

An Integrated Face–Body Representation in the Fusiform Gyrus but Not the Lateral Occipital Cortex

Michal Bernstein, Jonathan Oron, Boaz Sadeh, and Galit Yovel

Abstract

■ Faces and bodies are processed by distinct category-selective brain areas. Neuroimaging studies have so far presented isolated faces and headless bodies, and therefore little is known on whether and where faces and headless bodies are grouped together to one object, as they appear in the real world. The current study examined whether a face presented above a body are represented as two separate images or as an integrated face–body representation in face and body-selective brain areas by employing a fMRI competition paradigm. This paradigm has been shown to reveal higher fMRI response to sequential than simultaneous presentation of multiple stimuli (i.e., the competition effect), indicating competitive interactions among simultaneously presented multiple stimuli. We therefore hypothesized that if a

face above a body is integrated to an image of a person whereas a body above a face is represented as two separate objects, the competition effect will be larger for the latter than the former. Consistent with our hypothesis, our findings reveal a competition effect when a body is presented above a face, but not when a face is presented above a body, suggesting that a body above a face is represented as two separate objects whereas a face above a body is represented as an integrated image of a person. Interestingly, this integration of a face and a body to an image of a person was found in the fusiform, but not the lateral-occipital face and body areas. We conclude that faces and bodies are processed separately at early stages and are integrated to a unified image of a person at mid-level stages of object processing. ■

INTRODUCTION

Neuroimaging studies have revealed category-selective responses to images of faces and headless bodies in the lateral-occipital cortex and the fusiform gyrus. Numerous studies have examined the representation of faces and bodies in their category-selective cortex, the extrastriate and fusiform body areas (EBA and FBA, respectively) and the occipital and fusiform face areas (OFA and FFA, respectively; for reviews, see Peelen & Downing, 2007; Kanwisher & Yovel, 2006). These studies presented isolated faces or isolated headless bodies, and therefore little is known on where in the brain these two objects that typically appear together in real world are grouped to one object of a person.

Recent studies suggest a close interaction between the representation of faces and bodies. In an elegant study, Ghuman, McDaniel, and Martin (2010) have shown that behavioral adaptation to the identity of a headless body influences the perceived identity of a face (i.e., perceptual aftereffect). In other words, long exposure to headless body A biased the perception of a morphed face between face A and face B toward the identity of face B. The same effect was also found for body and face gender. Interestingly, this behavioral adaptation was specific to face and body images and was not found for other gender-related objects such as male and female shoes. An interaction between faces and bodies has also been

reported in studies that show activation of face processing mechanisms to faceless heads presented in body context. Cox, Meyers, and Sinha (2004) revealed a high fMRI response in the FFA to blurred faces attached to bodies, which was as large as the response to faces and significantly larger than the response to blurred faces presented alone. In line with these findings, Brandman and Yovel (2012) revealed similar inversion effect and face detection rating for faceless heads in body context as to faces, but these effects were reduced for faceless heads presented alone. Thus, body context can activate face mechanisms for head stimuli with no facial features. Finally, de Gelder and colleagues (Van den Stock, Righart, & de Gelder, 2007; Meeren, van Heijnsbergen, & de Gelder, 2005) showed integration of facial expression and emotional body language. They showed that, when face and body convey conflicting emotional information, the judgment of facial expression is biased toward the emotion expressed by the body. Aviezer and colleagues (Aviezer, Trope, & Todorov, 2012; Aviezer et al., 2008) similarly demonstrated that the emotion expressed by bodies influenced the perceived expression of a face, suggesting that faces and bodies are not processed independently but interactively. Taken together, these findings suggest that despite the fact that faces and headless bodies are processed by separate neural mechanisms, they strongly interact with one another.

Given the interaction found between faces and bodies on one hand and their distinct neural mechanisms on the other hand, it is important to determine whether face

and body-selective areas are involved in the grouping of a face and a headless body to an image of a person. This question was recently addressed in three fMRI studies that compared the representation of a face and a body to the representation of a whole person in face and body-selective areas (Kaiser, Strnad, Seidl, Kastner, & Peelen, 2014; Song, Luo, Li, Xu, & Liu, 2013; Schmalzl, Zopf, & Williams, 2012). Schmalzl et al. (2012) addressed the question of whether there are neural populations that exclusively code for whole individuals, by comparing fMRI adaptation effects when either both the face and the body, just the face, just the body, or neither the face nor the body of an individual were repeated. They found that, in each of the category-selective areas, a small percentage of voxels significantly adapted only to the repetition of the same whole individual and not to the repetition of the same face or same body alone. This adaptation effect was superadditive, meaning that response selectivity to whole individuals could not be explained merely by the sum of face and body selective responses. These results suggest preferential coding for whole individuals, in accordance with findings in single unit recording in monkeys that showed neural populations that are selective to whole individuals and unresponsive to isolated faces or bodies (Wachsmuth, Oram, & Perrett, 1994).

In a second recent study, Song et al. (2013) investigated the representation of faces and bodies using four experimental conditions: isolated face, isolated body, a face above a body (termed “FB condition”), and a body above a face (termed “scrambled condition”). They defined a “context effect” as the response difference between the FB condition and the face condition in the FFA and OFA and the response difference between the FB condition and the body condition in the FBA and EBA. Context effect was observed in the FFA, showing higher response to FB than to the face condition, and in the FBA, showing higher response to FB than to the body condition, whereas in the OFA and EBA the response to FB did not differ significantly from that of the preferred (face or body) stimulus. The authors conclude that context effect is larger in anterior than posterior regions and relate this finding to previous work suggesting that object recognition is achieved through hierarchical processing stages, so that low-level and local stimulus properties are processed in the posterior portion of the ventral pathway while more abstract and global properties are processed in the anterior portion (Taylor & Downing, 2011). Song et al. also used multivariate pattern analysis to examine which spatial pattern of neural activation, the FB condition or the scrambled condition, was more similar to the “synthetic” pattern created by averaging the spatial pattern for the face condition and that for the body condition (see, e.g., MacEvoy & Epstein, 2009, 2011). Results showed that, in the right FFA, the synthetic pattern was more similar to the pattern of the scrambled condition than to that of FB condition, suggesting a linear combination when a body is presented above a face but a holistic (nonlinear) representation when a face is

presented above a body. It is noteworthy that whereas in the scrambled condition the body and face stimuli were separated, the FB condition presented a whole person in which the face and body are attached. Therefore, these conditions are not comparable in terms of multiple object representation, such that the scrambled condition presents two stimuli whereas in the FB condition there are no multiple objects but a single stimulus of a person.

