J Neurophysiol 111: 82–90, 2014.
First published October 9, 2013; doi:10.1152/jn.00371.2013.

Whole person-evoked fMRI activity patterns in human fusiform gyrus are accurately modeled by a linear combination of face- and body-evoked activity patterns

Daniel Kaiser,[1]* Lukas Strnad,[1]* Katharina N. Seidl,[2] Sabine Kastner,[2] and Marius V. Peelen[1]
[1]*Center for Mind/Brain Sciences, University of Trento, Rovereto (TN), Italy; and* [2]*Department of Psychology, Princeton University, Princeton, New Jersey*

Submitted 21 May 2013; accepted in final form 2 October 2013

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. *J Neurophysiol* 111: 82–90, 2014. First published October 9, 2013; doi:10.1152/jn.00371.2013.—Visual cues from the face and the body provide information about another's identity, emotional state, and intentions. Previous neuroimaging studies that investigated neural responses to (bodiless) faces and (headless) bodies have reported overlapping face- and body-selective brain regions in right fusiform gyrus (FG). In daily life, however, faces and bodies are typically perceived together and are effortlessly integrated into the percept of a whole person, raising the possibility that neural responses to whole persons are qualitatively different than responses to isolated faces and bodies. The present study used fMRI to examine how FG activity in response to a whole person relates to activity in response to the same face and body but presented in isolation. Using multivoxel pattern analysis, we modeled person-evoked response patterns in right FG through a linear combination of face- and body-evoked response patterns. We found that these synthetic patterns were able to accurately approximate the response patterns to whole persons, with face and body patterns each adding unique information to the response patterns evoked by whole person stimuli. These results suggest that whole person responses in FG primarily arise from the coactivation of independent face- and body-selective neural populations.

fusiform face area; fusiform body area; extrastriate body area; object perception; category selectivity

THE GESTALT PSYCHOLOGISTS first pointed toward the interactive nature of the grouping of distinct, basic forms into wholes by formulating rules for these integration processes [Wertheimer 1938 (original work published 1923)]. Gestalt grouping not only has drastic effects on perception (e.g., as seen in visual illusions) but also influences neural processing and attentional competition (McMains and Kastner 2010; Spillmann 2006). A large body of literature focuses on low-level effects that are explained via the laws of Gestalt processing (for a recent review, see Wagemans et al. 2012). More recently, studies have started to investigate the integration of object parts into neural representations of whole objects. Typically, complex objects are composed of parts, which on their own might also be considered objects (e.g., the tires and frame of a car). Nevertheless, at a conceptual level objects are different from the combination of their parts. This raises the question of whether neural representations of complex objects are qualitatively different from the combined neural representations of their constituent parts. Single-cell recording studies of monkey inferotemporal cortex (IT) have provided evidence for nonlinear neural response summation to feature conjunctions (Perrett et al. 1982; Tanaka et al. 1991) but also have provided evidence for linear response summation of features of shapes and faces (Freiwald et al. 2009; Sripati and Olsen 2010). It is not yet clear to what extent these neural underpinnings of feature grouping in IT apply to the grouping of larger parts into complex objects in the human brain.

In the present study, we investigate a special case of the grouping of components into a whole: the composition of the human body. A person consists of two main parts: a body and a face. These parts are, however, grouped together into the percept of a single person (McArthur and Baron 1983). Despite the natural and seamless integration of the body and face into a person, imaging studies have typically focused on investigating the neural correlates of either face or body perception (Downing et al. 2001; Kanwisher et al. 1997). It is undoubtedly the case that faces and bodies convey partly different kinds of information. For example, faces are more informative about an individual's identity, whereas bodies carry more information about actions that the individual performs. Nevertheless, information extracted from faces and bodies is rapidly integrated for judgments about expressed emotions, intentions, and attractiveness (Aviezer et al. 2012; de Gelder 2006; Meeren et al. 2005; van den Stock et al. 2007). One explanation for this integration is that face and body representations are combined into a larger perceptual unit (Aviezer et al. 2012). An alternative explanation is that the integration at the behavioral level is the result of a weighted sum of body and face information, each of which is processed more or less independently in sensory systems (Ekman et al. 1982; Wallbott 1998).

At the neural level, the perception of faces and bodies has been associated with selective responses in high-level visual cortex. Interestingly, both faces and bodies evoke selective responses in the right fusiform gyrus (FG) at approximately the same anatomic location (Kanwisher et al. 1997; Peelen and Downing 2005). It has been suggested that the body-selective FG response reflects a postperceptual recruitment of face-selective mechanisms (Meeren et al. 2013), possibly caused by mental imagery of a face when a blurred face is seen on top of a body (see Cox et al. 2004). However, several studies using headless bodies have provided evidence for separable face and body selectivity in FG (Peelen et al. 2006; Schwarzlose et al. 2005), suggesting that body-evoked responses cannot be fully explained by activity in face-selective neurons. The overlap between face- and body-evoked FG responses partly persists at high spatial resolution (Schwarzlose et al. 2005), raising the

*D. Kaiser and L. Strnad contributed equally to this work.

Address for reprint requests and other correspondence: M. V. Peelen, Center for Mind/Brain Sciences, Univ. of Trento, 38068 Rovereto (TN), Italy (e-mail: marius.peelen@unitn.it).

possibility that FG contains holistic person representations. Face- and body-selective responses are also found posterior to FG, with faces activating the "occipital face area" (OFA; Gauthier et al. 2000) and bodies activating the "extrastriate body area" (EBA; Downing et al. 2001). Unlike the face- and body-selective regions in FG, these more posterior regions are anatomically separate from each other and are thus less likely candidates for housing integrated person representations.

In the present study, we asked how FG responses to whole persons relate to FG responses to faces and bodies presented in isolation. Our analysis approach was based on previous studies using intracranial recordings in monkeys and humans (Agam et al. 2010; Zoccolan et al. 2005, 2007) and functional magnetic resonance imaging (fMRI) in humans (MacEvoy and Epstein 2009; Reddy et al. 2009). These studies have shown that visual cortex activity patterns in response to displays of multiple unrelated objects can be very accurately modeled by a linear combination of response patterns evoked by the individual objects (but see e.g., Heuer and Britten 2002, for a nonlinear relationship). The properties of this linear relationship are not yet clear. Although a number of studies suggest the mean as the best operator for combining single-object responses (MacEvoy and Epstein 2009; Zoccolan et al. 2005), others provide evidence for linear models with different weights depending on the strength of the response to the individual objects (Baeck et al. 2013) or even a pure MAX operator, only taking the response of the strongest stimulus into account (Gawne and Martin 2002). By contrast to the present study, previous fMRI studies investigating multiple object responses have used unrelated (e.g., pairs of cars and chairs; MacEvoy and Epstein 2009) and always clearly separate objects, even in the case of action grouping (Baeck et al. 2013).

