



Differential selectivity for dynamic versus static information in face-selective cortical regions

David Pitcher^{a,b,*}, Daniel D. Dilks^a, Rebecca R. Saxe^c, Christina Triantafyllou^d, Nancy Kanwisher^{a,c}

^a McGovern Institute for Brain Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

^b Institute of Cognitive Neuroscience, University College London, 17 Queen Square, London WC1N 3AR, UK

^c Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

^d Athinoula A. Martinos Imaging Center at McGovern Institute for Brain Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

ARTICLE INFO

Article history:

Received 28 October 2010

Revised 18 March 2011

Accepted 24 March 2011

Available online 5 April 2011

Keywords:

Dynamic stimuli

Face perception

fMRI

Superior temporal sulcus (STS)

Fusiform face area (FFA)

Occipital face area (OFA)

ABSTRACT

Neuroimaging studies have identified multiple face-selective regions in human cortex but the functional division of labor between these regions is not yet clear. A central hypothesis, with some empirical support, is that face-selective regions in the superior temporal sulcus (STS) are particularly responsive to dynamic information in faces, whereas the fusiform face area (FFA) computes the static or invariant properties of faces. Here we directly tested this hypothesis by measuring the magnitude of response in each region to both dynamic and static stimuli. Consistent with the hypothesis, we found that the response to movies of faces was not significantly different from the response to static images of faces from these same movies in the right FFA and right occipital face area (OFA). By contrast the face-selective region in the right posterior STS (pSTS) responded nearly three times as strongly to dynamic faces as to static faces, and a face-selective region in the right anterior STS (aSTS) responded to dynamic faces only. Both of these regions also responded more strongly to moving faces than to moving bodies, indicating that they are preferentially engaged in processing dynamic information from faces, not in more general processing of any dynamic social stimuli. The response to dynamic and static faces was not significantly different in a third face-selective region in the posterior continuation of the STS (pcSTS). The strong selectivity of face-selective regions in the pSTS and aSTS, but not the FFA, OFA or pcSTS, for dynamic face information demonstrates a clear functional dissociation between different face-selective regions, and provides further clues into their function.

© 2011 Elsevier Inc. All rights reserved.

Introduction

Functional magnetic resonance imaging (fMRI) studies of face perception reliably identify multiple face-selective cortical regions (Kanwisher and Yovel, 2006; Ishai, 2008; Fox et al., 2009; Pinsk et al., 2009), but the functional operations performed in these regions are not yet clearly understood. Two of the most commonly studied regions are the fusiform face area (FFA) (Kanwisher et al., 1997), found on the ventral surface of the occipitotemporal cortex, and a face-selective region in the posterior superior temporal sulcus (pSTS) (Phillips et al., 1997; Puce et al., 1998). In addition to being located in different areas of the brain, the FFA and pSTS are also thought to perform different functional roles in the perception of faces. The FFA has been implicated in the representation of static or invariant properties of faces (Kanwisher et al., 1997; McCarthy et al., 1997; Haxby et al., 2000), such as facial identity (Grill-Spector et al., 2004; Yovel and Kanwisher,

2004; Rotshtein et al., 2005), whereas the pSTS region has been implicated in the representation of the dynamic properties of faces (Allison et al., 2000; Haxby et al., 2000; Gobbini et al., in press), such as eye, mouth and head movements (Puce et al., 1998; Lee et al., 2010) and facial expression (Phillips et al., 1997; Winston et al., 2004). Fox et al. (2009) have further demonstrated that face-selective regions can be more robustly identified using dynamic stimuli compared with static stimuli (see also Hasson et al., 2010; Scherf et al., 2010). However, no prior study has directly compared the response to dynamic versus static stimuli in these and other face-selective regions using data independent of that used to define the regions of interest (Saxe et al., 2006; Vul and Kanwisher, 2011). Here we do just that, by measuring the response of each face-selective region to short movies of faces, bodies, scenes, objects and scrambled objects and to static images taken from these same movies. In addition to the widely studied FFA, occipital face area (OFA) (Gauthier et al., 2000), and pSTS regions, we also examined the response profile of face-selective regions in the posterior continuation of the STS (pcSTS), the anterior STS (aSTS) (Pinsk et al., 2009; Said et al., 2010), inferior frontal gyrus (IFG) (Ishai et al., 2002; Fox et al., 2009), and motor cortex (Adolphs, 2002; Keysers et al., 2010).

