Published in final edited form as:

Soc Neurosci. 2016 December ; 11(6): 627–636. doi:10.1080/17470919.2015.1131194.

Task influences pattern discriminability for faces and bodies in ventral occipitotemporal cortex

Na Yeon Kim and Gregory McCarthy

Department of Psychology, Yale University, New Haven, Connecticut, USA

Abstract

Our prior research showed that faces and bodies activate overlapping regions of the ventral occipitotemporal cortex (VOTC). However, faces and bodies were nonetheless discriminable in these same overlapping regions when their spatial patterns of activity were classified using multi-voxel pattern analysis. Here we investigated whether these spatial patterns and their time courses were influenced by different categorization tasks. Participants viewed pictures of faces or headless bodies depicting a happy or fearful emotion. In one task, they categorized the picture as a face or a body regardless of emotion. In the other task, they categorized the emotion regardless of whether it was depicted by a face or body. Using a classifier trained on independent data, we found higher face-body classification accuracy for the emotion categorization task. The classifier was applied to each post-stimulus time-point to characterize the temporal course of classification. Accuracy initially rose equivalently above chance for both tasks, but then increased over a longer duration when participants categorized emotions. Thus, the temporal course of pattern differences between faces and bodies in VOTC were modulated by the behavioral goal of the observer, suggesting the top-down modulatory effect of task context on the category-selectivity activity in the VOTC.

Keywords

face perception; body perception; MVPA; ventral occipitotemporal cortex; fusiform gyrus

Neuroimaging research has demonstrated that images of faces and bodies activate regions in the ventral (VOTC) and lateral occipitotemporal cortices (LOT). Within the VOTC, the perception of faces strongly activates the mid-lateral fusiform gyrus (e.g., Puce, Allison, Gore, & McCarthy, 1995). The apparent face selectivity of this spatially circumscribed region has led researchers to propose that this region represents a neural module for face processing (“Fusiform Face Area”, or FFA, Kanwisher, McDermott, & Chun, 1997). The perception of bodies or body parts has also been reported to activate discrete regions in the VOTC (“Fusiform Body Area”, or FBA, Peelen & Downing, 2005) and LOTC (“Extrastriate Body Area”, or EBA, Downing, Jiang, Shuman, & Kanwisher, 2001) – suggesting that highly specific regions of cortex may selectively process different features of these corporeal forms.

A different view suggests that faces, bodies, and other object categories are represented in spatial patterns of activity within more extensive and overlapping regions of VOTC and LOTC (Haxby et al., 2001). Support for this perspective has been obtained by multi-voxel pattern analysis (MVPA) of functional MRI (fMRI) data. A recent paper from our laboratory provided support for this view by showing that spatial patterns of activity identified by MVPA can discriminate faces from bodies in VOTC regions that are strongly activated by *both* stimulus categories (i.e., in the extensive overlap of the FFA and FBA regions) (Kim, Lee, Erlendsdottir, & McCarthy, 2014). In this region of overlap, a univariate analytic approach revealed no significant difference in their mean activations. These findings indicate that the category-selectivity of neural responses can be defined at a finer resolution than can be easily examined in conventional univariate analysis (Downing, Wiggett, & Peelen, 2007; Kaiser, Strnad, Seidl, Kastner, & Peelen, 2014; Weiner & Grill-Spector, 2010).

It is as yet unclear what face and body features drive these different spatial patterns of activity in the VOTC. The patterns may rely on low-level visual properties (e.g., overall shape differences), as in early visual areas where orientation can be discriminated (Haynes & Rees, 2006; Kamitani & Tong, 2005). Alternatively, spatial activation patterns within the VOTC could represent more subtle facial or bodily information. For instance, some studies of face processing have reported that activation patterns in face-specific regions in the VOTC discriminate race (Brosch, Bar-David, & Phelps, 2013; Ratner, Kaul, & Van Bavel, 2013) and emotional expressions (Harry, Williams, Davis, & Kim, 2013).

The features of faces and bodies that are captured in the spatial activity patterns may reflect a patchy, but enduring, representation of features across cortex. But, given the myriad dimensions on which faces can be categorized, we wondered whether the spatial patterns that discriminate among faces, or which discriminate faces from other corporeal categories, such as bodies, are affected by the task in which the participant is engaged. For example, would a spatial pattern of activity in the VOTC that discriminated faces from bodies, continue to discriminate faces from bodies when a task required participants to categorize along a difference dimension, where some faces and bodies are categorized into one group, and other faces and bodies into a different group? That is, are the patterns sensitive to task context, even if the participant views the same stimuli?