A third study, by Kaiser and colleagues (2014), employed a paradigm very similar to that of Song et al. (2013): A synthetic person pattern was computed as the average of an isolated face pattern and an isolated body pattern. Kaiser et al. compared the multivoxel pattern correlations between the synthetic patterns and the actual patterns associated with faces, bodies, and whole persons. They found, in contrast to Song et al. (2013), that the synthetic pattern was statistically indistinguishable from the person pattern, meaning that the person pattern 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 the fusiform gyrus and not found in lateral-occipital cortex. In the lateral-occipital area, the synthetic pattern was distinguishable from the person pattern, reflected by higher person–person correlation than person–synthetic correlation. Given that response patterns to whole persons could be accurately modeled by a linear combination of response patterns evoked by isolated faces and bodies, the authors concluded that the fusiform gyrus encodes persons in a part-based rather than an integrated (holistic) manner. Importantly, Kaiser et al. (2014) did not include a body-above-face condition as in Song et al. (2013) study, which is expected to generate a linear combination and therefore is an important control condition to which the face above a body condition should be compared.

Here we used a different approach to ask whether a face presented above a body is grouped to an integrated image or represented as two separate images. In particular, we adapted the paradigm introduced by Kastner and colleagues (Kastner et al., 2001; Kastner, De Weerd, Desimone, & Ungerleider, 1998) in which the fMRI response is measured for simultaneous versus sequential presentation of multiple stimuli. This paradigm was inspired by monkey electrophysiology studies that have shown neural competition between two unattended simultaneously presented stimuli (Miller, Gochin, & Gross, 1993). In multiple fMRI studies, Kastner and colleagues have shown that simultaneous presentation of multiple stimuli generates an fMRI response that is lower than the sequential presentation of the same stimuli. They have suggested that the lower neural response to simultaneous than sequential presentation reflects suppressive interactions among the simultaneously presented multiple stimuli similar to those that were reported in the single unit recording studies of multiple stimuli (for a review, see Desimone & Duncan, 1995). Indeed, consistent with monkey electrophysiology studies (e.g., Reynolds, Chelazzi, & Desimone, 1999; Luck,

Chelazzi, Hillyard, & Desimone, 1997) attending to one of the simultaneously presented stimuli reduced competition effects (Kastner et al., 1998). Competition effects were also reduced when increasing the distance among the stimuli and more so in early visual areas with small receptive fields (Kastner et al., 2001) consistent with reduced competition when stimuli are not presented within the same receptive field. These findings suggest the difference in fMRI response to simultaneous and sequential presentations do not merely reflect differences in stimulus presentation parameters (e.g., number of onsets offsets, total presentation time) but are consistent with top-down (e.g., attention) and bottom-up (e.g., distance) effects on competition among multiple stimuli.

Most relevant to the current study, McMains and Kastner (2010) investigated the effect of grouping on neural competition among multiple stimuli. They presented two displays: a grouping display, in which multiple stimuli were arranged such that an illusory figure (e.g., a square) was formed, and a nongrouping display, in which the same stimuli were rotated outward, so that no figure was formed (McMains & Kastner, 2010). Both displays were presented either sequentially or simultaneously, and the difference in response between the two presentation conditions was measured. McMains and Kastner found that the competition effect was reduced for the grouping display relative to the non-grouping display. They concluded that perceptual grouping of simultaneously presented stimuli reduced suppressive interaction among them.

The current study applies a similar paradigm to ask whether an image of a face and a body are represented as two separated objects or as an integrated face-body representation. Because this paradigm has so far been used with low-level visual stimuli in low-level visual cortex,

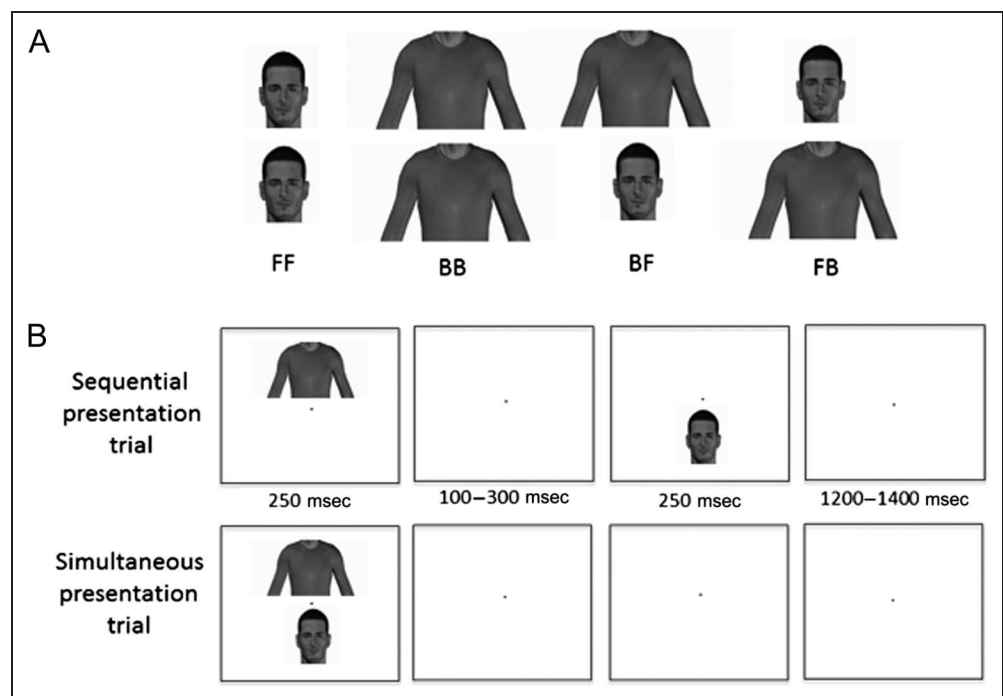
we first assessed if the competition effect is found for pairs of two faces or two bodies in their category-selective areas. We then asked whether a combination of a face above a body, which can be grouped to an image of a person, would generate less competition than the combination of a body above a face, which is perceived as two independent objects (see Figure 1). Furthermore, our design allows us to directly assess the effect of perceptual grouping (i.e., the unification of two separate objects into one), because we present separated faces and bodies both in the face-above-body and in the body-above-face conditions. In addition, in Song et al. the stimuli were attended so effects found may be attributed to greater attention to the familiar image of a person relative to the scrambled version. In the current study, participants performed a task on a central fixation dot so the face and body stimuli were unattended. Finally, based on findings that the representation of objects in the occipital object-selective areas is more primitive and local whereas the ventral-temporal object areas generate a more global representation (Song et al., 2013; Taylor & Downing, 2011), we asked whether grouping would appear at later stages in the hierarchical processing of faces and bodies. To test that, we compared the magnitude of neural suppression in the face and body-selective areas in the lateral occipital cortex and fusiform gyrus (Figure 2).