* Corresponding author at: McGovern Institute for Brain Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA. Fax: +1 617 258 8654.
E-mail address: dpitcher@mit.edu (D. Pitcher).

Method

Participants

Fourteen individuals (all right-handed; seven females) participated in this experiment. Participants were college or graduate students in the Boston area. All were neurologically normal and were paid for their participation. Informed consent was obtained and the committee on the use of humans as experimental participants at the Massachusetts Institute of Technology approved all procedures. One male participant was discarded from further analysis after moving more than 3 mm over the course of the scanning session.

Stimuli

Dynamic stimuli were 3-s movie clips of faces, bodies, scenes, objects and scrambled objects. There were 60 movie clips for each category. Movies of faces and bodies were filmed on a black background, and framed close-up to reveal only the faces or bodies of 7 children as they danced or played with toys or adults (who were out of frame). Fifteen different locations were used for the scene stimuli which were mostly pastoral scenes shot from a car window while driving slowly through leafy suburbs, along with some other films taken while flying through canyons or walking through tunnels that were included for variety. Fifteen different moving objects were selected that minimized any suggestion of animacy of the object itself or of a hidden actor pushing the object (these included mobiles, windup toys, toy planes and tractors, balls rolling down sloped inclines). Note that the smaller number of individuals in the face (7) and body (7) movie clips compared to the number of object (15) and scene (15) movie clips is conservative with respect to our hypothesis regarding face selectivity. Scrambled objects were constructed by dividing each object movie clip into a 15 by 15 box grid and spatially rearranging the location of each of the resulting movie frames. Within each block stimuli were randomly selected from within the entire set for that stimulus category (faces, bodies, scenes, objects, scrambled objects). This meant that the same movie clip could appear within the same block but given the number of stimuli this did not occur frequently.

Static stimuli were identical in design to the dynamic stimuli except that in place of each 3-s movie we presented three different still images taken from the beginning, middle and end of the corresponding movie clip (see [Supplemental Fig. 1](#)). Each image was presented for 1 s with no ISI, to equate the total presentation time with the corresponding dynamic movie clip.

Procedure

Each scanning session began with the acquisition of a high-resolution T-1 weighted anatomical scan. Functional data were acquired over 12 blocked-design functional runs lasting 234 s each. Each functional run contained three 18-s rest blocks, at the beginning, middle, and end of the run, during which a series of six uniform color fields were presented for 3 s each (these color fields were designed to maintain the interest of children, for whom the dynamic localizer was originally designed, while approximating a fixation baseline condition by avoiding any pattern visual input). Participants were instructed to watch the movies and static images but were not asked to perform any overt task. We did not use eye tracking because any differences in eye movements between dynamic and static stimuli would likely be a main effect and we predicted that only dynamic faces would produce a greater response in face-selective STS regions.

Each run contained two sets of five consecutive stimulus blocks (faces, bodies, scenes, objects or scrambled objects) sandwiched between these rest blocks, to make two blocks per stimulus category per run. Each block lasted 18 s and contained stimuli from one of the

five stimulus categories. The order of stimulus category blocks in each run was palindromic (e.g., fixation, faces, objects, scenes, bodies, scrambled objects, fixation, scrambled objects, bodies, scenes, objects, faces, fixation) and was randomized across runs. Functional runs presented either movie clips (the eight dynamic runs) or sets of static images taken from the same movies (the four static runs). For the dynamic runs, each 18-s block contained six 3-s movie clips from that category. For the static runs, each 18-s block contained eighteen 1-s still snapshots, composed of six triplets of snapshots taken at 1-s intervals from the same movie clip. Dynamic/static runs were run in the following order: 4 dynamic, 2 static, 2 dynamic, 2 static, 2 dynamic. The first 4 runs of the dynamic stimuli were used to define the studied ROIs (see [Data analysis](#) section).