Two recent studies demonstrated top-down modulatory effects on the fine-grained representations of objects. Harel and colleagues compared the object decoding accuracies within and across six different tasks and reported that task context had significant influence on object representations in the ventral and lateral object-selective regions (Harel, Kravitz, & Baker, 2014). Çukur and colleagues demonstrated that attention shifts voxel-wise category tuning in occipito-temporal and fronto-parietal cortices to expand the representation of attended categories and compress the representation of unattended categories during natural vision (Çukur, Nishimoto, Huth, & Gallant, 2013).

Here, we investigated how different behavioral tasks might influence classification of faces from bodies in the VOTC and control regions. We presented participants with pictures of faces and (headless) bodies that depict either a happy or fearful emotion. In one task, participants were required to discriminate faces from bodies regardless of emotional

expression. In a second task, participants were required to discriminate happy from fearful emotions, regardless of whether the emotion was depicted by a face or a body. If the happy-fearful discrimination task alters the processing within VOTC, we might expect this to be reflected in changes in the classification accuracy for faces and bodies. Indeed, if neural activity in the VOTC is repurposed to discriminate emotional categories rather than corporeal categories, we might expect a classifier trained to discriminate faces from bodies to perform less well when participants are discriminating emotions *across* face-body categories. To examine this, we performed MVPA on a set of neutral faces and bodies in an independent data set, and then applied the resulting classifier to both the face-body discrimination task and the happy-fearful discrimination task. If pattern discriminability between faces and bodies is task invariant in the VOTC regions tested, then we should observe no differences in classification accuracy in these two tasks.

We also considered the temporal evolution of classification performance. Our group has previously suggested, on the basis of subdural electrophysiological recordings, that the processing occurring within VOTC face regions is temporally dissociable, such that early processing might be category-specific (with faces being the category studied) while later processing at the same neural loci might be more sensitive to task and attentional manipulations (see, Engell & McCarthy, 2010; Puce, Allison, & McCarthy, 1999). To test for this possibility, we investigated task-related differences in classification performance at each time point to determine whether classifier accuracy diverged between the two task conditions across time.

Our primary analysis depended upon applying a classifier trained to discriminate neutral faces and bodies from an independent data set, and thus the features identified by this independent classifier were assessed in each task condition. To examine whether different features might be selected in different tasks, we also conducted a subsidiary analysis in which we compared the performance of classifiers that were trained and tested within each task condition. This allowed us to ask whether particular features identified by each classifier differed as a function of task and to determine whether emotional expressions can be classified by activation patterns in VOTC.

Material and Methods

Participants

Twenty healthy adults (11 female, mean age 23.5 ± 4.9 years, all right-handed) participated in this study. All had normal or corrected-to-normal vision and no history of neurological or psychiatric illnesses. All participants gave written informed consent. The study was approved by the Yale Human Investigations Committee.

Stimuli

Figure 1A presents exemplar stimuli for the main experiment and for the localizer task. The stimuli for the main experiment consisted of pictures of human faces and bodies that depicted a happy or fearful emotion. There was a total of 120 pictures, consisting of 30 pictures for each of the four stimulus categories: Face-Happy, Face-Fear, Body-Happy,

Body-Fear. Face stimuli were selected from the NimStim database of facial expressions (Tottenham et al., 2009). Body stimuli were selected from the Bodily Expressive Action Stimulus Set (BEAST; de Gelder & Van den Stock, 2011). The original body stimuli presented whole bodies with blurred faces expressing four basic emotions (i.e., Happiness, Fear, Anger, Sadness). Each of the body stimuli presented a person expressing a “highly recognizable” emotion without specific instructions about gestures, and the original images of the BEAST were validated through an emotion categorization experiment. To avoid possible confounding effects from the blurred faces, we cropped the original images to generate images of bodies without head contours. We only used Happy and Fear emotion categories because those two categories were most distinguishable in the headless bodies. For example, bodies in the Happiness category showed welcoming gestures with open arms, whereas those in the Fear were leaning backward with defensive hand gestures. We selected 30 individuals (15 males and 15 females) from original databases for face and body stimuli in order to control for identity and gender. All stimuli were presented on a gray background.

A set of computer-generated faces and bodies with neutral expressions was used in an independent localizer task. In the primary analysis, an MVPA was conducted on these stimuli (see below) and the classifier was applied to the main experiment tasks.

Procedures

Stimuli were presented using Psychophysics Toolbox (Brainard, 1997) in MATLAB (The Mathworks Inc., Natick, MA, USA).