In this study, we tested how FG responses to whole persons relate to responses to a person's main parts (face, body) presented in isolation. To do so, we identified a person-selective FG region in each participant and tested how person-evoked responses in this region relate to responses evoked by faces and bodies shown in isolation. We considered two possible outcomes (Fig. 1). First, FG response patterns to whole persons might be well approximated by a linear combination of response patterns to isolated face- and body-evoked response patterns, similar to response averaging previously observed for unrelated object pairs presented simultaneously (MacEvoy and Epstein 2009). This outcome would suggest that FG encodes

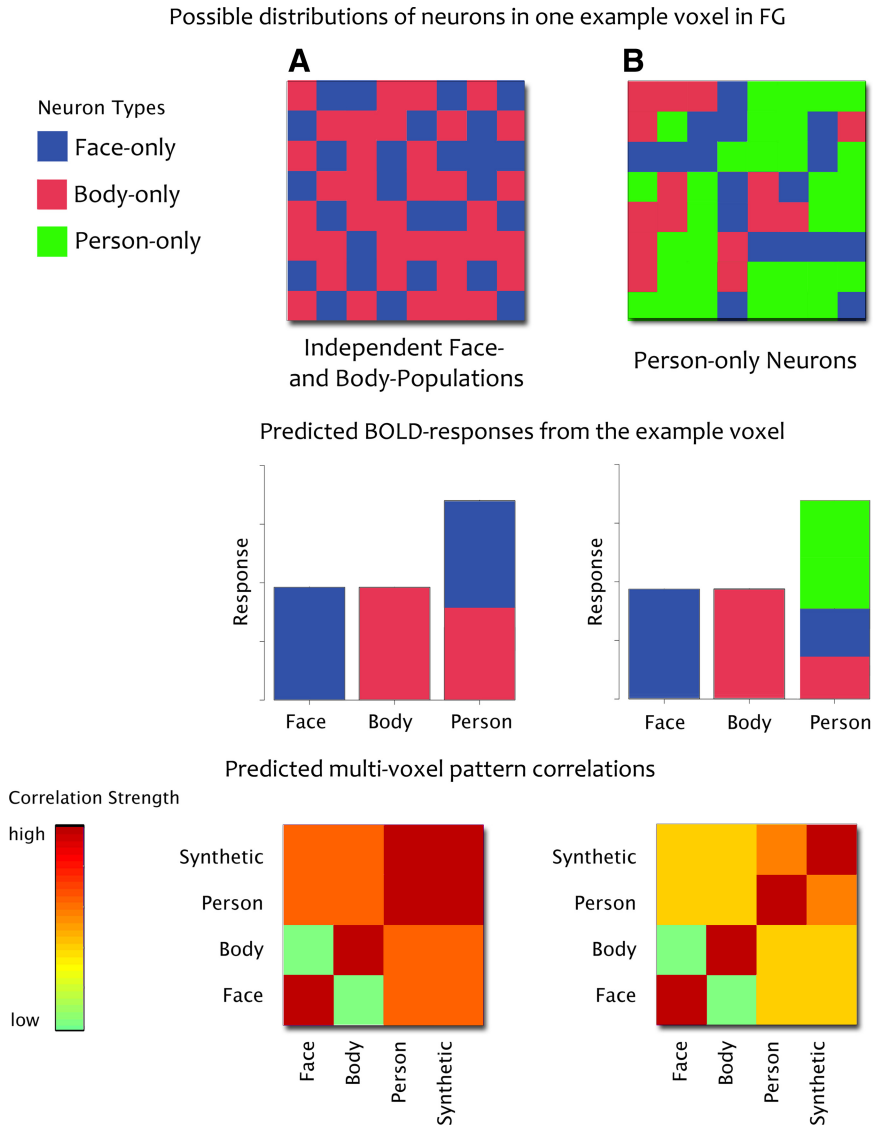

Fig. 1. Two hypothetical scenarios for the nature of whole person responses in person-selective fusiform gyrus (pFG). *A*: person perception activates distinct but intermixed face- and body-selective neural populations. *B*: person perception activates, in addition to face- and body-selective neurons, neurons that are exclusively responsive to whole persons. These scenarios make similar predictions for overall blood oxygen level-dependent (BOLD) responses but different predictions for multivoxel pattern correlations. Specifically, in the scenario shown in *A*, person-evoked responses are accurately modeled by a linear combination of face- and body-evoked responses, whereas in the scenario shown in *B* this would not be the case.

persons in a part-based manner. Alternatively, FG response patterns to whole persons might be distinct from a linear combination of face- and body-evoked response patterns, reflecting an integrated (holistic) person representation (Fig. 1*B*).

METHODS

Participants. Thirteen participants participated in the study (6 female; mean age: 24.6 yr, range: 19–32 yr). All participants had no history of neurological or psychiatric diseases and normal or corrected-to-normal vision. The study was approved by the Institutional Review Panel of Princeton University. Participants gave written consent to their participation.

Stimuli and paradigm. Stimuli consisted of grayscale exemplars from six object categories (trees, cars, headless bodies, faces, whole persons, outdoor scenes) depicted on a white background. Blocks of tree and scene stimuli were related to another experiment (Seidl et al. 2012) and are modeled but not analyzed in this report. See Fig. 2 for stimulus examples. Stimuli subtended ~12° × 12° visual angle. Stimuli were presented centrally, with a random jitter of 1°. Stimuli were presented on a translucent screen located at the end of the scanner bore with the use of a projector outside the scanner room. Participants could see the screen via a tilted mirror mounted to the head coil.

Participants performed four runs. One run consisted of 29 blocks of 14 s each. *Blocks 1*, *8*, *15*, *22*, and *29* were fixation-only blocks, in which a central fixation cross was presented on a white background. One block of each condition was presented between two fixation epochs. For the first half of the experiment, the order of blocks was determined randomly. The block order for the second half of the run was the mirror order of the first half of the run to equate the mean serial position of each condition within the run. Each block contained 10 sequential presentations of exemplars from one category. Stimuli were shown for 400 ms and followed by a white fixation-only screen of equal duration. Subjects performed a one-back repetition detection task. Stimulus repetitions occurred twice during each block.

Data acquisition. Data were acquired using a Siemens Allegra 3T MRI Scanner (Erlangen, Germany) equipped with a standard head coil (Nova Medical). Functional images were obtained in T2*-weighted echo-planar imaging sequences (TR = 2,000 ms, TE = 30 ms, 90° flip angle, 34 slices, axial orientation, 3 × 3 × 3-mm voxel size, 1-mm gap, field of view: 192 mm, 64 × 64 in-plane matrix). In addition, a high-resolution T1-weighted scan was acquired using a MPRAGE sequence (TR = 2,500 ms, TE = 4.3 ms, 8° flip angle, 1-mm voxel size, 256 × 256 in-plane matrix).

fMRI data preprocessing. Data analysis was performed using the AFNI software package (Cox 1996) and custom-written software in Python, MATLAB, and R. For each subject, a high-resolution anatomic image was aligned with the first volume of the first functional run. Functional data were corrected for slice timing and head motion, and transformed into Talairach space using parameters derived from the warping of the high-resolution anatomic image. The time series in each voxel was scaled to a common mean. No spatial smoothing was applied to the data used for multivoxel pattern analysis (MVPA). However, we also created a copy of the data that was spatially smoothed with a Gaussian kernel (4-mm full-width half-maximum). This version of the data was used for region of interest (ROI) definition and for univariate analyses.