Brain imaging

Scanning was performed in a 3.0 T Siemens Trio scanner at the A. A. Martinos Imaging Center at the McGovern Institute for Brain Research at the Massachusetts Institute of Technology. Functional images were acquired with a Siemens 32-channel phased array head-coil and a gradient-echo EPI sequence (32 slices, repetition time (TR) = 2 s, echo time = 30 ms, voxel size = 3 × 3 × 3 mm, and 0.6 mm interslice gap) providing whole brain coverage (slices were aligned with the anterior/posterior commissure). In addition, a high-resolution T-1 weighted MPRAGE anatomical scan was acquired for anatomically localizing the functional activations.

Data analysis

Data were analyzed with FS-FAST, Freesurfer (<http://www.surfer.nmr.mgh.harvard.edu/>) (Dale et al., 1999; Fischl et al., 1999). Before statistical analysis, images were motion corrected (Cox and Jesmanowicz, 1999), and (for those runs used in defining ROIs only) smoothed (3 mm FWHM Gaussian kernel), detrended, and fit using a gamma function ($\delta = 2.25$ and $\tau = 1.25$). The pre-processing did not involve any spatial normalization of participants in a common reference space (e.g., Talairach transformations). The first four dynamic runs were used to define ROIs using a contrast of dynamic faces greater than dynamic objects, using the same statistical threshold ($p = 10^{-4}$, uncorrected) for all participants. Within each functionally defined ROI we then calculated the magnitude of response (percent signal change, or PSC, from a fixation baseline) to the dynamic and static conditions of each of the five stimulus categories (faces, bodies, scenes, objects and scrambled objects), using the data collected from runs 5–12 in which pairs of dynamic and static runs were alternated. All of the data used to calculate PSC was independent of the data used to define the ROIs (Saxe et al., 2006; Vul and Kanwisher, 2011).

Results

Identifying ROIs

Face-selective ROIs in the right hemisphere were identified based on the data from four runs of the dynamic localizer using a contrast of dynamic faces greater than dynamic objects. The most robust ROIs across participants were found in the mid-fusiform gyrus (FFA) in 13/13 participants, in or around the inferior occipital gyrus (OFA) in 13/13 participants, at or slightly anterior to the junction of the ascending and descending limbs of the STS (posterior STS, or pSTS) in 13/13 participants, and in an anterior region of the STS (aSTS) in 11/13 participants. To illustrate the location of these face-selective ROIs we have included data from three example participants (see [Supplemental Fig. 2](#)). In addition, we also identified face-selective ROIs that were less robust across participants in the inferior frontal gyrus (IFG) (in 7/13 participants), in the posterior continuation of the STS (in or around

the ascending and descending limbs of the STS, referred to here as the pcSTS, in 6/13 participants), and in motor cortex (in 6/13 participants). The Talairach co-ordinates of the peak voxel in each ROI for all participants are shown in Table 1.

Consistent with prior literature face-selective ROIs in the left hemisphere (LH) were less robust across participants (Kanwisher et al., 1997; Barton et al., 2002; Young et al., 1985; Pitcher et al., 2007). The number of participants exhibiting face-selective ROIs in the LH is shown in Table 1.

Face-selective ROIs were more robustly identified using dynamic images than static images (see also Fox et al., 2009). The number of face-selective ROIs identified using dynamic versus static stimuli is reported in Table 1. Importantly, however, the pattern of the response to all stimulus categories within a given ROI did not differ when we defined face-selective regions using dynamic stimuli or static stimuli (see supplemental materials).

ROI response profiles

To begin, we examined the response profiles of each of the most robust face-selective ROIs (i.e., rFFA, rOFA, rpSTS, and raSTS) to short movies of faces, bodies, scenes, objects and scrambled objects and to static images taken from these same movies using the independently calculated PSC data (see Fig. 1).

rFFA

Consistent with the hypothesis that the rFFA is sensitive to the static or invariant properties of faces, we found no difference in the magnitude of the response to dynamic versus static faces in the rFFA. A 2 (motion: dynamic, static) \times 5 (category: faces, bodies, scenes, objects, scrambled objects) repeated-measures ANOVA (Greenhouse–Geisser corrected) revealed a significant main effect of category ($F(4, 48) = 83.41$, $p < 0.001$), with a significantly greater response to faces than any other category (Bonferroni corrected post-hoc comparisons, all $ps < 0.05$), demonstrating the known face selectivity of this region. Crucially, we found no significant main effect of motion ($F(1, 12) = 1.37$, $p = 0.27$), and no significant interaction between motion and category ($F(4, 48) = 1.86$, $p = 0.18$), demonstrating that there was no significant difference in the response to dynamic versus static faces in the rFFA.