The experiment (Figure 1B) consisted of two task conditions: a face-body discrimination task and a happy-fearful discrimination task. Task conditions were indicated by a text cue (i.e., “Face or Body?” or “Happy or Fearful?”) that appeared directly above the stimulus picture. Participants were instructed to pay attention to the text cue and make a button response based on the stimulus by using a four-button response pad. Each button on the response pad was labeled as Face, Body, Happy, or Fearful. All stimuli were presented twice per participant, once for each task condition.

Each participant completed 6 runs for the main experiment, and each run lasted 6 minutes. Each run consisted of four task blocks in a randomized order, two face-body task blocks and two happy-fearful task blocks. The cue appeared in the beginning of each block and stayed on the screen for the entire block. Each task block consisted of 10 stimulus pictures. Stimuli were presented in a randomized order. Each stimulus was presented for 2 seconds and was separated by a 4, 6, or 8-s fixation interval. There was a 10-s fixation interval between blocks. Participants viewed each stimulus picture twice, once for each task condition. Each session included 12 face-body task blocks and 12 happy-fearful task blocks in total.

Participants completed a localizer task after the main experiment. Sixteen participants completed four localizer runs, and four finished two localizer runs. The localizer task procedures closely resembled those used in our prior study (Kim et al. 2014). Briefly, face, body, and house pictures were presented in a block design, where 12-s stimulus blocks consisting of rapidly presented pictures of a single category were interleaved with 12-s

fixation blocks. Participants were instructed to press a button when they saw the same picture twice consecutively.

fMRI image acquisition and preprocessing

Data were acquired at the Magnetic Resonance Research Center at Yale University using a 3.0 T Siemens TIM Trio scanner with a 32-channel head coil. Functional images were acquired using a multiband imaging sequence (TR = 2000 ms, TE = 32 ms, flip angle = 62°, FOV = 210 × 202 mm, matrix = 104 × 100, slice thickness = 2.0 mm, 60 slices, multiband accelerate factor = 3) yielding isotropic voxels that were 2 mm³. Two structural images were acquired for registration: T1 coplanar images were acquired using a T1 Flash sequence (TR = 335 ms, TE = 2.61 ms, flip angle = 70°, FOV = 240 mm, matrix = 192 × 192, slice thickness = 2.0 mm, 60 slices), and high-resolution images were acquired using a 3D MP-RAGE sequence (TR = 2530 ms, TE = 2.77 ms, flip angle = 7°, FOV = 256 mm, matrix = 256 × 256, slice thickness = 1 mm, 176 slices).

Image preprocessing

Image preprocessing for both the main experiment and the localizer task was performed using the FMRIB Software Library (FSL, <http://www.fmrib.ox.ac.uk/fsl>) and AFNI tools (Cox, 1996). Structural and functional images were skull-stripped using the Brain Extraction Tool (BET). The first three volumes (6 s) of each functional dataset were discarded to allow for MR equilibration. Functional images were then corrected for motion using the MCFLIRT linear realignment, and high-pass filtered with a 0.01 Hz cut-off to remove low-frequency drift. Data were not spatially smoothed. The functional data were registered to the coplanar images, which were in turn registered to the high-resolution structural images, using non-linear registration, and then normalized to the Montreal Neurological Institute's template (MNI152).

Multi-voxel pattern analysis (MVPA)

In our main analysis, we compared the face-body pattern discriminability between tasks by applying a classifier trained on the localizer data set to data from each task condition of the main experiment. A linear support vector machine (SVM) classifier (LIBSVM; Chang & Lin, 2011), as implemented in PyMVPA (Hanke et al., 2009), was trained using the Face and Body examples of the localizer task. For our main analyses, the analysis was constrained to voxels comprising the temporal occipital fusiform cortex as defined by the TOFC overlay from the Harvard-Oxford Structural Atlas. This mask included 7,485 voxels and covered the collateral sulcus and inferior temporal gyrus as well as the mid-fusiform gyrus in both hemispheres ("TOFC ROI"). Smaller and anatomically restricted masks were used in secondary analyses to drill down into the main results, as described below.

We obtained parameter estimates (or betas) for each trial and for each participant by using hemodynamic model fitting. Regression analyses were performed on the preprocessed functional data using AFNI's 3dDeconvolve and 3dREMLfit functions, where each trial was modeled using the BLOCK5 basis function with duration of 2 seconds. The resulting beta volumes were concatenated into a single beta series for each participant. This yielded a

series of 8 beta volumes (2 localizer runs) for each of the first four participants and a series of 16 beta volumes (4 localizer runs) for the other participants.