Univariate analyses. To ensure independence of data used to define ROIs from data used to compare blood oxygen level-dependent (BOLD) responses to the different stimulus categories in these ROIs, the four runs for every participant were separated into two groups. Two runs were used for ROI definition, and BOLD responses associated with the stimulus categories were then estimated from the other two runs. The results were then averaged over all six possible splits of the four runs into two groups. BOLD responses were estimated using the general linear model (GLM). The GLM included regressors for each stimulus category (created by convolving the boxcar function indicating when each stimulus type was presented with the double-gamma canonical hemodynamic response function), as well as six nuisance regressors modeling head motion. For each run, four additional regressors were included to account for mean, linear, quadratic, and cubic trends in the data.

Because our hypotheses (Fig. 1) were about how person-evoked responses relate to face- and body-evoked responses, we investigated responses in ROIs that were most strongly responsive to the person stimuli and that could thus potentially house person-only representations (Fig. 1*B*). We did not localize face- or body-selective regions, because these regions are less likely to house person-only neurons. Two right-hemisphere person-selective ROIs, one in FG and one in

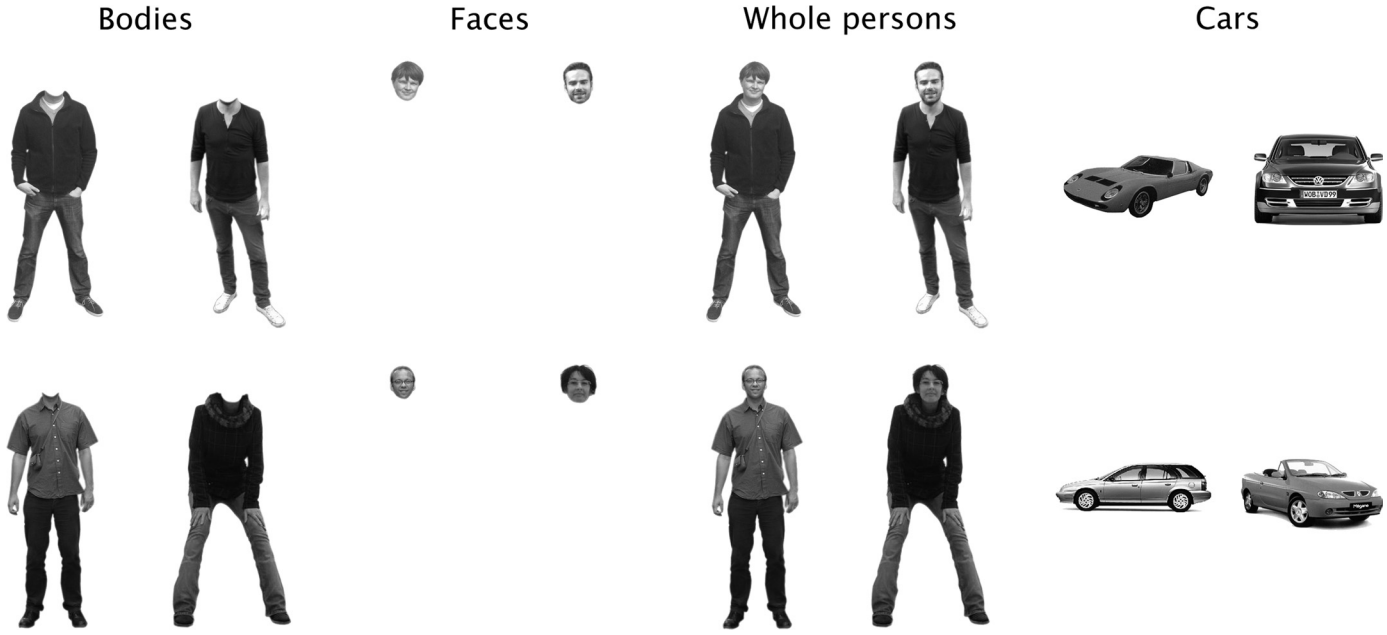

Fig. 2. Example stimuli. Whole persons were always a compound of one of the faces and one of the bodies presented. Stimuli were presented for 400 ms each in blocks of 14 s.

lateral occipitotemporal cortex (LOTC), were functionally defined using an objective (automated) method, selecting all voxels that showed significant person selectivity (person > cars; $P < 0.001$, uncorrected) within box-shaped FG ($30 < x < 50$, $55 < y < 30$, $25 < z < 10$; Peelen et al. 2009) and LOTC regions ($35 < x < 60$, $80 < y < 55$, $10 < z < 15$; Peelen et al. 2009). The person-selective FG (pFG) ROI contained on average 38 voxels (mean Talairach coordinates: $x = 41.0$, $y = -45.4$, $z = -15.5$), and the person-selective LOTC (pLOTC) ROI contained on average 67 voxels (mean Talairach coordinates: $x = 46.5$, $y = -68.6$, $z = 3.6$). Univariate ROI analyses were based on the mean value of the regressor coefficients (betas) for each stimulus category, estimated from the second half of the data in the split and extracted from the regions defined above. In addition to the functionally defined regions, we also anatomically defined an occipital cortex (OC) region corresponding to primary visual cortex (V1) using the Talairach atlas (Brodmann area 17). For congruency with the functional ROIs, only right-hemisphere OC was analyzed.

MVPA. The procedure for ROI definition for MVPA was similar to that used for the univariate analysis. However, for MVPA, we used all four runs to define ROIs. Participants whose FG ROI contained fewer than 15 voxels were excluded from the analyses. Two participants were excluded on the basis of this criterion. For MVPA, each condition was modeled by a separate regressor for each run. All other parameters were identical to those of the GLMs used for univariate analyses. Multivoxel patterns for subsequent analyses were created from the beta values of the category regressors. A single pattern would consist of the beta values of a single regressor extracted from all voxels of a given ROI.

MVPA was performed in a split-half manner. All analyses were performed for every possible two-way split of the four runs (6 possible splits altogether), and these results were combined by averaging. In each half, a synthetic person pattern was computed as the average of the face pattern and the body pattern. The average was taken following previous work on object pairs (MacEvoy and Epstein 2009). To assess the similarity of response patterns, we computed correlations between all possible pairs of patterns A–B such that A comes from one of the halves and B from the other. Correlations between pairs that included the same categories stemming from different runs were averaged (e.g., A_face–B_body and A_body–B_face correlations were averaged). For present purposes, only the correlations involving face, body, person, and synthetic conditions were relevant. Before group-level statistical analysis was carried out, all correlations calculated on the single-subject level were Fisher transformed.