rOFA

As with the rFFA, the rOFA did not exhibit a differential response to dynamic and static faces. A 2 (motion: dynamic, static) \times 5 (category: faces, bodies, scenes, objects, scrambled objects) repeated-measures ANOVA (Greenhouse–Geisser corrected) revealed no significant main effect of motion ($F(1, 12) = 2.52$, $p = 0.14$), but did show a significant main effect of category ($F(4, 48) = 25.08$, $p < 0.0001$), with a significantly greater response to faces than to any other category (Bonferroni corrected post-hoc comparisons, all $ps < 0.05$). There was also no significant interaction between motion and category ($F(4, 48) = 2.20$, $p = 0.14$). These results demonstrate the face selectivity of the rOFA, and also show that there was no significant difference in the response to dynamic versus static faces in the rOFA.

rpSTS

Consistent with the hypothesis that face-selective regions in the STS are particularly responsive to dynamic information in faces, we found a significantly greater response to dynamic faces than static faces in the rpSTS. A 2 (motion: dynamic, static) \times 5 (category: faces, bodies, scenes, objects, scrambled objects) repeated-measures ANOVA (Greenhouse–Geisser corrected) revealed a significant main effect of motion ($F(1, 12) = 35.93$, $p < 0.001$), with a significantly greater response to dynamic than static stimuli. We also found a significant main effect of category ($F(4, 48) = 54.62$, $p < 0.001$), with a significantly greater response to faces than to all other categories (Bonferroni corrected post-hoc comparisons, all $ps < 0.05$), demonstrating the face selectivity of this region. Importantly, we also found a significant interaction between motion and category ($F(4, 48) = 27.08$, $p < 0.001$). Bonferroni corrected post-hoc comparisons revealed that dynamic faces produced a significantly greater response than static faces ($t(12) = 8.5$, $p < 0.001$), revealing the strong dependence of the rpSTS on dynamic information from faces. Bonferroni corrected post-hoc tests also revealed a significantly greater response to dynamic bodies compared to static bodies ($t(12) = 7.3$, $p < 0.001$).

raSTS

As with the rpSTS, we found a significantly greater response to dynamic than static faces in the raSTS. A 2 (motion: dynamic, static) \times 5 (category: faces, bodies, scenes, objects, scrambled objects) repeated-

Table 1
Table showing the number of face-selective ROIs (faces greater than objects) identified across participants.

ROI table	FFA	OFA	pSTS	aSTS	IFG	pcSTS	Motor cortex
Right hemisphere 4 dynamic runs	13/13	13/13	13/13	11/13	7/13	6/13	6/13
Right hemisphere 8 dynamic runs	13/13	13/13	13/13	12/13	10/13	9/13	9/13
Right hemisphere 4 static runs	12/13	12/13	10/13	2/13	4/13	5/13	0/13
Left hemisphere 4 dynamic runs	11/13	8/13	9/13	2/13	0/13	7/13	1/13
Talairach co-ordinates							
	FFA	OFA	pSTS	aSTS	IFG	pcSTS	Motor cortex
Participant 1	42, -58, -18	44, 80, -4	60, -41, 7	55, 9, -18	60, 23, -6		46, 3, 43
Participant 2	42, -50, -13	42, -76, -8	57, -28, 1	62, -4, -17	56, 27, 4	59, -30, 15	50, 2, 44
Participant 3	39, -59, -18	40, -84, -3	61, -22, -4			65, -45, 22	
Participant 4	39, -48, -16	41, -83, 3	58, -46, 8				
Participant 5	42, -45, -17	36, -85, -2	53, -37, 9	53, -2, -19		42, -56, 15	
Participant 6	43, -48, -16	45, -74, -10	54, -40, 3	64, 6, -19		59, -51, 22	
Participant 7	40, -50, -13	40, 79, -10	51, -34, 0	62, -9, -3	46, 30, 5		54, -4, 45
Participant 8	43, -50, 11	41, -76, -1	48, -40, 10	56, 9, -8	47, 27, 1		
Participant 9	45, -44, -15	41, -78, 5	51, -40, 15	62, -1, -11	48, 31, -8	50, -54, 10	50, 9, 43
Participant 10	42, -47, -16	42, -76, -13	57, -35, 1	61, -7, -10			45, -7, 41
Participant 11	40, -51, -15	39, -80, -7	49, -40, 4	62, -8, -16	47, 25, -2		
Participant 12	43, -40, -24	51, -77, 7	53, -43, 2	57, 1, -23	56, 22, 5	63, -55, 7	56, 8, 29
Participant 13	40, -45, -15	39, 75, 5	54, -43, 0	60, -4, -16			