We then tested how accurately the trained classifier could predict the Face and Body examples of the face-body discrimination task and the Face and Body examples of the happy-fearful discrimination task. We used PyMVPA's Balancer function to ensure that the number of Face examples was equal to that of Body examples and that the datasets from two task conditions contained the same number of Face and Body examples. The data were z-scored using the mean and standard deviation for each stimulus type within each task condition (i.e., localizer and two tasks of the main experiment). This also matched the scale of the data used for the training and testing processes. These analyses were restricted to trials in which correct responses were made.

Time course analysis

We compared the temporal evolution of classification performance between task conditions by obtaining classification accuracy at each volume (which was acquired over a 2-s TR, or repetition time). We examined activation patterns using percent signal changes from five volumes for each trial (i.e., up to 10 seconds after stimulus onset). We generated a separate dataset for each time point by concatenating values from a specific time point of all trials into a single series. For example, the first volumes of all trials were concatenated into a single series, and the second volumes of all trials were obtained to create another series. As in the cross-task classification, a classifier was trained on the localizer dataset and tested on each of the five time points within individual participants. Group-level accuracies were obtained for each time point by averaging the accuracies from individual participants. A paired-sample t-test was performed on each time point in order to compare the accuracies between the two task conditions.

Within-task pattern classification

Subsequent to our primary analysis using a classifier trained on an independent localizer task, we also performed within-task pattern classification analysis with different labels on the same datasets. Specifically, we first tested how patterns of activity in the anatomical ROI could discriminate faces from bodies regardless of emotional expression *within* each task condition. We then tested how those patterns could discriminate happy from fearful emotions regardless of category within each task condition. Finally, we compared those classification results between the two task conditions. SVM classifiers were trained and tested within each task condition using a leave-one-run-out cross-validation strategy. The number of samples for each condition was balanced in each cross-validation fold using PyMVPA's Balancer function. Classification accuracies and *p*-values were obtained for individual participants. A paired-sample t test was performed to compare group-level classification accuracies between tasks.

Additional Regions of Interest (ROIs)

As described above, our primary analyses included all voxels from the brain region masked by the Harvard-Oxford TOFC anatomical ROI. However, following that primary analysis, we wished to further explicate our results by examining more spatially restricted ROIs and

control regions. We used FreeSurfer's probabilistic atlases (Fischl et al., 2002) to restrict our analysis to the fusiform gyrus. The fusiform ROI from FreeSurfer includes gray matter of the bilateral fusiform gyrus and extends to the anterior fusiform gyrus. This ROI includes 2,912 voxels in total. We also used FreeSurfer to define V1 and V2 to conduct control analyses to examine whether these regions could discriminate faces from bodies, and more importantly, whether discrimination, if found, was influenced by task.

Finally, to restrict analysis to the FFA, we defined an ROI based upon the Atlas of Social Agent Perception (ASAP; Engell and McCarthy, 2013), a probabilistic atlas developed from a large localizer dataset ($n = 124$) comparing activations to faces versus scenes ("Face-scene ROI"). Similar to our prior study (Kim et al. 2014), we selected all voxels with P (the probability that a participant show a category-specific response to faces at that specific voxel) $> .53$. This yielded a mask of 180 voxels, which consisted primarily of the right and left FFAs.

Results

Behavioral results

Participants made correct responses for 116 of 120 trials (96.7%) in the face-body discrimination task, and 106 of 120 trials (88.3%) in the happy-fearful discrimination task. The difference in response accuracy between tasks was significant ($p < .001$). Our main analysis was performed on the trials in which the participant pressed the correct button.

Multi-voxel Pattern Analysis

Our primary analysis was conducted within the TOFC ROI. A classifier trained on the independent localizer data successfully discriminated Face and Body examples from the face-body task ($M = 68.3\%$, $SD = 9.4$) and from the happy-fearful task ($M = 75.3\%$, $SD = 10.7$). The classification accuracy was significantly higher in the happy-fearful task than in the face-body task ($t(19) = 5.21$, $p < .0001$).

Figure 2 presents cross-task classification accuracies from the primary TOFC analysis and from the additional anatomical ROIs as described in Methods. The accuracies were higher in the happy-fearful task than in the face-body task within all VOTC ROIs. V1 and V2 showed lower accuracies in face-body classification compared to VOTC regions, but paired-sample t tests revealed that the accuracies were significantly above chance level (50%), $p < .0001$ for both V1 and V2. However, there was no task-related difference in V1 and V2. Thus, faces and bodies were successfully discriminated by the patterns within posterior visual areas, but task did not influence the classification performance.