METHODS

Participants

Twenty-two healthy volunteers (eight men, ages 21–34) with normal or corrected-to-normal vision participated

Figure 1. (A) The four types of stimulus pairs presented in the experiment. Each of these pairs of stimuli was presented either simultaneously (both stimuli at the same time) or sequentially (one after the other), creating a total of eight conditions. (B) Each trial lasted for 2 sec and included two images presented either sequentially or simultaneously.



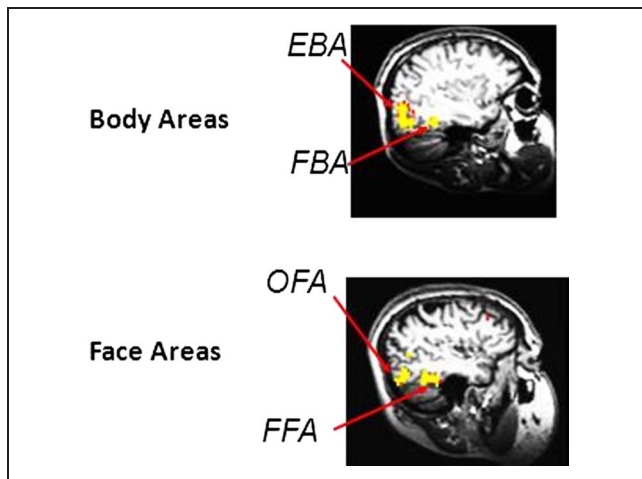


Figure 2. The face and body selective areas were localized in each participant based on a localizer scan. The four areas are shown for a representative participant. Body areas include the EBA in the lateral occipital cortex and the FBA in the fusiform gyrus. The face areas include the OFA in the lateral occipital cortex and the FFA in the fusiform gyrus.

in the study for either course credit or \$15/hr. Two participants were excluded because of technical difficulties (scanner overheating), leaving a total of 20 participants in the final analysis. All participants provided a written informed consent to participate in the study, which was approved by the ethics committees of the Tel Aviv Sourasky Medical Center and Tel Aviv University.

Stimuli

The experiment included two parts: a functional localizer and the main experiment. Stimuli of the functional localizer were 80 grayscale photographs of faces, headless bodies, objects, and scrambled objects (i.e., of each of the object images). Stimuli of the main experiment were four grayscale photographs: two different faces and two different bodies (torsos). The main experiment included four types of stimuli: two faces (FF), two bodies (BB), face above body (FB), and body above face (BF; Figure 1), each presented either simultaneously or sequentially. Pairs of two faces or two bodies were of the same individual. A pair of a face and a body was composed of stimuli of the same individual (two different pairs in total: the face of individual A with the body of individual A and the face of individual B with the body of individual B).

Apparatus and Procedure

fMRI data were collected in a 3T GE MRI scanner, using an eight-channel head coil. Echo-planar volumes were acquired with the following parameters: repetition time (TR) = 2 sec, echo time = 35 msec, 23 slices per TR, slice thickness = 2.4 mm, 96×96 matrix, field of view = 20 cm. Stimuli were presented with Matlab (Psychtoolbox; Brainard, 1997) and projected onto a screen viewed by

the participants through a mirror located in the scanner. Anatomical SPGR images were collected with $1 \times 1 \times 1$ mm resolution, echo time = 3.52 msec, TR = 9.104 msec.

Functional localizer included three scans. Each scan included the four object categories, faces, headless bodies, objects, and scrambled objects, in different blocks. Each block lasted 16 sec and included 20 stimuli, each presented for 350 msec with a 450 msec ISI. Category block order was counterbalanced within and across scans. Each localizer scan consisted of four blocks for each category and five blocks of a baseline fixation point resulting in a total of 21 blocks (336 sec). To maintain vigilance, participants were instructed to press a response box button, whenever two identical images appeared consecutively (a 1-back task). This happened twice per block in a random location.

The main experiment included four scans. Each scan included eight different conditions presented in different blocks. In each condition, we presented pairs of images above and below fixation: two faces, two bodies, a face above a body, and a body above a face (see Figure 1A). The four types of pairs were presented either sequentially or simultaneously (Figure 1). Each image was presented for 250 msec. In the simultaneous presentation, two images were presented together for 250 msec; in the sequential presentation the first image was presented for 250 msec, and after an ISI of 100–300 msec the second image was presented for the same duration of time. The duration of each trial was 2 sec, and each block lasted 16 sec for all the conditions.

Each scan consisted of two blocks of each of the eight conditions and five blocks of baseline fixation point resulting in a total of 21 blocks (336 sec). The order of the eight conditions was counterbalanced within and across scans. Participants were asked to press a response box button, whenever a fixation point changed its color at random intervals (color changes were simultaneous with stimulus onset, on average twice per block). This was done to maintain vigilance and fixation without drawing attention to the stimuli themselves. The scans of the localizer and the main experiment were interleaved.

Data Analysis

Behavioral Data Analysis

We calculated the proportion of misses and false alarms in the fixation point color change task during each of the eight conditions.

fMRI Data Analysis

fMRI analysis was accomplished using statistical parametric mapping (SPM5). The first six volumes in the localizer scans and the first four volumes in the main experiment scans were acquired during a blank screen display and were discarded from the analysis (“dummy scans”). The

data were then preprocessed using slice timing correction and realignment. All slices that were acquired during the localizer and experimental tasks were realigned to the first slice. Spatial smoothing was done for the localizer data only (5 mm).