To test for the optimal weights for combining face and body patterns to approximate the person pattern, we obtained the values of the coefficients maximizing the correlation between a linear combination of the face and body patterns with the person pattern through a simple optimization procedure. Since for the calculation of correlation only the relative magnitude of the coefficients associated with the face and body patterns matters, we imposed a further constraint that the two coefficients must add up to 1. Note that, therefore, this analysis only reveals information about the relative contribution of face and body patterns, but not about their absolute contributions. We denote the face coefficient α and the body coefficient β, but note that β is determined by α as $\beta = (1 - \alpha)$. The "synthetic" pattern approximating the person pattern was thus equal to α * face pattern + β * body pattern. To compute the coefficients, data from all four runs were first combined by averaging. We then tested, separately for each participant, for the value of α yielding the highest correlation between the synthetic face-body pattern and the person pattern. Note that this analysis was done separately from the analyses involving the split-half correlations between the conditions. In those analyses, the α was always set to 0.5.

RESULTS

Univariate analyses. We first compared the mean BOLD signal elicited by faces, bodies, and persons within pFG, pLOTC, and anatomically defined OC (Fig. 3). In each of the three ROIs, there was a main effect of stimulus category [all $F(3,36) > 8.04$, $P < 0.001$]. In pFG, persons elicited higher responses than faces, bodies, and cars [all $t(12) > 5.00$, $P < 0.001$]. Faces and bodies each elicited higher responses than cars [both $t(12) > 6.70$, $P < 0.001$], whereas the difference between face- and body-evoked responses did not reach significance [$t(12) = 2.09$, $P = 0.058$]. pLOTC responses were higher for faces, bodies, and persons compared with cars [all $t(12) > 3.00$, $P < 0.011$]. However, in contrast to pFG, responses in pLOTC were higher for persons and bodies than for faces [both $t(12) > 8.60$, $P < 0.001$], and the region did not differentiate between bodies and persons [$t(12) = 0.60$, $P = 0.56$]. The lack of difference between the body and person categories suggests that pLOTC is only sensitive to body information present in both body and person stimuli, and that it is not influenced by the additional face information present only in person stimuli (see Morris et

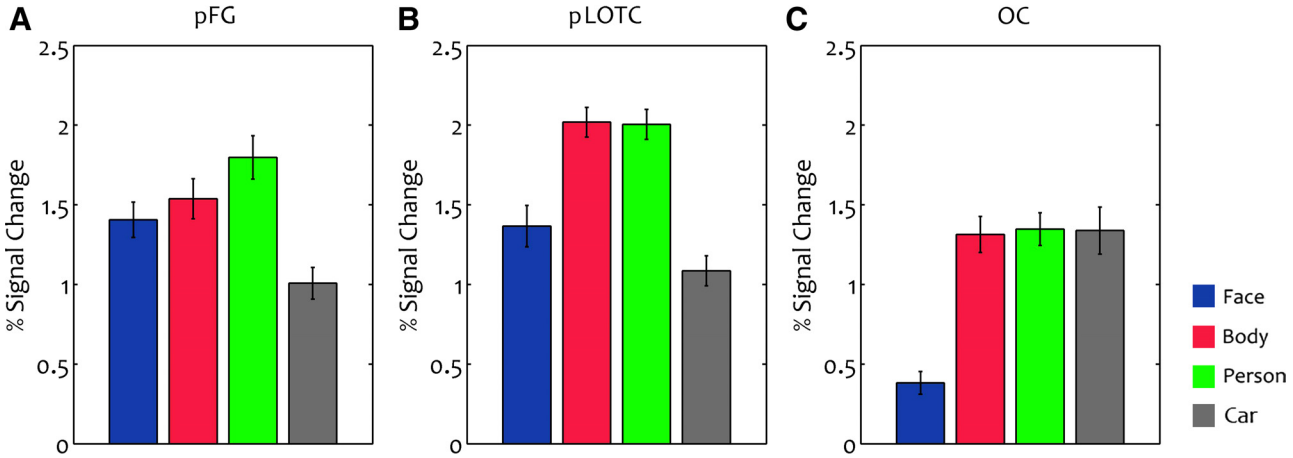

Fig. 3. Univariate results. BOLD signal changes in response to stimuli of different categories were averaged over all voxels in each region of interest (ROI). In the pFG ROI, the person stimuli elicited the highest response, whereas faces and bodies evoked comparable responses. In person-selective lateral occipitotemporal cortex (pLOTC), by contrast, the responses to the person stimuli did not differ significantly from the responses to the body stimuli, which were both higher than responses to faces. Occipital cortex (OC) showed weaker responses for faces, because of their overall smaller retinal sizes. Error bars indicate SE.

al. 2006 for evidence that the presence of a face may even reduce activity in body-selective cortex). In OC, the face condition led to lower responses than all other conditions [all $t(12) > 6.73$, $P < 0.001$], with no difference between bodies, persons, and cars [all $t(12) < 0.92$, $P > 0.38$]. The low response to faces is likely due to the smaller retinal size of the face condition relative to the other conditions (Fig. 2).

MVPA. In the MVPA, we aimed to test hypotheses about how person-evoked response patterns in pFG relate to face- and body-evoked response patterns (Fig. 1). To assess the degree of specificity of the effects found in pFG, we also analyzed results in pLOTC and OC.

Previous studies investigating neural responses in the ventral stream to displays consisting of multiple distinct objects report that these responses (measured by single-cell recordings and multivoxel fMRI patterns) can be very accurately approximated as a simple average of responses to the individual constituent objects (MacEvoy and Epstein 2009; Zoccolan et al. 2005). Because the face and the body can be thought of as constituent objects of a person image, we hypothesized that it may be possible to approximate the person pattern in FG as a linear combination (such as the average) of the face and body patterns. To test directly for the optimal coefficients for approximating person patterns, we computed the correlation between synthetic and person patterns as a function of the coefficients (Fig. 4). For every subject and all three ROIs, we used an optimization procedure to estimate the values of the linear coefficients for which this correlation is the highest (see MATERIALS AND METHODS for details). We then tested whether, at the group level, these coefficients are different from 0.5. In pFG, this difference was not significant [$t(10) = 0.034$, $P = 0.74$], indicating that the linear combination of the face and body patterns that most successfully approximates the person pattern comprised equal weights for faces and bodies. By contrast, in both pLOTC and OC, the body coefficient was significantly greater than 0.5 [pLOTC: $t(10) = 6.60$, $P < 0.001$; OC: $t(10) = 4.35$, $P = 0.0014$].