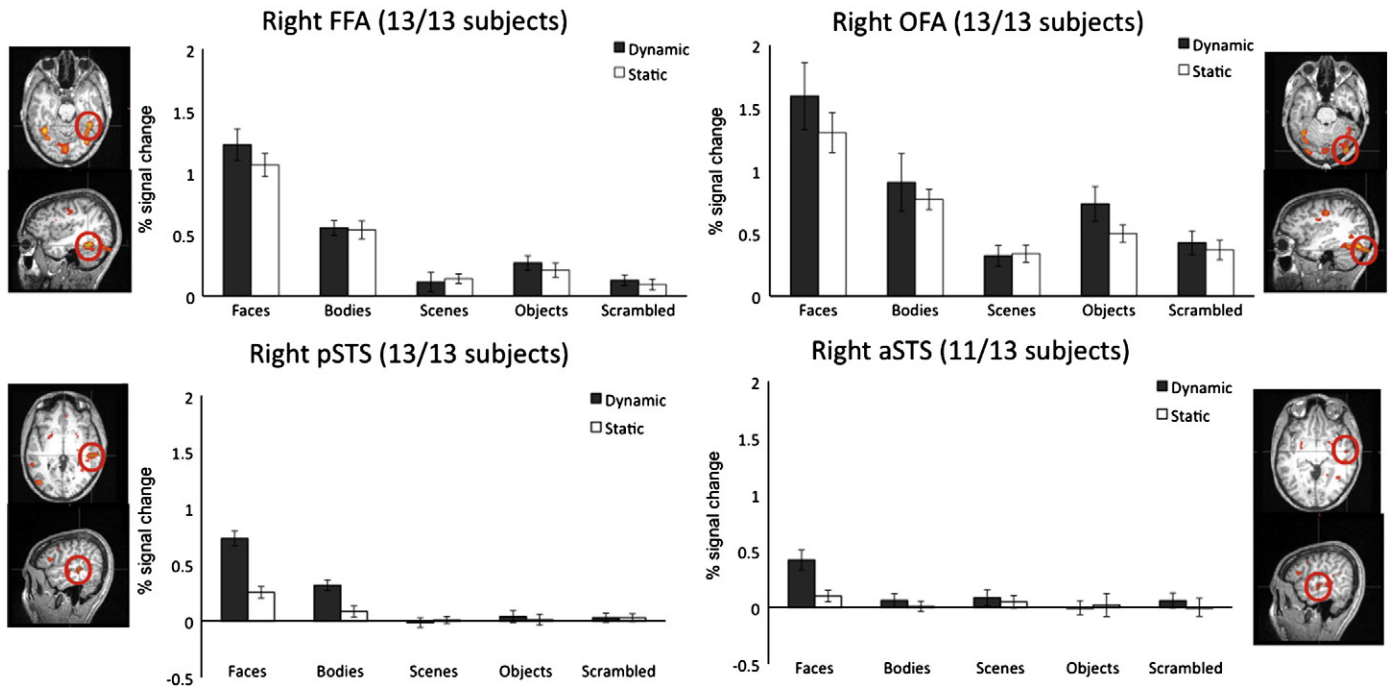


Fig. 1. Percent signal change (PSC) data for the dynamic and static stimuli from all five categories (faces, bodies, scenes, objects and scrambled objects) in the rFFA, rOFA, rpSTS and raSTS. All four regions showed a significantly greater response to faces than all other categories. The rpSTS and raSTS also showed a significantly greater response to dynamic faces than to static faces. Data shown are independent of the data used to define the ROIs.

measures ANOVA (Greenhouse–Geisser corrected) revealed no significant main effect of motion ($F(1, 10) = 1.46, p = 0.25$), but a significant main effect of category ($F(4, 40) = 11.56, p < 0.001$), with a significantly greater response to faces than any other category (Bonferroni corrected post-hoc comparisons, all $ps < 0.05$). We also found a significant interaction between motion and category ($F(4, 40) = 3.7, p = 0.03$). Bonferroni corrected post-hoc tests revealed that dynamic faces produced a significantly greater response than static faces ($t(10) = 5.3, p = 0.002$). These results show both the face selectivity of the region as well as the specificity to dynamic information in faces in the raSTS.