ROI Analysis

In the localizer experiment, category-selective voxels were defined using combinations of *t* contrasts to assure their specific selectivity. Face-selective areas in the fusiform gyrus (FFA) and the lateral occipital cortex (OFA) were extracted from a full conjunction of faces > objects and faces > scrambled-objects *t* contrast maps ($p < .00001$), masking out bodies > objects *t* contrast map ($p < .05$). Similarly, body-selective areas in the fusiform gyrus (FBA) and in the lateral occipital cortex (EBA) were extracted from a full conjunction of bodies > objects and bodies > scrambled-objects *t* contrast maps ($p < .00001$), excluding faces > objects *t* contrast map ($p > .05$). The lateral occipital cortex object general areas were defined as voxels in the lateral occipital (LO) and posterior fusiform gyrus (pFs) that showed higher response to objects than to the scrambled images of these objects ($p < .00001$), excluding faces > objects and bodies > objects *t* contrast maps ($p > .05$). ROIs were defined individually for each participant. Table 1 indicates the volume and the number of participants that showed each of these ROIs.

Main Experiment

Time courses were extracted for each of the eight conditions using the MarsBaR ROI toolbox of SPM (Brett, Anton, Valabregue, & Poline, 2002) within each of the predefined ROIs. The peak of the BOLD signal was within the third/fourth TRs and the eighth TR, which is when the 16-sec block ended. Accordingly, the dependent measure was an average of the values of TRs 3–8. Statistical analysis of the average response to each of the eight conditions within each ROI was performed with Statistica 9.

We first assessed for each face and body areas (EBA, FBA, OFA, and FFA) whether there was a significant interaction of the effects of interest with hemisphere.

Table 1. The Volume (Number of Voxels) of the Face and Body Areas (after Excluding Voxels Selective to the Other Category) and the Number of Participants in which Each of These Areas Was Found

	Face Areas		Body Areas	
	OFA	FFA	EBA	FBA
Right	87 ($n = 17$)	51 ($n = 20$)	801 ($n = 20$)	53 ($n = 17$)
Left	74 ($n = 17$)	48 ($n = 17$)	660 ($n = 20$)	40 ($n = 15$)

The size of each voxel was (2 mm × 2 mm × 2.4 mm).

Table 2. Behavioral Performance (Proportion of Misses) on the Orthogonal Color Change Detection Task during the Presentation of Each of the Experimental Conditions

	FF	BB	BF	FB
Sequential	8% (3.2)	8% (3.6)	7% (4.4)	8% (3.1)
Simultaneous	7% (4.3)	6% (3.1)	7% (2.8)	7% (3.6)

Because we found no interaction of any of the factors of interest with Hemisphere ($p > .1$), data from the two hemispheres were averaged using weighted averages based on the volume of each ROI. In the object general areas (LO and pFs), we did find interactions with hemisphere; therefore, we report the results separately for the right and left ROIs.

RESULTS

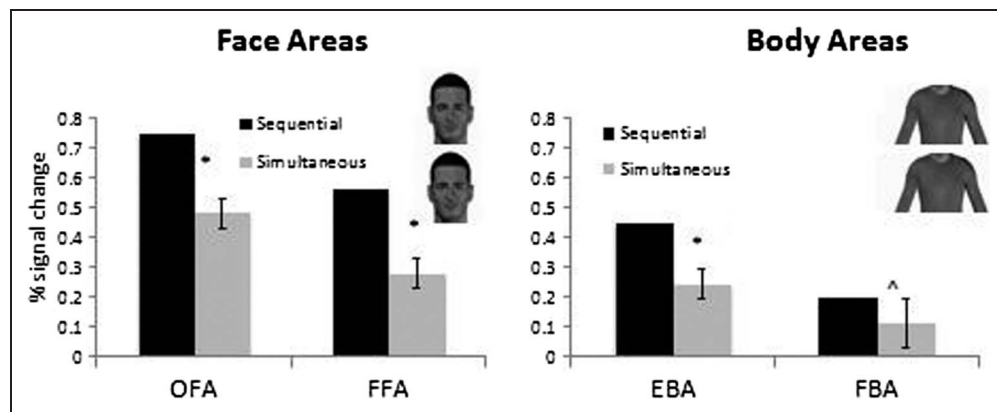
Behavioral Analysis

We calculated the proportion of misses and false alarms of a dot color change during the presentation of the eight experimental conditions. The mean proportion of misses was 7.3% and was similar across the eight experimental conditions (Table 2). A two-way ANOVA with Stimulus Type (FF, BB, BF, FB) and Presentation Type (sequential, simultaneous) did not reveal any significant differences. The proportion of false alarms was very low (<1%) and also did not differ across the different conditions. These findings suggest that the type of stimuli presented did not influence performance on the dot task.

fMRI Data Analysis

We first asked whether neural competition is found for two stimuli from the same category in their category-selective areas. Our data revealed lower responses for two simultaneously presented than two sequentially presented preferred stimuli in each of the category-selective areas (Figure 3). A three-way repeated-measure ANOVA of Brain Area (occipital, temporal), Category Area (face area, body area), and Presentation Type (sequential, simultaneous) was used to assess whether simultaneous presentation of two faces in the face areas or two bodies in the body areas generate lower response than sequential presentation of the same stimuli. This analysis revealed a main effect of Presentation Type, $F(1, 17) = 30.81$, $p < .00001$, which reflects higher response to sequential than simultaneous presentation (Figure 3). We found no interaction of Presentation Type with Brain Area or Category Area, suggesting that the magnitude of suppression did not differ in the different ROIs. Paired *t* tests confirmed these findings showing effect of suppression for pairs of faces in the face areas (OFA $t(18) = 5.2$, $p < .0001$ and FFA $t(19) = 4.8$, $p < .0001$) and for

Figure 3. The response to sequential and simultaneous presentation of two faces and two bodies in the face-selective (left side) and body selective (right side) areas indicates neural suppression for pairs of simultaneously presented preferred stimuli (faces in face areas and bodies in body areas). * $<.05$, ^ $=.07$. Error bars indicate the standard error of the difference between the simultaneous and sequential conditions.