By combining the face and body patterns through a linear combination with the coefficients equal to 0.5, we next created synthetic multivoxel patterns approximating the person patterns in FG for every participant. We then compared multivoxel pattern correlations between these synthetic patterns and the patterns associated with faces, bodies, and whole persons (Fig. 5). We found that the synthetic pattern was better at approximating the person pattern than were the individual face or body patterns. In particular, we tested whether the person pattern was reliably more similar to the synthetic pattern than to the face or body pattern alone. A one-way ANOVA including the person-synthetic, person-face, and person-body correlations yielded a main effect of category pair [$F(2,20) = 5.11$, $P = 0.016$]. In follow-up tests, we found a positive difference between person-synthetic and person-face correlations [$t(10) = 4.42$, $P = 0.0013$] and also between person-synthetic and person-body correlations [$t(10) = 3.22$, $P = 0.0091$]. Therefore, the high degree of similarity between the synthetic pattern and the person pattern can be ascribed to the combination of distinct pieces of information contained in the individual face and body patterns.

In addition, we tested whether the synthetic pattern could at all be distinguished from the person pattern in pFG. We computed a one-way ANOVA including the person-person, person-synthetic, and synthetic-synthetic correlations. This ANOVA did not yield a significant main effect of category pair [$F(2,20) = 1.36$, $P = 0.28$], suggesting that the synthetic pattern is statistically indistinguishable from the person pattern. To underline this point, we also directly tested for a difference between the person-person and person-synthetic correlations [$t(10) = 1.14$, $P = 0.28$] and between the synthet-

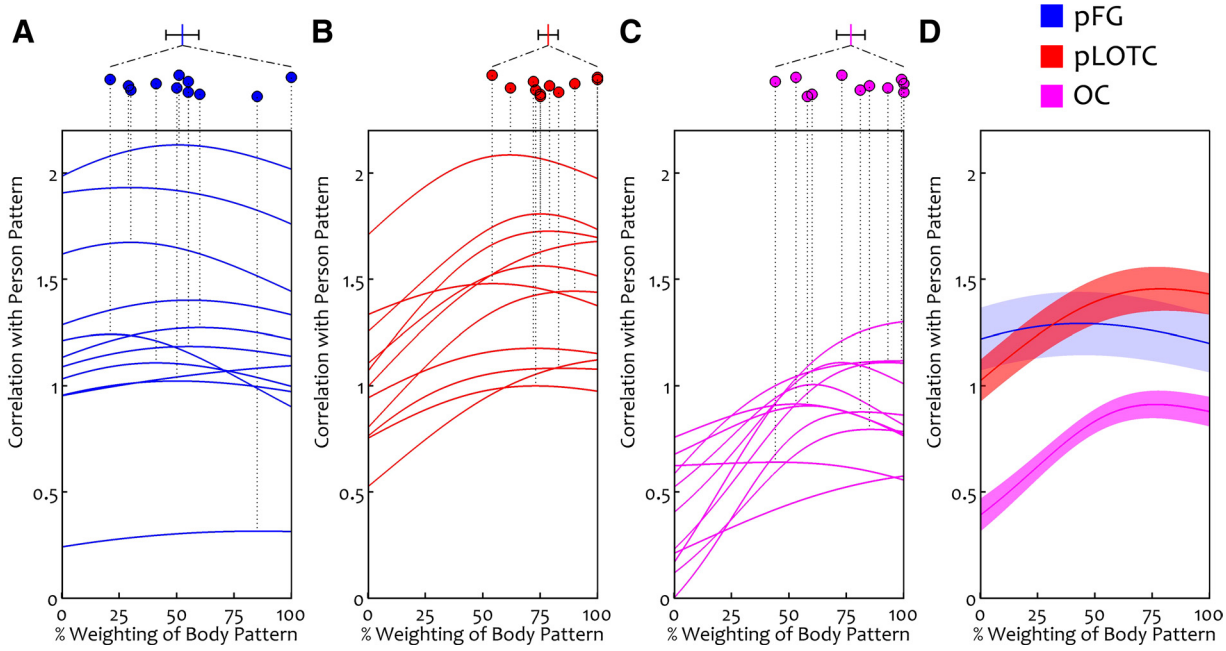

Fig. 4. Correlations between the person pattern as a function of different linear combinations of face and body patterns in pFG, pLOTC, and OC. Single-subject estimates for the optimal coefficients are represented by circles in *A–C*, with the mean and SE of the optimal estimate represented above; *D* shows the mean results across participants for all regions (indicated by colors corresponding to the circles in *A–C*). For pFG, the highest correlation was obtained when the weights for face and body patterns were equal. By contrast, the person patterns in pLOTC and OC were best approximated by a synthetic pattern that consisted almost exclusively of the body pattern. All correlations shown were Fisher transformed.

