (e.g., only frontal human faces; Tsao et al., 2006) or only frontal monkey faces (Issa and DiCarlo, 2012) the stronger the apparent category selectivity will be. Our body (and face) stimuli were rather heterogeneous compared with the face stimulus sets used in previous studies, sampling a broad range of bodies (different identities and postures). This might have contributed to both the relatively low category selectivity indices and the strong within-category selectivity. We argue that such a broader sampling of the category space provides a more ecologically valid assessment of the category selectivity of the neurons. Note that in general bodies can vary in shape and posture much more than faces, possibly leading to more selective responses within the body patch.

The relatively low category selectivity and the strong within-category selectivity of the body-patch neurons, combined with their stronger average response to bodies compared with nonbodies suggests that these neurons respond to features that happen to be present more often in images of bodies than of other objects. In other words, these neurons may not respond to bodies or body parts per se, but to features present in body images. The cluster analysis in which a few nonbody images, in particular the face with the limb-like hair style, were present in the body cluster suggests that local shape features that occur frequently in body images play an important role in determining the neural response in this patch. Note that each of these features need not be shared by all body images (or orientations), explaining the within-category selectivity. The identification of these features needs further work.

Bell et al. (2011) recorded in a more anterior STS region that was activated more strongly by body parts compared with faces, objects, and places. This region may correspond to the anterior body patch of Popivanov et al. (2012). Bell et al. (2011) reported that approximately half of the neurons in that anterior body part selective region responded stronger to body parts compared with the other three classes. This is less than what we observed in the present sample of midSTS body-patch neurons (78%), using the

same liberal criterion for category selectivity as Bell et al. (2011; BSI on raw responses >0). However, it remains to be seen whether this is a genuine difference between the two body patches or instead is due to dismembered body parts being less effective stimuli, compared with full (headless) bodies, in the anterior body patch, unlike what we observed in the midSTS body patch. Future research should compare the stimulus selectivity of the neurons between the two body patches.

Kiani et al. (2007) found a hierarchical representation of categories with a major distinction between animate (faces and bodies) and inanimate objects when analyzing the responses of a large number of neurons recorded at random locations within anterior IT. This differs from the bodies versus other classes (including faces) clustering that we observed here for the midSTS body patch. The clustering that we observed is very likely specific to the body-selective patches, mainly resulting from the weaker responses to stimulus classes other than bodies. In fact, at more lateral locations where responses to faces were at least as prominent as to bodies; faces became distinct from all other classes. The implication is that the category representation strongly varies with location within IT. It is possible that the animate versus inanimate distinction of Kiani et al. (2007) resulted from a random sampling over a wide expanse of IT cortex that masked strong regional differences in the hierarchical representations.

The correlation of the category selectivity between the LFP gamma power >60 Hz and spiking activity agrees with previous studies that observed a high correlation of the spiking activity and power in this band (Liu and Newsome, 2006; Belitski et al., 2008; Ray et al., 2008; De Baene and Vogels, 2010). Interestingly, the stronger fMRI activation for bodies compared with other stimulus classes agreed well with the category selectivity of the LFP gamma power but not with the power at lower frequencies, which is in line with some studies that observed a positive correlation between gamma band power and the BOLD response in primates (Mukamel et al., 2005; Niessing et al., 2005; Magri et al., 2012).

Huth et al. (2012) recently showed smooth gradients of semantic, category selectivity in human cortex with fMRI. Because of the low spatial resolution of fMRI, it could not be excluded that the category maps seen in that study appeared smoother than what is actually the case at a finer spatial scale. However, our data showing a transition between body-selective and a combination of face and body-selective population responses (for both spiking activity and LFP gamma band power) within the STS and the heterogeneous stimulus selectivity within the body patch supports the notion of smooth category-selective gradients. Indeed, the relative proportions of body-selective and face-selective neurons changed smoothly within STS, on a millimeter scale. The presence of face-selective neurons inside the body patch also agrees with a previous study demonstrating that face-selective neurons can be found outside the face patches (Bell et al., 2011).

We showed that the body selectivity seen at the fMRI and gamma power LFP level originates from averaging highly selective neurons that are biased, on average, to respond stronger to bodies than other object classes. This finding has implications for the interpretation of category-selectivity as measured with fMRI

Figure 12. Selectivity for body parts. **A**, Spiking activity matrix where each row represents the normalized responses of a neuron to each of the 105 body part stimuli. The horizontal line at cell number 26 separates the neurons of the two animals (below horizontal line, monkey B). Each column corresponds to a stimulus. The stimuli were first grouped per body part class, as indicated by the example images, then by body part exemplar (three exemplars per class) and subsequently by orientation. Short and long tick marks at the top and bottom of the matrix indicate the divisions between the three exemplars of a class and between the different classes, respectively. Normalized response strength is indicated by a color code (see colored bar).