Time course analysis

Figure 3 shows group-average accuracies at each time point from the cross-task classification analysis on the TOFC ROI. The accuracy from the first post-stimulus TR (a volume obtained 0 to 2 seconds after stimulus onset, and plotted at the mid-point of the interval) was nearly chance-level (50%) and did not differ between tasks. The accuracies were above chance both in the face-body task ($M = 59\%$, $SD = 7.0$) and in the happy-fearful

task ($M = 59\%$, $SD = 8.1$) at the second volumes (2 to 4 seconds after onset), but did not differ between tasks. The accuracy peaked at the third volume (4 to 6 seconds after onset), and was higher in the happy-fearful task ($M = 76.6\%$, $SD = 9.6$) than in the face-body task ($M = 73.5\%$, $SD = 11.7$); $t(19) = 2.43$, $p = 0.025$. The difference between the two tasks was greater at the fourth volume (6 to 8 seconds after onset), where the accuracy in the happy-fearful task ($M = 75.7\%$, $SD = 11.8$) was higher than that in the face-body task ($M = 67.3\%$, $SD = 12.8$); $t(19) = 4.02$, $p = 0.0007$. The accuracies decreased at the fifth volume (8 to 10 seconds after onset), but the difference between tasks remained significant (59.6% in the happy-fearful task and 54.1% in the face-body task; $t(19) = 2.91$, $p = 0.009$).

Within-task pattern classification

In the TOFC ROI, faces and bodies were successfully discriminated by the patterns of activity both in the face-body task ($M = 66.1\%$, $SD = 8.1$) and in the happy-fearful task conditions ($M = 72.7\%$, $SD = 7.0$). As in the classification trained upon the localizer data, the classification accuracy was higher in the happy-fearful task than in the face-body task; $t(19) = 4.12$, $p = 0.00058$. Because mean differences between conditions could influence classification performance (Coutanche, 2013), the same classification analysis was performed on the z-scored datasets, where the mean difference between stimulus categories was removed. The results from the classification analysis on the z-scored data were consistent: the accuracy was higher in the happy-fearful task ($M = 72.6\%$, $SD = 8.7$), compared to the face-body task ($M = 66.6\%$, $SD = 7.6$); $t(19) = 3.34$, $p = 0.0035$.

Figure 4 displays within-task classification accuracies from the TOFC and additional ROIs tested. Consistent with the cross-task classification results, the accuracy differed between tasks in the ROIs within VOTC, but not in V1 and V2. In the VOTC regions the accuracy of face-body classification was higher in the happy-fearful task than in the face-body task.

In the TOFC ROI, there was no difference in the accuracy of happy-fearful classification between the face-body task ($M = 53.3\%$, $SD = 5.08$) and happy-fearful task ($M = 51.8\%$, $SD = 8.90$), $t(19) = 0.75$, $p = 0.464$. The accuracy for emotional expression classification did not exceed chance in either task.

Discussion

Here we showed that the classification accuracy for discriminating faces and bodies based upon spatial patterns of fMRI activity within the VOTC was influenced by task. The performance accuracy of the pattern classifier increased when participants categorized emotional expressions depicted by faces *and* bodies, compared to when they simply discriminated faces from bodies regardless of emotional expression. This result held regardless of whether the pattern classifier was trained on an independent localizer data set of neutral faces or bodies, or trained on the main task data using a leave-one-out procedure. Since mean activation differences could facilitate classification performance (Coutanche, 2013), we compared classification analyses on the same data sets before and after mean normalization. The results were consistent regardless of the presence of mean differences, suggesting higher classification accuracies in the emotional expression task cannot be explained by mean activation differences. The task differences in classifier performance

were evident in our primary ROI, the TOFC, but were also present in smaller and more restricted VOTC ROIs including a small ROI that mainly consisted of the left and right FFA. However, no task-related change in pattern discriminability was observed in posterior visual areas V1 and V2.

How can we interpret these robust task differences in classification performance? One hypothesis inconsistent with our present results is that activity patterns within VOTC regions are altered to represent either generic category (i.e., face or body) or specific information (i.e., emotional expressions) depending on which information is required by the task. While response patterns for object categories have been shown to vary depending on task context (Harel et al., 2014), our results show that this is not the case for face- or body-evoked activity patterns in VOTC. On the contrary, face- and body-specific pattern information in VOTC became sharper when participants were discriminating happy from fearful emotions presented in a face or body image. Furthermore, our additional within-task analyses revealed no patterns of activity in VOTC that discriminated between happy and fearful emotions in *either* of the task conditions. This finding contradicts the results of Harry and colleagues (Harry et al. 2013). It is consistent, however, with the idea that emotional content is represented in a different region of the face perception network, such as superior temporal sulcus (STS) (Said, Moore, Engell, & Haxby, 2010).