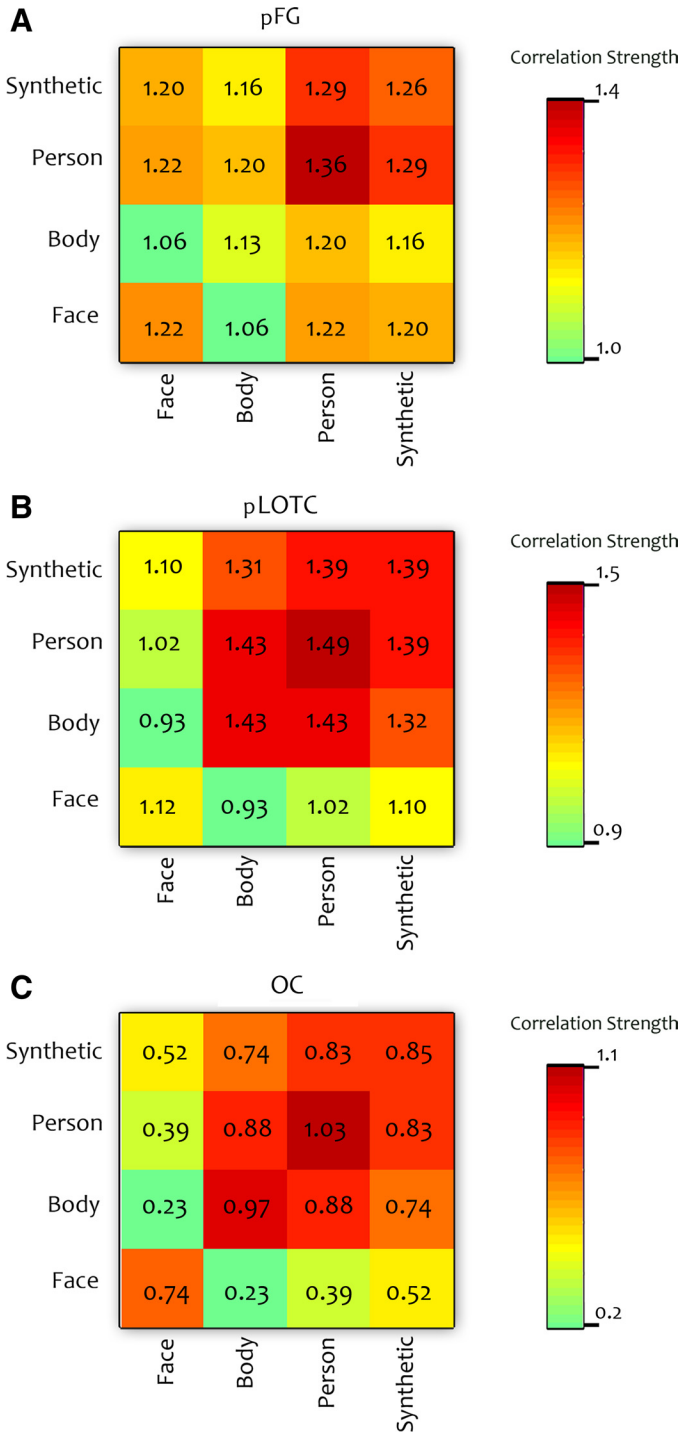

Fig. 5. Results of multivoxel pattern analysis. *A–C*: correlations in pFG, pLOTC, and OC, respectively, between the multivoxel patterns associated with the face, body, and person stimuli, as well as the synthetic patterns created by averaging face- and body-related patterns. The pFG results best match the scenario shown in Fig. 1*A*. All correlations shown were Fisher transformed.

ic-synthetic and person-synthetic correlations [$t(10) = 1.04$, $P = 0.32$], neither of which yielded a significant result. Taken together, these analyses indicate that the person pattern in pFG can be accurately modeled as a linear combination of the face and body patterns, with equal weights for both categories.

These results were specific to pFG. Unlike in pFG, the person patterns in pLOTC and OC were equally correlated with the body pattern as with the synthetic pattern [LOTC: $t(10) = 1.25$, $P = 0.24$; OC: $t(10) = 0.95$, $P = 0.37$], suggesting that the synthetic pattern did not contain any meaningful information but the one already included in the body pattern. Furthermore, unlike in pFG, the synthetic patterns were distinguishable from the person patterns: whereas the synthetic-synthetic correlation was equal to the synthetic-person correlation [pLOTC: $t(10) = 0.083$, $P = 0.93$; OC: $t(10) = 0.36$, $P = 0.73$], the person-person correlation was significantly higher than the person-synthetic correlation [pLOTC: $t(10) = 2.71$, $P = 0.022$; OC: $t(10) = 5.64$, $P < 0.001$].

DISCUSSION

In the present study, we tested how FG responses to whole persons relate to responses to a person's main parts (face, body) presented in isolation. In a univariate analysis, we found that person-selective FG was more responsive to pictures of whole persons than to pictures of faces and bodies shown in isolation. In contrast to FG, we observed comparable responses to whole persons and isolated bodies in person-selective LOTC and anatomically defined OC. Interestingly, multivoxel pattern analysis showed that FG response patterns to whole persons could be accurately modeled by a linear combination of response patterns evoked by isolated faces and bodies, with equal weights for both categories. Taken together, these results suggest that FG encodes persons in a part-based rather than an integrated (holistic) manner.

Distinct representations of faces and bodies in FG. The finding that response patterns to faces and bodies contributed unique information to whole person-evoked response patterns provide further evidence that face and body representations in FG are separable, thus arguing against the possibility that body-selective FG responses are fully accounted for by late activation of face-selective neural populations (Meeren et al. 2013). These conclusions are in line with several previous findings. For example, high-resolution fMRI allows for isolating FG voxels that are only selective for faces or for bodies (Schwarzlose et al. 2005, 2008), although an overlap between face- and body-selective responses persists even at high spatial resolution. Other studies have used MVPA in FG to show that body-selective response patterns, but not face-selective response patterns, correlate with response patterns evoked by point-light biological motion displays of whole body actions (Atkinson et al. 2012; Peelen et al. 2006) and with activity evoked by emotional (vs. neutral) body expressions (Peelen et al. 2007). Developmental work has shown that face- and body-selective fusiform regions develop along different trajectories (Peelen et al. 2009). In the monkey, adjacent and partly overlapping face- and body-selective regions have also been identified using fMRI (Pinsk et al. 2005, 2009; Tsao et al. 2003). Single-cell recording studies in monkey IT have found neural populations that respond selectively to either faces (Desimone et al. 1984; Tsao et al. 2006) or body parts (e.g., hands; Gross et al. 1972). Furthermore, the pattern of responses across hundreds of IT neurons showed distinct clusters for faces, hands, and bodies (Kiani et al. 2007). Together with the present findings, these results suggest that representations of faces and bodies in FG are likely to be largely distinct at the neuronal level. This raises the question of why faces and bodies are represented so close together. One possibility is that

the consistent co-occurrence and mutual relevance of faces and bodies results in these categories being represented nearby, because this minimizes axon length and thus optimizes efficient interactions between these representations (e.g., Aflalo and Graziano 2011; Van Essen 1997). Another possibility is that innate anatomic connectivity patterns between visual cortex and downstream areas involved in social cognition are strongest in the part of FG where representations of faces and bodies are located, thus optimizing this part of visual cortex for representing socially relevant stimuli (Mahon and Caramazza 2011; Simmons and Martin 2012). Finally, it has been proposed that neurons in FG may be well suited to perform particular computations such as subordinate-level discrimination (Tarr and Gauthier 2000), which is highly relevant for both face and body processing.

No evidence for holistic person representations in FG. The present results showed that response patterns to whole persons could be accurately approximated by a linear combination of response patterns to faces and bodies shown in isolation. This result is most in line with the scenario depicted in Fig. 1*A*, in which person-evoked responses are accounted for by concurrent activity in face- and body-selective neural populations, thus without assuming additional populations tuned preferentially or exclusively to whole persons. It should be noted, however, that the absence of evidence for holistic person representations in FG does not conclusively prove that such populations do not exist in this region. First, it may be that such populations are largely outnumbered by populations of face- and body-selective neurons and are therefore hard to detect. For example, a study recording from cells in monkey superior temporal sulcus (Wachsmuth et al. 1994) found that of the cells responding to whole persons, only 17% responded exclusively to whole persons (the other cells responded to either the face or the body, or to both). Second, person-selective neurons could be evenly distributed across voxels, unlike face- and body-selective neurons. In this case, these neurons would not form a pattern across voxels and therefore would not be captured by MVPA.