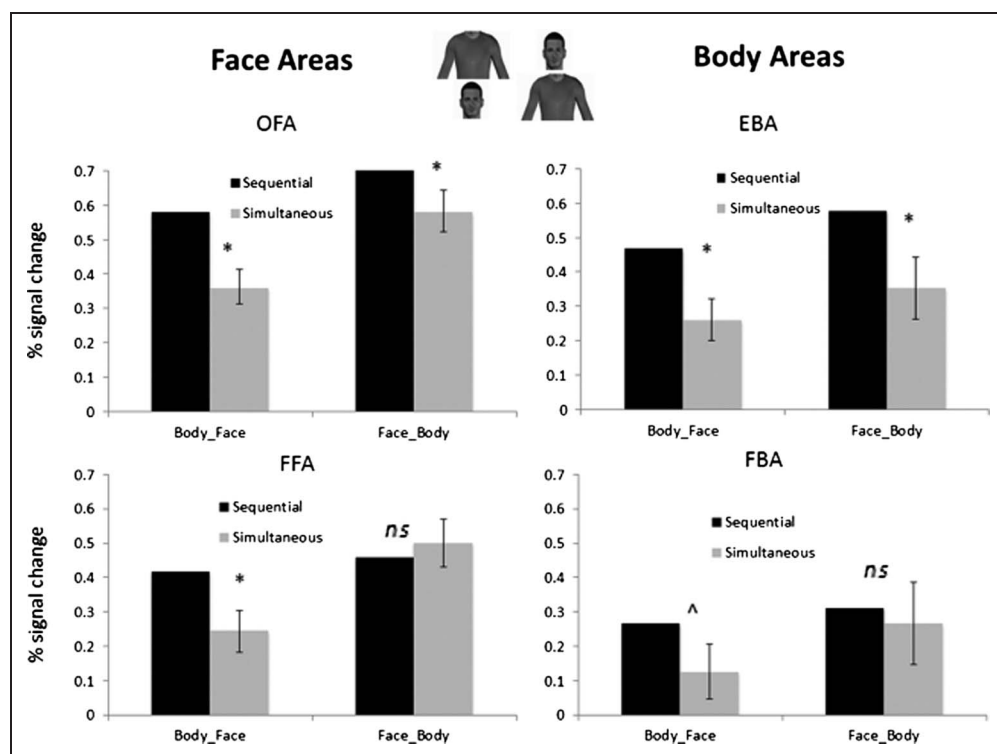


pairs of bodies in the body areas (EBA $t(19) = 3.2, p < .005$ and marginally significant in FBA $t(18) = 1.9, p = .07$).

We next asked how grouping may influence the magnitude of neural suppression among pairs of objects. In other words, would a combination of a face above body (Figure 1), which can be perceptually grouped to one image of a person, generate less suppression than the body above face, which is represented as two independent stimuli. Our data reveal significant neural suppression when a body is presented above face. In contrast, when a face is presented above a body, neural suppression was found in the lateral occipital face and body areas but not in the face and body areas in the fusiform gyrus (see Figure 4). A four-way repeated-measure ANOVA of Brain Area (occipital, temporal), Category Area (face area, body area), Stimulus Type (FB, BF), and Presentation

Type (sequential, simultaneous) revealed a significant three-way interaction of Brain Area, Stimulus Type, and Presentation Type, $F(1, 17) = 8.4, p < .01$, partial eta squared = 0.33. This reflects neural suppression for simultaneous presentation of a body above a face but not for a face above a body in the face and body areas in the fusiform gyrus (a significant interaction of Stimulus Type and Presentation Type, $F(1, 18) = 7.3, p < .02$, partial eta squared = 0.29, in the face and body areas in the fusiform gyrus) but not in the face and body areas in the lateral occipital cortex. Paired t tests indicate neural suppression for a body above a face in the FFA, $t(19) = 3.1, p < .01$, and a marginally significant effect in the FBA, $t(19) = 1.8, p = .08$, no suppression was found for a face above a body in the FFA, $t(19) = .64, p = .52$, and FBA, $t(18) = .39, p = .7$. In contrast, in the

Figure 4. The response to sequential and simultaneously presented body above a face and a face above body in the face and body areas shows neural suppression for both types of face-body configurations in the face and body areas in the lateral occipital cortex. In the fusiform gyrus face and body areas, neural suppression was found for a body above a face but not for a display of a face above a body. * $<.05$, ^ $=.08$. Error bars indicate the standard error of the difference between the simultaneous and sequential conditions.



face and body areas in the lateral occipital cortex, there was a main effect of Stimulus Presentation, $F(1, 18) = 19.8, p < .0001$, partial eta squared = 0.50, and no interaction with Condition ($F < 1$), indicating neural suppression to both types of stimulus presentations. Paired t tests reveal significant suppression for body above face in the EBA, $t(19) = 3.1, p < .005$, and the OFA, $t(18) = 4.0, p < .001$, and for a face above a body in the EBA, $t(19) = 2.5, p < .02$, and the OFA, $t(18) = 2.8, p < .02$. Finally, paired t test reveal no difference between the sequential presentation of a body above a face than a face above a body in the FFA and FBA ($p > .1$) but a higher response to a body above a face than a face above a body in the OFA and EBA ($p < .03$). The response to simultaneous presentation of a face above a body was higher than a body above a face in all face and body areas ($p < .03$).

To examine whether our findings reflect a mechanism specific to face and body regions or a general distinction between ventral and lateral cortices, we next examined the responses of object general areas. A four-way repeated-measures ANOVA of Brain Area (occipital, temporal), Hemisphere (right, left), Stimulus Type (FF, BB), and Presentation Type (sequential, simultaneous) revealed a main effect of Stimulus Type, $F(1, 12) = 17.5, p < .001$, reflecting higher response to bodies (BB) than to faces (FF) in the object general areas (right LO, $F(1, 18) = 24.80, p < .00009$; left LO, $F(1, 19) = 13.11, p < .001$; right pFs, $F(1, 18) = 12.08, p < .002$). The four-way ANOVA also showed an interaction of Area and Presentation Type, $F(1, 12) = 6.45, p < .02$, and an interaction of Hemisphere and Presentation Type, $F(1, 12) = 6.5, p < .02$. Further examination of these interactions revealed that only right LO showed a main effect of Presentation Type, $F(1, 18) = 6.46, p = .02$, which reflects higher response to sequential than simultaneous presentation. A four-way repeated-measures ANOVA of Brain Area (occipital, temporal), Hemisphere (right, left), Stimulus Type (FB, BF), and Presentation Type (sequential, simultaneous) revealed no significant main effects or interactions, suggesting that, unlike the face and body-selective areas, no consistent patterns of competition were found in the different object general areas.