←

The mean normalized responses, averaged across neurons, for each stimulus are shown in the bar below the matrices (average). **B**, Top, The dissimilarity matrix of pairwise Euclidean distances for all stimulus pairs, based on the spiking responses of 52 body-patch neurons. Bottom, The dissimilarity matrix based on the gray level difference in the corresponding pixels. The same conventions are used as in Figure 8A. The two images that demonstrated a relatively high neural dissimilarity are shown above the top panel and the arrows indicate their mean response (**A**) and dissimilarity values (**B**).

(Mur et al., 2012; Vul et al., 2012) and LFP studies (Liu et al., 2009). The category selectivity measured with these techniques can overestimate the category selectivity that is actually present at a finer spatial scale simply due to the averaged activity of a large population of neurons, which may have heterogeneous stimulus selectivity and strong within-category selectivity.

The present study shows that categorization of superordinate categories ("bodies" versus "nonbodies") can be performed quite accurately based on the responses of a small population of neurons in the midSTS body patch. The heterogeneous but biased selectivity within the body patch allows both the classification of bodies versus other categories by a weighted sum of the responses (as shown by the SVM classification analysis) and the identification of bodies by differentiating the responses of different units within the patch. Responses of the same neuronal population can also categorize faces versus other objects and even carry information about other inanimate object classes. How this rich and diverse repertoire of responses eventually relates to behavioral categorization and identification of bodies and perhaps of other stimuli, however, will require the application of causal techniques.

References

Afraz SR, Kiani R, Esteky H (2006) Microstimulation of inferotemporal cortex influences face categorization. Nature 442:692–695. CrossRef Medline

Belitski A, Gretton A, Magri C, Murayama Y, Montemurro MA, Logothetis NK, Panzeri S (2008) Low-frequency local field potentials and spikes in primary visual cortex convey independent visual information. J Neurosci 28:5696–5709. CrossRef Medline

Bell AH, Hadj-Bouziane F, Frihauf JB, Tootell RB, Ungerleider LG (2009) Object representations in the temporal cortex of monkeys and humans as revealed by functional magnetic resonance imaging. J Neurophysiol 101:688–700. CrossRef Medline

Bell AH, Malecek NJ, Morin EL, Hadj-Bouziane F, Tootell RB, Ungerleider LG (2011) Relationship between functional magnetic resonance imaging-identified regions and neuronal category selectivity. J Neurosci 31:12229–12240. CrossRef Medline

Bindemann M, Scheepers C, Ferguson HJ, Burton AM (2010) Face, body, and center of gravity mediate person detection in natural scenes. J Exp Psychol Hum Percept Perform 36:1477–1485. CrossRef Medline

De Baene W, Vogels R (2010) Effects of adaptation on the stimulus selectivity of macaque inferior temporal spiking activity and local field potentials. Cereb Cortex 20:2145–2165. CrossRef Medline

De Baene W, Premereur E, Vogels R (2007) Properties of shape tuning of macaque inferior temporal neurons examined using rapid serial visual presentation. J Neurophysiol 97:2900–2916. CrossRef Medline

de Gelder B, Van den Stock J, Meeren HK, Sinke CB, Kret ME, Tamietto M (2010) Standing up for the body: recent progress in uncovering the networks involved in the perception of bodies and bodily expressions. Neurosci Biobehav Rev 34:513–527. CrossRef Medline

Downing PE, Jiang Y, Shuman M, Kanwisher N (2001) A cortical area selective for visual processing of the human body. Science 293:2470–2473. CrossRef Medline

Ekstrom LB, Roelfsema PR, Arsenault JT, Bonmassar G, Vanduffel W (2008) Bottom-up dependent gating of frontal signals in early visual cortex. Science 321:414–417. CrossRef Medline

Fabre-Thorpe M, Richard G, Thorpe SJ (1998) Rapid categorization of natural images by rhesus monkeys. Neuroreport 9:303–308. CrossRef Medline

Freiwald WA, Tsao DY (2010) Functional compartmentalization and viewpoint generalization within the macaque face-processing system. Science 330:845–851. CrossRef Medline

Freiwald WA, Tsao DY, Livingstone MS (2009) A face feature space in the macaque temporal lobe. Nat Neurosci 12:1187–1196. CrossRef Medline

Giese MA, Poggio T (2003) Neural mechanisms for the recognition of biological movements. Nat Rev Neurosci 4:179–192. CrossRef Medline

Hall M, Frank E, Holmes G, Pfahringer B, Reutemann P, Witten IH (2009) The WEKA data mining software: an update. SIGKDD Explorations 11:10–18.