Our behavioral data corroborates our intuition that the emotional expression categorization task is more difficult than the face-body categorization task. Faces and bodies could be discriminated by their overall shape without the need for detailed processing of specific features. In contrast, to discriminate happy from fearful expressions, participants had to recognize more subtle features, such as the shape of eyes and orientation of the hands. One explanation for the improved pattern discrimination, then, is that face- and body-specific patterns increased in signal-to-noise, or became sharper, because visual details in a face or body image were processed further when participants categorized emotional expressions compared to when they categorized corporeal categories. That is, participants were required to attend more closely to the faces and bodies when categorizing emotional expressions.

It has been demonstrated that attention enhances not only the overall magnitude of activity, but also the similarity between the patterns of activation to a repeated stimulus (Moore, Yi, & Chun, 2013). However, such task-based differences were not observed in occipital visual areas, such as V1 and V2, suggesting that better discrimination between face- and body-evoked activity patterns is not simply driven by enhanced processing of low-level perceptual details. The time course analysis revealed that face-body classifier accuracy in the TOFC was above chance in the 2–4 seconds after stimulus onset for both tasks, but did not differ between tasks. The classification accuracy increased at 4–6 seconds for both tasks, but the accuracy at this latency was greater for the emotion task than for the face-body task. Moreover, the increased classification accuracy for the emotional expression task persisted to 8–10 sec. Although fMRI is not suitable to study temporal dynamics at a fine scale, the similarity of classification performance for both tasks in the 2–4 sec range are at least consistent with our data from subdural electrophysiology that the initial response of category-specific processing regions in VOTC is relatively immune to task and attention (Engell & McCarthy, 2010; Puce, Allison, & McCarthy, 1999). That the difference in

classification accuracy emerges later in the epoch suggests that task effects modulate the persistence and strength of the classifier features that discriminate face and bodies in the later portion of the post-stimulus epoch. This suggests the operation of task-related top-down influences upon VOTC.

Other studies reported task rule-specific activation patterns in the ventrolateral frontal and parietal cortices (Zhang, Kriegeskorte, Carlin, & Rowe, 2013) and the sequential recruitment of such regions during task preparation (Bode & Haynes, 2009). Future research may be able to examine the neural mechanisms that underlie the task-related differences by investigating how the interaction between VOTC regions and fronto-parietal attentional network contributes to the task-dependent variation in the neural responses.

In summary, the present study demonstrates the effect of task and attentional manipulations on the discriminability of face- and body-evoked activity patterns. Our findings provide further evidence for stimulus-driven and task-driven components of the BOLD responses to faces and bodies, and suggest the role of top-down signals in defining the category-specificity of the pattern representations in VOTC.

Acknowledgments

We thank Irene Jiang and Nicole Minkina for assistance with stimulus creation and data collection.

This work was supported by National Institute of Mental Health grant MH-005286 (G.M.) and the Yale University Faculty of Arts and Sciences Imaging Fund.

References

- Bode S, Haynes JD. Decoding sequential stages of task preparation in the human brain. *NeuroImage*. 2009; 45(2):606–13. DOI: 10.1016/j.neuroimage.2008.11.031 [PubMed: 19111624]
- Brainard DH. The psychophysics toolbox. *Spatial Vision*. 1997; 10:433–436. [PubMed: 9176952]
- Brosch T, Bar-David E, Phelps EA. Implicit race bias decreases the similarity of neural representations of black and white faces. *Psychological Science*. 2013; 24(2):160–6. DOI: 10.1177/0956797612451465 [PubMed: 23300228]
- Chang C, Lin C. LIBSVM: A Library for Support Vector Machines. *ACM Transactions on Intelligent Systems and Technology*. 2011; 2(3):1–27. DOI: 10.1145/1961189.1961199
- Coutanche MN. Distinguishing multi-voxel patterns and mean activation: why, how, and what does it tell us? *Cognitive, Affective & Behavioral Neuroscience*. 2013; 13(3):667–73. DOI: 10.3758/s13415-013-0186-2
- Cox RW. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research, an International Journal*. 1996; 29(3):162–73.
- Çukur T, Nishimoto S, Huth AG, Gallant JL. Attention during natural vision warps semantic representation across the human brain. *Nature Neuroscience*. 2013; 16(6):763–70. DOI: 10.1038/nn.3381 [PubMed: 23603707]
- De Gelder B, Van den Stock J. The Bodily Expressive Action Stimulus Test (BEAST). Construction and Validation of a Stimulus Basis for Measuring Perception of Whole Body Expression of Emotions. *Frontiers in Psychology*. 2011 Aug;2:181.doi: 10.3389/fpsyg.2011.00181 [PubMed: 21886632]
- Downing PE, Wiggett AJ, Peelen MV. Functional magnetic resonance imaging investigation of overlapping lateral occipitotemporal activations using multi-voxel pattern analysis. *The Journal of Neuroscience*. 2007; 27(1):226–33. DOI: 10.1523/JNEUROSCI.3619-06.2007 [PubMed: 17202490]