Next, we examined the response profiles of the three less robust face-selective ROIs (i.e., rpcSTS, rIFG, and right motor cortex) to short movies of faces, bodies, scenes, objects and scrambled objects and to static images from the same movies (see Fig. 2).

rpcSTS

Unlike the other face-selective regions in the STS, the rpcSTS did not exhibit a greater response to dynamic than static faces. A 2 (motion: dynamic, static) \times 5 (category: faces, bodies, scenes, objects, scrambled objects) repeated-measures ANOVA (Greenhouse–Geisser corrected) revealed no significant main effect of motion ($F(1, 5) = 0.01, p = 0.93$), but did show a significant main effect of category ($F(4, 20) = 6.56, p = 0.04$), with a significantly greater response to faces than to scenes, objects and scrambled objects (Bonferroni corrected post-hoc comparisons, all $ps < 0.05$). We did not find a significant interaction between motion and category ($F(4, 20) = 0.87, p = 0.47$). Finally, a direct contrast of moving faces versus static faces found no significant difference ($t(5) = 1.9, p = 0.11$). This pattern of results demonstrates that there was no significant difference in the response to dynamic and static faces in the rpcSTS.

rIFG

We found a significantly greater response to dynamic than static faces in the rIFG. A 2 (motion: dynamic, static) \times 5 (category: faces, bodies, scenes, objects, scrambled objects) repeated-measures ANOVA (Greenhouse–Geisser corrected) revealed no significant main effect of motion ($F(1, 6) = 0.78, p = 0.41$), but did show a

significant main effect of category ($F(4, 24) = 7.82, p = 0.007$), with a significantly greater response to faces than to scenes, objects and scrambled objects (Bonferroni corrected comparisons, all $ps < 0.05$). We also found a significant interaction between motion and category ($F(4, 24) = 3.66, p = 0.05$). A Bonferroni corrected post-hoc test revealed that dynamic faces produced a significantly greater response than static faces ($t(6) = 2.5, p = 0.047$).

Right motor cortex

The face-selective region in the right motor cortex exhibited a greater response to dynamic than static faces. A 2 (motion: dynamic, static) \times 5 (category: faces, bodies, scenes, objects, scrambled objects) repeated-measures ANOVA (Greenhouse–Geisser corrected) revealed no significant main effect of motion ($F(1, 5) = 4.16, p = 0.1$), but did show a significant main effect of category ($F(4, 20) = 9.83, p = 0.014$), with a significantly greater response to faces than to objects and scrambled objects (Bonferroni corrected comparisons, all $ps < 0.05$). We also found a significant interaction between motion and category ($F(4, 20) = 4.4, p = 0.05$). Bonferroni corrected post-hoc tests revealed that dynamic faces produced a significantly greater response than static faces ($t(5) = 5.3, p = 0.003$) and that dynamic bodies produced a greater response than static bodies ($t(5) = 3.2, p = 0.022$).

Functional dissociation across face-selective regions

The above analyses suggest a functional dissociation across the most robust face-selective regions, with the rFFA and rOFA exhibiting similar responses to dynamic and static faces, while the rpSTS and raSTS exhibited a strong selectivity for dynamic faces over static faces. To directly test for this functional dissociation, we conducted a 4 (ROI: rFFA, rOFA, rpSTS, raSTS) \times 2 (motion: dynamic, static) \times 2 (category: faces, objects) repeated-measures ANOVA (Greenhouse–Geisser corrected). We found a main effect of category ($F(1, 10) = 136, p < 0.001$) demonstrating a larger overall response to faces than to objects. Critically, we also found a significant 3-way interaction between ROI, motion, and category ($F(3, 30) = 7.35, p = 0.011$).