Huth AG, Nishimoto S, Vu AT, Gallant JL (2012) A continuous semantic space describes the representation of thousands of object and action categories across the human brain. Neuron 76:1210–1224. CrossRef Medline

Issa EB, DiCarlo JJ (2012) Precedence of the eye region in neural processing of faces. J Neurosci 32:16666–16682. CrossRef Medline

Kayaert G, Biederman I, Op de Beeck HP, Vogels R (2005) Tuning for shape dimensions in macaque inferior temporal cortex. Eur J Neurosci 22:212–224. CrossRef Medline

Kiani R, Esteky H, Mirpour K, Tanaka K (2007) Object category structure in response patterns of neuronal population in monkey inferior temporal cortex. J Neurophysiol 97:4296–4309. CrossRef Medline

Liu H, Agam Y, Madsen JR, Kreiman G (2009) Timing, timing, timing: fast decoding of object information from intracranial field potentials in human visual cortex. Neuron 62:281–290. CrossRef Medline

Liu J, Newsome WT (2006) Local field potential in cortical area MT: stimulus tuning and behavioral correlations. J Neurosci 26:7779–7790. CrossRef Medline

Magri C, Schridde U, Murayama Y, Panzeri S, Logothetis NK (2012) The amplitude and timing of the BOLD signal reflects the relationship between local field potential power at different frequencies. J Neurosci 32:1395–1407. CrossRef Medline

Mukamel R, Gelbard H, Arieli A, Hasson U, Fried I, Malach R (2005) Coupling between neuronal firing, field potentials, and fMRI in human auditory cortex. Science 309:951–954. CrossRef Medline

Mur M, Ruff DA, Bodurka J, De Weerd P, Bandettini PA, Kriegeskorte N (2012) Categorical, yet graded: single-image activation profiles of human category-selective cortical regions. J Neurosci 32:8649–8662. CrossRef Medline

Niessing J, Ebisch B, Schmidt KE, Niessing M, Singer W, Galuske RA (2005) Hemodynamic signals correlate tightly with synchronized gamma oscillations. Science 309:948–951. CrossRef Medline

Ohayon S, Freiwald WA, Tsao DY (2012) What makes a cell face selective? The importance of contrast. Neuron 74:567–581. CrossRef Medline

Op de Beeck H, Wagemans J, Vogels R (2001) Inferotemporal neurons represent low-dimensional configurations of parameterized shapes. Nat Neurosci 4:1244–1252. CrossRef Medline

Pinsk MA, DeSimone K, Moore T, Gross CG, Kastner S (2005) Representations of faces and body parts in macaque temporal cortex: a functional MRI study. Proc Natl Acad Sci U S A 102:6996–7001. CrossRef Medline

Pinsk MA, Arcaro M, Weiner KS, Kalkus JF, Inati SJ, Gross CG, Kastner S (2009) Neural representations of faces and body parts in macaque and human cortex: a comparative FMRI study. J Neurophysiol 101:2581–2600. CrossRef Medline

Popivanov ID, Jastorff J, Vanduffel W, Vogels R (2012) Stimulus representations in body-selective regions of the macaque cortex assessed with event-related fMRI. Neuroimage 63:723–741. CrossRef Medline

Ray S, Maunsell JH (2011) Different origins of gamma rhythm and high-gamma activity in macaque visual cortex. PLoS Biol 9:e1000610. CrossRef Medline

Ray S, Hsiao SS, Crone NE, Franaszczuk PJ, Niebur E (2008) Effect of stimulus intensity on the spike-local field potential relationship in the secondary somatosensory cortex. J Neurosci 28:7334–7343. CrossRef Medline

Sawamura H, Orban GA, Vogels R (2006) Selectivity of neuronal adaptation does not match response selectivity: a single-cell study of the FMRI adaptation paradigm. Neuron 49:307–318. CrossRef Medline

Tallon-Baudry C, Bertrand O (1999) Oscillatory gamma activity in humans and its role in object representation. Trends Cogn Sci 3:151–162. CrossRef Medline

Tsao DY, Freiwald WA, Knutsen TA, Mandeville JB, Tootell RB (2003) Faces and objects in macaque cerebral cortex. Nat Neurosci 6:989–995. CrossRef Medline

Tsao DY, Freiwald WA, Tootell RB, Livingstone MS (2006) A cortical region consisting entirely of face-selective cells. Science 311:670–674. CrossRef Medline

Vanduffel W, Fize D, Mandeville JB, Nelissen K, Van Hecke P, Rosen BR, Tootell RB, Orban GA (2001) Visual motion processing investigated using contrast agent-enhanced fMRI in awake behaving monkeys. Neuron 32:565–577. CrossRef Medline

Vangeneugden J, De Mazière PA, Van Hulle MM, Jaeggli T, Van Gool L, Vogels R (2011) Distinct mechanisms for coding of visual actions in macaque temporal cortex. J Neurosci 31:385–401. CrossRef Medline

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