- Engell AD, McCarthy G. Selective attention modulates face-specific induced gamma oscillations recorded from ventral occipitotemporal cortex. *Journal of Neuroscience*. 2010; 30(26):8780–6. DOI: 10.1523/JNEUROSCI.1575-10.2010 [PubMed: 20592199]
- Fischl B, Salat D, Busa E, Albert M, Dieterich M, Haselgrove C, Dale. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron*. 2002; 33:341–355. [PubMed: 11832223]
- Hanke M, Halchenko YO, Sederberg PB, Hanson SJ, Haxby JV, Pollmann S. PyMVPA: A python toolbox for multivariate pattern analysis of fMRI data. *Neuroinformatics*. 2009; 7(1):37–53. DOI: 10.1007/s12021-008-9041-y [PubMed: 19184561]
- Harel A, Kravitz DJ, Baker CI. Task context impacts visual object processing differentially across the cortex. *Proceedings of the National Academy of Sciences of the United States of America*. 2014; 111(10):E962–71. DOI: 10.1073/pnas.1312567111 [PubMed: 24567402]
- Harry B, Williams MA, Davis C, Kim J. Emotional expressions evoke a differential response in the fusiform face area. *Frontiers in Human Neuroscience*. 2013 Oct.7:692.doi: 10.3389/fnhum.2013.00692 [PubMed: 24194707]
- Haxby JV, Gobbini MI, Furey ML, Ishai A, Schouten JL, Pietrini P. Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science (New York, NY)*. 2001; 293(5539):2425–30. DOI: 10.1126/science.1063736
- Haynes JD, Rees G. Decoding mental states from brain activity in humans. *Nature Reviews Neuroscience*. 2006; 7(7):523–34. DOI: 10.1038/nrn1931 [PubMed: 16791142]
- Jehee JFM, Brady DK, Tong F. Attention improves encoding of task-relevant features in the human visual cortex. *The Journal of Neuroscience*. 2011; 31(22):8210–9. DOI: 10.1523/JNEUROSCI.6153-09.2011 [PubMed: 21632942]
- Kaiser D, Strnad L, Seidl KN, Kastner S, Peelen MV. Whole person-evoked fMRI activity patterns in human fusiform gyrus are accurately modeled by a linear combination of face- and body-evoked activity patterns. *Journal of Neurophysiology*. 2014; 111(1):82–90. DOI: 10.1152/jn.00371.2013 [PubMed: 24108794]
- Kamitani Y, Tong F. Decoding the visual and subjective contents of the human brain. *Nature Neuroscience*. 2005; 8(5):679–85. DOI: 10.1038/nn1444 [PubMed: 15852014]
- Kanwisher N, McDermott J, Chun MM. The fusiform face area: a module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*. 1997; 17(11):4302–11. [PubMed: 9151747]
- Kim NY, Lee SM, Erlendsdottir M, McCarthy G. Discriminable spatial patterns of activation for faces and bodies in the fusiform gyrus. *Frontiers in Human Neuroscience*. 2014 Aug.8:1–12. DOI: 10.3389/fnhum.2014.00632 [PubMed: 24474914]
- Moore KS, Yi DJ, Chun MM. The effect of attention on repetition suppression and multivoxel pattern similarity. *Journal of Cognitive Neuroscience*. 2013; :1305–1314. DOI: 10.1162/jocn [PubMed: 23489143]
- Peelen MV, Downing PE. Selectivity for the human body in the fusiform gyrus. *Journal of Neurophysiology*. 2005; 93(1):603–8. DOI: 10.1152/jn.00513.2004 [PubMed: 15295012]
- Puce A, Allison T, Gore JC, McCarthy G. Face-sensitive regions in human extrastriate cortex studied by functional MRI. *Journal of Neurophysiology*. 1995; 74(3):1192–9. [PubMed: 7500143]
- Puce A, Allison T, McCarthy G. Electrophysiological studies of human face perception. III: Effects of top-down processing on face-specific potentials. *Cerebral Cortex (New York, NY)*. 1991. 1999; 9(5):445–58.
- Ratner KG, Kaul C, Van Bavel JJ. Is race erased? Decoding race from patterns of neural activity when skin color is not diagnostic of group boundaries. *Social Cognitive and Affective Neuroscience*. 2013; 8(7):750–5. DOI: 10.1093/scan/nss063 [PubMed: 22661619]
- Said CP, Moore CD, Engell AD, Haxby JV. Distributed representations of dynamic facial expressions in the superior temporal sulcus. *Journal of Vision*. 2010; 10:1–12. DOI: 10.1167/10.5.11.Introduction
- Soon CS, Namburi P, Chee MWL. Preparatory patterns of neural activity predict visual category search speed. *NeuroImage*. 2012; 66C:215–222. DOI: 10.1016/j.neuroimage.2012.10.036