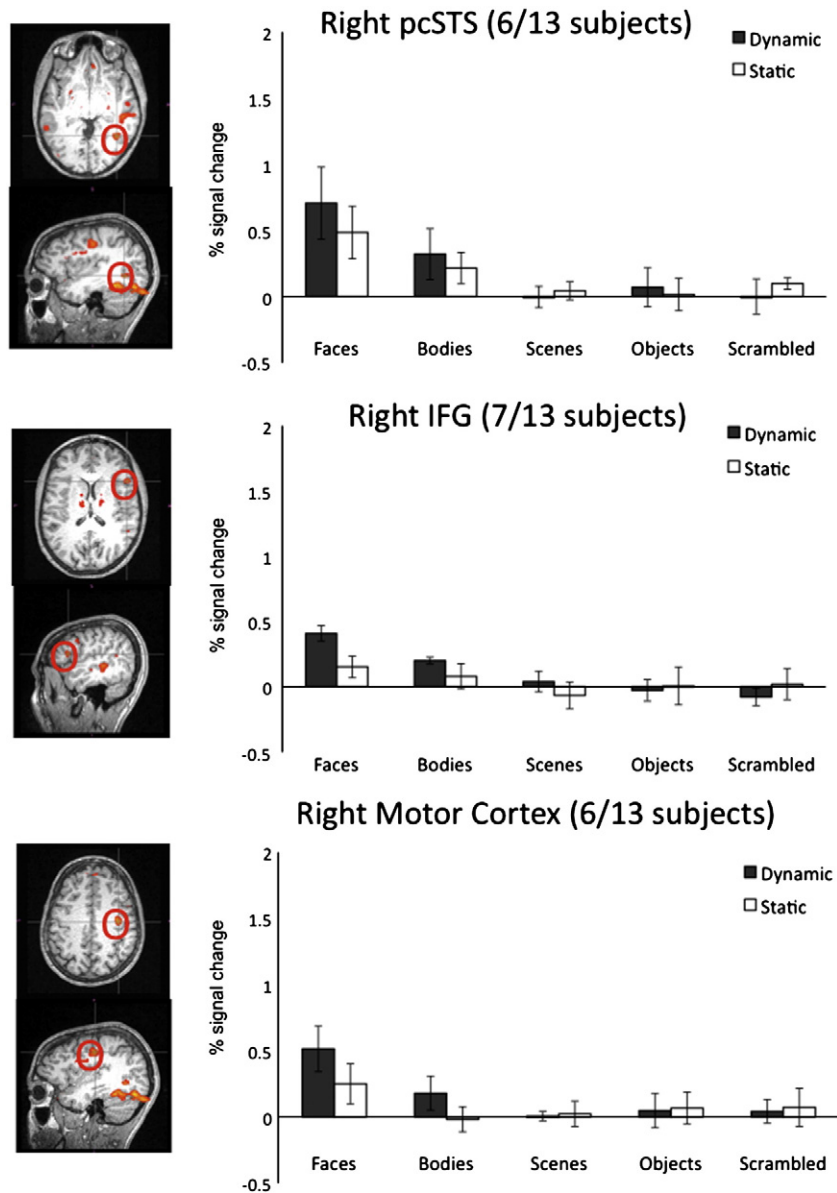


Fig. 2. Percent signal change (PSC) data for the dynamic and static stimuli from all five categories (faces, bodies, scenes, objects and scrambled objects) in the rpcSTS, right IFG and right motor cortex. All three regions showed a significantly greater response to faces than all other categories. The right IFG and motor cortex also showed a significantly greater response to dynamic faces than to static faces. Data shown are independent of the data used to define the ROIs.

To further understand what effects were driving this three-way interaction we conducted a separate 4 (ROI: rFFA, rOFA, rpSTS, raSTS) \times 2 (motion: dynamic, static) repeated-measures ANOVA (Greenhouse–Geisser corrected) on the face data only. Results showed main effects of ROI ($F(3, 30) = 18.1, p < 0.001$) and motion ($F(1, 10) = 13.7, p = 0.004$) as well as a significant two-way interaction ($F(3, 30) = 3.7, p = 0.047$). As predicted by our hypothesis Bonferroni corrected post-hoc tests demonstrated that dynamic faces produced a greater response than static faces in the rpSTS ($t(10) = 7.2, p < 0.001$) and in the raSTS ($t(10) = 4.3, p = 0.002$) but not in the rFFA ($p = 0.26$) or rOFA ($p = 0.16$). This analysis reveals a significant functional dissociation between the rFFA and rOFA (which do not show different selectivity for dynamic and static faces), and the rpSTS and raSTS (which responded more strongly to dynamic than to static faces).