Our conclusions may appear to contradict those of a recent article that reported person-selective fMRI adaptation effects in right FG (Schmalzl et al. 2012). In this study, participants were shown two subsequently presented person pictures. The second picture could be an exact repetition of the first picture, a picture of a different person, or a partial repetition of the first picture, with either the face or the body being that of a different individual (with the other part repeating). Results showed that fMRI adaptation in right FG to exact-picture repetitions was larger than the sum of the adaptation effects to face-only and body-only repetitions. The interpretation of these results was that FG contains representations of whole persons in addition to representations of faces and bodies; if the whole person repeats, this would adapt neurons coding for the whole person in addition to neurons coding for the face and body, thereby explaining the superadditive adaptation effects. Alternatively, however, the partial change conditions (in which only the body or the face changed) may have given higher activity (i.e., less adaptation) partly because they attracted more attention than the condition in which the exact same picture was repeated. Differences in attention and expectation across conditions may complicate the interpretation of fMRI adaptation studies (e.g., Henson et al. 2003; Mur et al. 2010; Summerfield et al. 2008), and such effects may partly account for the results of the Schmalzl et al. (2012) study: First, a change of either a face or a body (with the other part unchanged) is highly unnatural and thus may increase overall attention and/or direct attention to the part that changed, thereby drawing attention away from the repeated part. Second, the task of the participants was to detect specific target faces and bodies occurring on 10% of trials. When the exact same picture repeated, participants could quickly establish, based on low-level cues, that the response (target or no target) to the second picture was the same as the response to the first picture, thus no longer requiring the identification of the face or body. More generally, the different conclusions reached by the Schmalzl et al. (2012) study and our study could be due to differences in methodological approaches (for differences between MVPA and fMRI adaptation results, see e.g., Drucker and Aguirre 2009; Epstein and Morgan 2012). Another important difference between our study and the study of Schmalzl et al. (2012) is that their experiment was focused on finding adaptation in neurons coding for specific identities of faces, bodies, and whole persons (whole persons were presented on every trial). By contrast, our study looked at the distribution of category-level selectivity to faces, bodies, and their combined presentation.

The absence of evidence for holistic person representations in FG in our study raises the question of whether information from these stimuli is integrated elsewhere in the brain, and if so, where. Candidate regions for person representations that are abstracted away from visual cues are the medial prefrontal cortex and posterior superior temporal cortex. These regions have been implicated in social cognition, mentalizing, and theory-of-mind (Adolphs 2009; Gallagher and Frith 2003; Mitchell et al. 2002; Saxe and Kanwisher 2003) and have been shown to contain representations of perceived mental states that generalize across face and body cues (Peelen et al. 2010). Future work is needed to test whether these regions may underlie the body-face integration effects observed in behavioral studies (Aviezer et al. 2012; de Gelder 2006; Ghuman et al. 2010; Meeren et al. 2005; van den Stock et al. 2007) or whether these effects instead arise as a result of rapid interactions between distinct but nearby representations of faces and bodies in fusiform gyrus.

GRANTS

This work was supported by National Institutes of Health Grants R01 EY017699 and R01 MH064043 and National Science Foundation Grant BCS-1025149.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the authors.

AUTHOR CONTRIBUTIONS

D.K., L.S., and M.V.P. analyzed data; D.K., L.S., and M.V.P. interpreted results of experiments; D.K. prepared figures; D.K., L.S., and M.V.P. drafted manuscript; D.K., L.S., K.N.S., S.K., and M.V.P. approved final version of manuscript; K.N.S., S.K., and M.V.P. conception and design of research; K.N.S. and M.V.P. performed experiments; K.N.S., S.K., and M.V.P. edited and revised manuscript.

REFERENCES

Adolphs R. The social brain: neural basis of social knowledge. *Annu Rev Psychol* 60: 693–716, 2009.

Aflalo TN, Graziano MS. The organization of the macaque extrastriate visual cortex re-examined using the principle of spatial continuity of function. *J Neurophysiol* 105: 305–320, 2011.

Agam Y, Liu H, Papanastassiou A, Buia C, Golby AJ, Madsen JR, Kreiman G. Robust selectivity to two-object images in human visual cortex. *Curr Biol* 20: 1–8, 2010.

Atkinson AP, Vuong QC, Smithson HE. Modulation of the face- and body-selective visual regions by the motion and emotion of point-light face and body stimuli. *Neuroimage* 59: 1700–1712, 2012.

Aviezer H, Trope Y, Todorov A. Holistic person processing: faces with bodies tell the whole story. *J Pers Soc Psychol* 103: 20–37, 2012.

Baeck A, Wagemans J, Op de Beeck HP. The distributed representation of random and meaningful object pairs in human occipitotemporal cortex: the weighted average as a general rule. *Neuroimage* 70: 37–47, 2013.

Cox D, Meyers E, Sinha P. Contextually-evoked object-specific responses in human visual cortex. *Science* 304: 115–117, 2004.

Cox RW. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput Biomed Res* 29: 162–173, 1996.

de Gelder B. Towards the neurobiology of emotional body language. *Nat Rev Neurosci* 7: 242–249, 2006.

Desimone R, Albright TD, Gross CG, Bruce C. Stimulus selective properties of inferior temporal neurons in the macaque. *J* Neurosci 4: 2051–2062, 1984.

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

Drucker DM, Aguirre GK. Different spatial scales of shape similarity representation in lateral and ventral LOC. *Cereb Cortex* 19: 2269–2280, 2009.

Ekman P, Friesen W, Ellsworth P. What are the relative contributions of facial behavior and contextual information to the judgment of emotion. In: *Emotion in the Human Face*, edited by Ekman P. New York: Cambridge University Press, 1982.

Epstein RA, Morgan LK. Neural responses to visual scenes reveals inconsistencies between fMRI adaptation and multivoxel pattern analysis. *Neuropsychologia* 50: 530–543, 2012.

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

Gallagher HL, Frith CD. Functional imaging of 'theory of mind'. *Trends Cogn Sci* 7: 77–83, 2003.

Gauthier I, Tarr MJ, Moylan J, Skudlarski P, Gore JC, Anderson AW. The fusiform "face area" is part of a network that processes faces at the individual level. *J Cogn Neurosci* 12: 495–504, 2000.

Gawne TJ, Martin JM. Responses of primate visual cortical V4 neurons to simultaneously presented stimuli. *J Neurophysiol* 88: 1128–1135, 2002.

Ghuman AS, McDaniel JR, Martin A. Face adaptation without a face. *Curr Biol* 20: 32–36, 2010.

Gross CG, Rocha-Miranda CE, Bender DB. Visual properties of neurons in inferotemporal cortex of the macaque. *J Neurophysiol* 35: 96–111, 1972.

Henson RNA. Neuroimaging studies of priming. *Prog Neurobiol 70*: 53–81, 2003.

Heuer HW, Britten KH. Contrast dependence of response normalization in area MT of the rhesus monkey. *J Neurophysiol* 88: 3398–3408, 2002.

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

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

MacEvoy SP, Epstein RA. Decoding the representation of multiple simultaneous objects in human occipitotemporal cortex. *Curr Biol* 19: 943–947, 2009.