DISCUSSION

Face and body stimuli hardly ever appear in isolation. Nevertheless, because they are processed by separate brain areas in the visual cortex, most neuroimaging studies have examined the neural representation of isolated faces and headless bodies in their category-selective areas. The current study aimed to take one step forward and assess the representation of simultaneously presented face and headless body as they typically appear in real life. In particular, we asked whether a face presented above a body, as in the image of a whole person, are represented as two separated objects or are grouped together. We first showed

that the neural response to pairs of two faces or two bodies presented simultaneously is lower than when the same stimuli were presented sequentially in face-selective areas (OFA and FFA) and body-selective areas (EBA and FBA), respectively (Figure 3). These results are consistent with previous neuroimaging studies with low-level visual stimuli (e.g., Beck & Kastner, 2005, 2007; Kastner et al., 1998, 2001), which found an fMRI competition effect for multiple stimuli presented simultaneously.

This fMRI competition effect allowed us to determine whether two objects presented simultaneously are represented as two separate objects or are grouped together. In particular, we asked whether a face presented above a body contains two objects, which can be grouped to an image of a person, in contrast to a body above a face, which are represented as two independent objects. Our results show that, in ventral temporal face- and body-selective regions, the competition effect was evident when a body was presented above a face but not when a face was presented above a body. In other words, whereas the response to sequential presentation of a body above a face was higher than the response to its simultaneous presentation, simultaneous presentation of a face above a body was not lower than its sequential presentation. The lack of competition for simultaneous presentation of the face above body, in contrast to the competition found for the other stimulus presentations, further suggests that differences in fMRI signal to sequential and simultaneous presentations do not merely reflect differences in stimulus presentation parameters (e.g., different number of onsets and offsets), which would generate similar effects across all conditions.

Interestingly, integration of a face and a body was found in the ventral temporal face and body areas (FFA and FBA) but not in the lateral occipital face and body areas (OFA and EBA; Figure 4). The dissociation between the response of the occipital and temporal brain areas suggests that perceptual grouping of a face and a body takes place at intermediate stages of visual processing whereas at earlier stages a face and a body are still represented as two separate objects. These findings are consistent with studies showing that the representation of objects in the lateral occipital object areas is more local whereas the ventral-medial object areas generate a more global representation (Taylor & Downing, 2011). For example, studies that examined the representation of whole and part-based stimuli (Taylor, Wiggett, & Downing, 2007; Lerner, Hendler, & Malach, 2002; Lerner, Hendler, Ben-Bashat, Harel, & Malach, 2001) show a preference for the representation of wholes relative to their parts in object category selective cortex. These studies also show that a preference to wholes is usually larger in object areas that are higher in the hierarchy than in lower-level object areas. Our results are also in line with the findings of Song et al. (2013) showing an integrated representation of a face and a body in the fusiform but not in the occipital regions. Song et al. have shown this for the right FFA only; here we expand their findings also

to the left FFA and right and left FBA. Importantly, as opposed to the face and body stimuli used by Song et al. that were combined to one image in the FB condition but not in the BF condition, here we presented separate face and body in both conditions, therefore allowing us to make conclusions regarding the representation of multiple contextually related objects. Finally, our findings also accord with the incremental grouping theory proposed by Roelfsema and colleagues (Korjoukov et al., 2012; Roelfsema & Houtkamp, 2011; Roelfsema, 2006), which suggests that when grouping involves integration of information about object parts that are processed by different brain areas it may take place after object categorization (Korjoukov et al., 2012). Thus, object categorization may take place in the face and body areas in the lateral occipital cortex, whereas the integration of the two objects may take place in the fusiform gyrus where face and body areas reside close together by lateral connections between them.

Interestingly, it has been suggested that the regularity in which faces and bodies are presented relative to one another in everyday life has been incorporated into the visual system resulting in a map of face- and body-selective regions throughout high-level visual cortex, similar to the well-known retinotopic map, where two adjacent points in visual space are projected to two adjacent points on the retina (Weiner & Grill-Spector, 2013). Importantly, the face and body regions in the fusiform gyrus are more adjacent to each other (and even overlap) than the face and body regions in the lateral-occipital area, which are anatomically separated from each other. The fusiform regions are therefore more likely to host integrated representations of faces and bodies than the occipital regions, consistent with our findings. Also consistent with this idea of the influence of perceptual experience on neural activations to faces and bodies is the work by Chan, Kravitz, Truong, Arizpe, and Baker (2010), demonstrating that the strength and distinctiveness of visual representations in body and face-selective cortex is determined by long-term natural visual experience. Specifically, they showed that response patterns in the EBA can be discriminated for right body parts in the left visual field and for left body parts in the right visual field. Similarly, in the FFA discrimination was strongest for right half-faces in the left and left half-faces in the right visual field. These combinations of visual field and stimulus side correspond to the commonly experienced configurations. Chan et al. (2010) suggested that representations in face and body regions directly capture the statistics with which complex stimuli occur.

It is noteworthy that the lack of competition for simultaneously presented face above body may reflect a mechanism of mutual facilitation, resulting from the fact that they always appear together in the real world. Such mutual facilitation mechanism may counteract the mutual inhibition between these two objects and reduce competition between them. Our data do not allow ruling out this alter-

native explanation in favor of a grouping account. However, taken together with Song et al. (2013), who found that the pattern of response to a face above a body (but not to a body above a face) is more than the linear sum of the pattern of response to each stimulus alone, we suggest that our findings are likely to reflect grouping mechanisms that generate an image which is more than the sum of its parts.

One possible alternative explanation to the higher response to simultaneous presentation of a face above a body that should be addressed is the possibility that this familiar combination of objects attracted more attention than the other conditions. Two sets of findings allow us to rule out this alternative explanation. First, increased attention to the face and body stimuli would result in larger proportion of misses in the orthogonal dot color change detection task that participants performed during the presentation of the stimuli. However, proportions of misses were similar across all experimental conditions (Table 1). Second, greater allocation of attention to a face and a body presented simultaneously would increase the fMRI response to this condition relative to all other conditions in both the lateral-occipital and the temporal-ventral areas. However, our findings show higher response to sequential than simultaneous presentation in the occipital face and body areas and no difference in the temporal-ventral areas (Figure 4).

One drawback of the paradigm used here is that it does not include the presentation of a face and a body alone and therefore does not allow us to predict the response to the compound stimuli (a face and a body) from the response to the single stimuli (see MacEvoy & Epstein, 2009). The study by Song et al. (2013), mentioned above, did present a face and body stimuli alone and showed that the response to a face above a body was higher than the response to each of them presented alone. Interestingly, this effect was found in the fusiform face and body areas but not in the OFA and the EBA, consistent with the dissociation between the lateral occipital and fusiform areas that we reported here. Most interesting was the finding that the linear combination of multivoxel patterns to the isolated face and body could not be discriminated from the pattern generated to a body above a face but was discriminated from the response to a face above a body. This effect was only found in the right FFA and suggest that the representation of a face above a body is different from the linear combination of its components as expected from a holistic representation. Taking a different approach, we reported here similar results of grouped (holistic) representation of simultaneously presented faces and bodies. The grouping effect was found in bilateral FFA and FBA, expanding Song et al.'s results also to the body areas.