- Tottenham N, Tanaka JW, Leon AC, McCarry T, Nurse M, Hare Ta, Nelson C. The NimStim set of facial expressions: judgments from untrained research participants. *Psychiatry Research*. 2009; 168(3):242–9. DOI: 10.1016/j.psychres.2008.05.006 [PubMed: 19564050]
- Weiner KS, Grill-Spector K. Sparsely-distributed organization of face and limb activations in human ventral temporal cortex. *NeuroImage*. 2010; 52(4):1559–73. DOI: 10.1016/j.neuroimage.2010.04.262 [PubMed: 20457261]
- Worsley KJ, Marrett S, Neelin P, Vandal AC, Friston KJ, Evans AC. A unified statistical approach for determining significant signals in images of cerebral activation. *Human Brain Mapping*. 1996; 4(1):58–73. DOI: 10.1002/(SICI)1097-0193(1996)4:1<58::AID-HBM4>3.0.CO;2-O [PubMed: 20408186]
- Zhang J, Kriegeskorte N, Carlin JD, Rowe JB. Choosing the rules: distinct and overlapping frontoparietal representations of task rules for perceptual decisions. *Journal of Neuroscience*. 2013; 33(29):11852–62. DOI: 10.1523/JNEUROSCI.5193-12.2013 [PubMed: 23864675]

Figure 1.

Experimental design. (A) Pictures of human faces and bodies with happy or fearful emotions were used in the main experiment, and computer-generated neutral faces and bodies were used in the localizer task. (B) The main experiment consisted of two task conditions: face-body discrimination and happy-fearful discrimination. Task conditions were indicated by text cues in the beginning of each task block. Ten stimuli were presented in a slow event-related design for each task block. Participants were instructed to make a button response for each stimulus based on the text cue (i.e., task condition).

Figure 2.

Classification accuracies. An SVM classifier was trained to discriminate neutral faces from neutral bodies of the independent localizer task, and then applied to the two main task conditions: face-body and emotional (happy-fearful) categorization tasks. The VOTC ROIs all showed higher accuracies in the happy- fearful task compared to the face-body task. Posterior visual areas (i.e., V1 and V2) did not show differences between tasks. Asterisks indicate p-values from paired t tests to examine differences between the two task conditions. (TOFC = Harvard-Oxford temporal-occipital fusiform cortex, ASAP = Atlas of Social Agent Perception mask for FFA, Fusiform = FreeSurfer fusiform gyrus ROI).

Figure 3.

Time course analysis. A pattern classifier trained using the localizer data set was applied to BOLD signals at each time point after stimulus onset. Task-related differences were observed in later response patterns (peaked 6–8 seconds after stimulus onset), indicating a temporal dissociation of initial face/body processing and task-based modulations. Data points are plotted at the mid-point of the 2 second acquisition period (i.e., TR = 2sec). The asterisks indicate significant differences in classification accuracy between conditions. The horizontal line at 50% indicates chance for classification, and both tasks exceed chance beginning at 2–4 seconds post-stimulus onset.

Figure 4.

Within-task classification accuracies. We trained and tested separate classifiers for each task condition. Similar to cross-task classification results shown in Figure 2, ROIs that include the fusiform gyrus exhibited task-related differences, whereas V1 and V2 did not show any differences between tasks. (TOFC = Harvard-Oxford temporal-occipital fusiform cortex, ASAP = Atlas of Social Agent Perception mask for FFA, Fusiform = FreeSurfer fusiform gyrus ROI).