Mahon BZ, Caramazza A. What drives the organization of object knowledge in the brain? *Trends Cogn Sci* 15: 97–103, 2011.

McArthur LZ, Baron RM. Toward an ecological theory of social perception. *Psychol Rev* 90: 215–238, 1983.

McMains SA, Kastner S. Defining the units of competition: influences of perceptual organization on competitive interactions in human visual cortex. *J Cogn Neurosci* 22: 2417–2426, 2010.

Meeren HKM, de Gelder B, Ahlfors SP, Hämäläinen MS, Hadjikhani N. Different cortical dynamics in face and body perception: an MEG study. *PLoS One* 8: e71408, 2013.

Meeren HKM, van Heijnsbergen CCRJ, de Gelder B. Rapid perceptual integration of facial expression and emotional body language. *Proc Natl Acad Sci USA* 102: 16518–16523, 2005.

Mitchell JP, Heatherton TF, Macrae CN. Distinct neural systems subserve person and object knowledge. *Proc Natl Acad Sci USA* 99: 15238–15243, 2002.

Morris JP, Pelphrey KA, McCarthy G. Occipitotemporal activation evoked by the perception of human bodies is modulated by the presence or absence of the face. *Neuropsychologia* 44: 1919–1927, 2006.

Mur M, Ruff DA, Bodurka J, Bandettini PA, Kriegeskorte N. Face-identity change activation outside the face system: "release from adaptation" may not always indicate neuronal selectivity. *Cereb Cortex* 20: 2027–2042, 2010.

Peelen MV, Atkinson AP, Andersson F, Vuilleumier P. Emotional modulation of body-selective visual areas. *Soc Cogn Affect Neurosci* 2: 274–283, 2007.

Peelen MV, Atkinson AP, Vuilleumier P. Supramodal representations of perceived emotions in the human brain. *J Neurosci* 30: 10127–10134, 2010.

Peelen MV, Downing PE. Selectivity for the human body in the fusiform gyrus. *J Neurophysiol* 93: 603–608, 2005.

Peelen MV, Glaser B, Vuilleumier P, Eliez S. Differential development of selectivity for faces and bodies in the fusiform gyrus. *Dev Sci* 12: F16–F25, 2009.

Peelen MV, Wiggett AJ, Downing PE. Patterns of fMRI activity dissociate overlapping functional brain areas that respond to biological motion. *Neuron* 49: 815–822, 2006.

Perrett DI, Rolls ET, Caan W. Visual neurones responsive to faces in the monkey temporal cortex. *Exp Brain Res* 47: 329–342, 1982.

Pinsk MA, Arcaro M, Weiner K, Kalkus J, Inati S, Gross CG, Kastner S. Neural representations of faces and body parts in macaque and human cortex: a comparative fMRI study. *J Neurophysiol* 101: 2581–2600, 2009.

Pinsk MA, DeSimone K, Moore T, Gross CG, Kastner S. Representations of faces and body parts in macaque temporal cortex: an fMRI study. *Proc Natl Acad Sci USA* 102: 6996–7001, 2005.

Reddy L, Kanwisher NG, VanRullen R. Attention and biased competition in multi-voxel object representations. *Proc Natl Acad Sci USA* 106: 21447–21452, 2009.

Saxe R, Kanwisher N. People thinking about thinking people: the role of the temporo-parietal junction in 'theory of mind'. *Neuroimage* 19: 1835–1842, 2003.

Schmalzl L, Zopf R, Williams M. From head to toe: evidence for response selectivity to whole individuals in human extrastriate cortex. *Front Hum Neurosci* 6: 108, 2012.

Schwarzlose R, Baker C, Kanwisher N. Separate face and body selectivity on the fusiform gyrus. *J Neurosci* 25: 11055–11059, 2005.

Schwarzlose RF, Swisher JD, Dang S, Kanwisher N. The distribution of category and location information across object-selective regions in human visual cortex. *Proc Natl Acad Sci USA* 105: 4447–4452, 2008.

Seidl KN, Peelen MV, Kastner S. Neural evidence for distracter suppression during visual search in real-world scenes. *J Neurosci* 32: 11812–11819, 2012.

Simmons WK, Martin A. Spontaneous resting-state BOLD fluctuations reveal persistent domain-specific neural networks. *Soc Cogn Affect Neurosci* 7: 467–475, 2012.

Spillmann L. From perceptive fields to Gestalt. *Prog Brain Res* 155: 67–92, 2006.

Sripati AP, Olson CR. Responses to compound objects in monkey inferotemporal cortex: the whole is equal to the sum of the discrete parts. *J Neurosci* 30: 7948–7960, 2010.

Summerfield C, Trittschuh EH, Monti JM, Mesulam MM, Egner T. Neural repetition suppression reflects fulfilled perceptual expectations. *Nat Neurosci* 11: 1004–1006, 2008.

Tanaka K, Saito H, Fukada Y, Moriya M. Coding visual images of objects in the inferotemporal cortex of the macaque monkey. *J Neurophysiol* 66: 170–189, 1991.

Tarr MJ, Gauthier I. FFA: a flexible fusiform area for subordinate-level visual processing automatized by expertise. *Nat Neurosci* 3: 764–769, 2000.

Tsao DY, Freiwald WA, Knutsen TA, Mandeville JB, Tootell RB. The representation of faces and objects in macaque cerebral cortex. *Nat Neurosci* 6: 989–95, 2003.

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

Van den Stock J, Righart R, de Gelder B. Body expressions influence recognition of emotions in the face and voice. *Emotion* 7: 487–494, 2007.

Van Essen DC. A tension-based theory of morphogenesis and compact wiring in the central nervous system. *Nature* 385: 313–318, 1997.

Wachsmuth E, Oram MW, Perrett DI. Recognition of objects and their component parts: responses of single units in the temporal cortex of the macaque. *Cereb Cortex* 5: 509–522, 1994.

Wagemans J, Elder JH, Kubovy M, Palmer SE, Peterson MA, Singh M, von der Heydt R. A century of Gestalt psychology in visual perception: I. Perceptual grouping and figure-ground organization. *Psychol Bull* 138: 1172–217, 2012.

Wallbott HG. Bodily expression of emotion. *Eur J Soc Psychol* 28: 879–896, 1998.

Wertheimer M. Laws of organization in perceptual forms. In: *A Source Book of Gestalt Psychology*, edited by Ellis WD. London: Harcourt Brace, 1938.

Zoccolan D, Cox DD, DiCarlo JJ. Multiple object response normalization in monkey inferotemporal cortex. *J Neurosci* 25: 8150–8164, 2005.

Zoccolan D, Kouh M, Poggio T, DiCarlo JJ. Trade-off between object selectivity and tolerance in monkey inferotemporal cortex. *J Neurosci* 27: 12292–12307, 2007.