Our findings are also consistent with the recent study by Schmalzl et al. (2012), mentioned above. This study found that in each of the category-selective ROIs, a small

percentage of voxels showed adaptation effect only to the repetition of the same whole individual and not to the repetition of the same face or same body alone. This adaptation effect was superadditive, meaning that response selectivity to whole individuals could not be explained merely by the sum of face and body selective responses. These results suggest preferential coding for whole individuals and are consistent with single-cell recordings in monkeys showing neural populations that are sensitive to whole-person stimuli and not to their component parts (Wachsmuth et al., 1994).

The recent study by Kaiser et al. (2014) mentioned above seems to contradict Schmalzl et al.'s findings of person-selective neural populations in the fusiform gyrus. Kaiser et al. showed that response patterns in the fusiform gyrus 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. They suggested that whole person responses in the fusiform gyrus primarily arise from the coactivation of independent face- and body-selective neural populations, rather than whole-person-selective neural populations. Kaiser et al. discuss some possible explanations for the contradiction between the two studies, for example, methodological differences (Schmalzl et al. used fMRI adaptation, whereas Kaiser et al. used multivariate pattern analysis) and attentional accounts. The findings of Kaiser et al. are also in contradiction with those of Song et al. (2013), though both studies employed similar paradigms. There are two major differences between these two studies: First, Kaiser et al. (2014) did not include a body-above-face condition as in Song et al. (2013) study, which is expected to generate a linear combination and therefore is an important control condition to which the face above a body condition should be compared. Second, the ROIs were defined differently: Song et al. defined the face-selective regions by the contrast of faces versus bodies, scenes, and objects and the body-selective regions by the contrast of bodies versus faces, scenes, and objects, as we also did in our study here. In contrast, Kaiser et al. did not localize the face- and body-selective regions, but rather defined "person-selective" regions by the contrast of persons versus cars. Given that they did present faces and bodies in isolation, it was important to examine their category-selective regions as well and assess the extent to which they respond similarly or differently to the person-selective areas.

One potential drawback of the competition paradigm that we applied here is that the sequential and simultaneous presentations differ in several ways including the length of stimulus presentation (1 sec in the sequential and 250 msec in the simultaneous), number of onsets or offsets (more for the sequential than the simultaneous), which could potentially account for the higher response to the sequential than simultaneous presentation (i.e., the competition effect). However, there are plenty of findings both in our study and in previous studies,

which suggest that presentation parameters alone may not account for the observed effect. First, as discussed above, several bottom-up and top-down effects (e.g., attention, distance among stimuli, pop-out, grouping) have been shown to decrease the competition effect, despite similar stimulus presentation parameters across all experiments (McMains & Kastner, 2010; Beck & Kastner, 2005, 2009; Kastner et al., 1998, 2001). Similar to previous studies, in our study the magnitude of the competition effect varied both across stimuli and across brain areas. In particular, the effect was abolished only when a face was presented above a body and only in the face and body areas in the fusiform gyrus, not in the lateral occipital cortex. Effects of stimulus presentation would generate similar effects across all stimuli and brain regions and therefore cannot account for the observed findings.

Our findings provide a possible neural locus for recent behavioral findings that show interactions between faces and bodies. Such interactions were shown in transfer of perceptual aftereffect between a headless body to a face (Ghuman et al., 2010) as well as in studies that revealed interaction between emotional information expressed by faces and bodies (Aviezer et al., 2008, 2012; Van den Stock et al., 2007; Meeren et al., 2005). These studies suggest that the representations of face and body in an image of a person are not independent but generate a new image in which both stimuli interact and influence the representation of the other. Further neuroimaging and electrophysiological studies, which has so far primarily focused on isolated faces and headless bodies, are needed to examine how these face body interactions are manifested at the neural level and at what stage of visual processing they occur. Finally, the majority of studies on person recognition focused on faces. Recent studies show that bodies may also contribute to person recognition in particular when face information is not available (Rice, Phillips, Natu, An, & O'Toole, 2013) or when dynamic rather than static stimuli are used (O'Toole et al., 2011). These findings further highlight the importance of studying both faces and bodies and the interaction among them to better understand how we recognize people.

To conclude, our findings clearly show that faces and bodies are processed interactively even in their category selective cortex. These findings indicate that better understanding of how faces and bodies are represented requires further research that includes both stimuli rather than presenting each of them in isolation (cf. O'Toole et al., 2011). More generally, these findings highlight the need to start studying the representation and interaction between multiple objects in object category selective cortex, which has been so far examined primarily with isolated objects.

Acknowledgments

This study was funded by Israeli Science Foundation grant 446/12 to G. Y.

Reprint requests should be sent to Galit Yovel, School of Psychological Sciences, Tel Aviv University, Tel Aviv, 69978, Israel, or via e-mail: gality@post.tau.ac.il.

REFERENCES

- Aviezer, H., Hassin, R. R., Ryan, J., Grady, C., Susskind, J., Anderson, A., et al. (2008). Angry, disgusted, or afraid? Studies on the malleability of emotion perception. *Psychological Science*, *19*, 724–732.
- Aviezer, H., Trope, Y., & Todorov, A. (2012). Body cues, not facial expressions, discriminate between intense positive and negative emotions. *Science*, *338*, 1225–1229.
- Beck, D. M., & Kastner, S. (2005). Stimulus context modulates competition in human extrastriate cortex. *Nature Neuroscience*, *8*, 1110–1116.
- Beck, D. M., & Kastner, S. (2007). Stimulus similarity modulates competitive interactions in human visual cortex. *Journal of Vision*, *7*, 19.1–19.12.
- Beck, D. M., & Kastner, S. (2009). Top-down and bottom-up mechanisms in biasing competition in the human brain. *Vision Research*, *49*, 1154–1165.
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*, 433–436.
- Brandman, T., & Yovel, G. (2012). A face inversion effect without a face. *Cognition*, *125*, 365–372.
- Brett, A., Anton, J., Valabregue, R., & Poline, J. (2002). Region of interest analysis using an SPM toolbox. *Neuroimage*, *16*, 1140–1141.
- Chan, A. W. Y., Kravitz, D. J., Truong, S., Arizpe, J., & Baker, C. I. (2010). Cortical representations of bodies and faces are strongest in commonly experienced configurations. *Nature Neuroscience*, *13*, 417–418.
- Cox, D., Meyers, E., & Sinha, P. (2004). Contextually evoked object-specific responses in human visual cortex. *Science*, *304*, 115–117.
- Desimone, R., & Duncan, J. (1995). Neural mechanisms of selective visual attention. *Annual Review of Neuroscience*, *18*, 193–222.
- Ghuman, A. S., McDaniel, J. R., & Martin, A. (2010). Face adaptation without a face. *Current Biology*, *20*, 32–36.
- Kaiser, D., Strnad, L., Seidl, K. N., Kastner, S., & Peelen, M. V. (2014). Whole person-evoked fMRI activity patterns in human fusiform gyrus are accurately modeled by a linear combination of face- and body-evoked activity patterns. *Journal of Neurophysiology*, *111*, 82–90.
- Kanwisher, N., & Yovel, G. (2006). The fusiform face area: A cortical region specialized for the perception of faces. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences*, *361*, 2109–2128.
- Kastner, S., De Weerd, P., Desimone, R., & Ungerleider, L. G. (1998). Mechanisms of directed attention in the human extrastriate cortex as revealed by functional MRI. *Science*, *282*, 108–111.
- Kastner, S., De Weerd, P., Pinsk, M. A., Elizondo, M. I., Desimone, R., & Ungerleider, L. G. (2001). Modulation of sensory suppression: Implications for receptive field sizes in the human visual cortex. *Journal of Neurophysiology*, *86*, 1398–1411.
- Korjoukov, I., Jeurissen, D., Kloosterman, N. A., Verhoeven, J. E., Scholte, H. S., & Roelfsema, P. R. (2012). The time course of perceptual grouping in natural scenes. *Psychological Science*, *23*, 1482–1489.
- Lerner, Y., Hendler, T., Ben-Bashat, D., Harel, M., & Malach, R. (2001). A hierarchical axis of object processing stages in the human visual cortex. *Cerebral Cortex*, *11*, 287–297.
- Lerner, Y., Hendler, T., & Malach, R. (2002). Object-completion effects in the human lateral occipital complex. *Cerebral Cortex*, *12*, 163–177.
- Luck, S. J., Chelazzi, L., Hillyard, S. A., & Desimone, R. (1997). Neural mechanisms of spatial selective attention in areas V1, V2, and V4 of macaque visual cortex. *Journal of Neurophysiology*, *77*, 24–42.
- MacEvoy, S. P., & Epstein, R. A. (2009). Decoding the representation of multiple simultaneous objects in human occipitotemporal cortex. *Current Biology*, *19*, 943–947.
- MacEvoy, S. P., & Epstein, R. A. (2011). Constructing scenes from objects in human occipitotemporal cortex. *Nature Neuroscience*, *14*, 1323–1329.
- McMains, S. A., & Kastner, S. (2010). Defining the units of competition: Influences of perceptual organization on competitive interactions in human visual cortex. *Journal of Cognitive Neuroscience*, *22*, 2417–2426.
- Meeren, H. K. M., van Heijnsbergen, C. C. R. J., & de Gelder, B. (2005). Rapid perceptual integration of facial expression and emotional body language. *Proceedings of the National Academy of Sciences, U.S.A.*, *102*, 16518–16523.
- Miller, E. K., Gochin, P. M., & Gross, C. G. (1993). Suppression of visual responses of neurons in inferior temporal cortex of the awake macaque by addition of a second stimulus. *Brain Research*, *616*, 25–29.
- O’Toole, A. J., Phillips, P. J., Weimer, S., Roark, D. A., Ayyad, J., Barwick, R., et al. (2011). Recognizing people from dynamic and static faces and bodies: Dissecting identity with a fusion approach. *Vision Research*, *51*, 74–83.
- Peelen, M. V., & Downing, P. E. (2007). The neural basis of visual body perception. *Nature Reviews Neuroscience*, *8*, 636–648.
- Reynolds, J. H., Chelazzi, L., & Desimone, R. (1999). Competitive mechanisms subserve attention in macaque areas V2 and V4. *Journal of Neuroscience*, *19*, 1736–1753.
- Rice, A., Phillips, P. J., Natu, V., An, X. B., & O’Toole, A. J. (2013). Unaware person recognition from the body when face identification fails. *Psychological Science*, *24*, 2235–2243.
- Roelfsema, P. R. (2006). Cortical algorithms for perceptual grouping. *Annual Review of Neuroscience*, *29*, 203–227.
- Roelfsema, P. R., & Houtkamp, R. (2011). Incremental grouping of image elements in vision. *Attention, Perception, & Psychophysics*, *73*, 2542–2572.
- Schmalzl, L., Zopf, R., & Williams, M. A. (2012). From head to toe: Evidence for selective brain activation reflecting visual perception of whole individuals. *Frontiers in Human Neuroscience*, *6*, article 108.
- Song, Y. Y., Luo, Y. L. L., Li, X. T., Xu, M., & Liu, J. (2013). Representation of contextually related multiple objects in the human ventral visual pathway. *Journal of Cognitive Neuroscience*, *25*, 1261–1269.
- Taylor, J. C., & Downing, P. E. (2011). Division of labor between lateral and ventral extrastriate representations of faces, bodies, and objects. *Journal of Cognitive Neuroscience*, *23*, 4122–4137.
- Taylor, J. C., Wiggett, A. J., & Downing, P. E. (2007). Functional MRI analysis of body and body part representations in the extrastriate and fusiform body areas. *Journal of Neurophysiology*, *98*, 1626–1633.
- Van den Stock, J., Righart, R., & de Gelder, B. (2007). Body expressions influence recognition of emotions in the face and voice. *Emotion*, *7*, 487–494.
- Wachsmuth, E., Oram, M. W., & Perrett, D. I. (1994). Recognition of objects and their component parts - Responses of single units in the temporal cortex of the macaque. *Cerebral Cortex*, *4*, 509–522.
- Weiner, K. S., & Grill-Spector, K. (2013). Neural representations of faces and limbs neighbor in human high-level visual cortex: Evidence for a new organization principle. *Psychological Research-Psychologische Forschung*, *77*, 74–97.