Functional dissociation between STS regions

In addition to the most robust face-selective STS regions described above, we also identified the right pcSTS region in 6 of 13 participants.

As discussed above, the response profile of this region differed from the other two STS regions (see Fig. 3). To directly test this apparent functional dissociation between the rpSTS and raSTS regions, which exhibited a strong preference for dynamic faces, and the pcSTS, which did not differentiate between dynamic and static faces, we conducted a 3 (ROI: rpcSTS, rpSTS, raSTS) \times 2 (motion: dynamic, static) \times 2 (category: faces, objects) repeated-measures ANOVA (Greenhouse–Geisser corrected). Before doing so, however, we wanted to increase the number of participants exhibiting the rpcSTS region (present in only 6/13 participants in the analysis reported above), and thus used six dynamic runs (runs 1, 2, 3, 4, 11 and 12) to define the STS ROIs instead of four runs as used in the previous analysis. This increased the total number of participants exhibiting the rpcSTS from 6/13 to 9/13. The remaining two dynamic runs (7 and 8) together with two static runs (6 and 9) were then used to independently calculate the magnitude of response to all stimulus categories in each ROI.

The ANOVA on the nine participants in whom all three STS ROIs could be localized revealed a main effect of category ($F(1, 8) = 45.5, p < 0.001$) demonstrating a larger overall response to faces than to

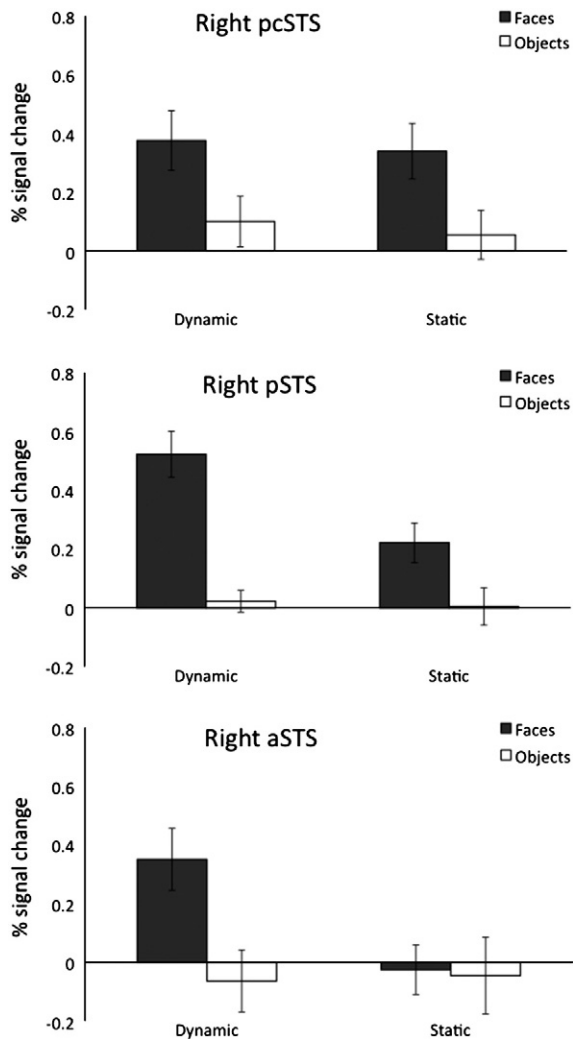


Fig. 3. Percent signal change (PSC) data for the dynamic and static stimuli from the face and object categories in the rpcSTS, rpSTS and raSTS. Results showed a functional dissociation between these three regions, with a significantly greater response to dynamic faces than static faces in the rpSTS and raSTS but not in the rpcSTS. Data shown are independent of the data used to define the ROIs.

objects. Crucially, we also found a significant 3-way interaction between ROI, motion, and category ($F(2, 16) = 4.9, p = 0.043$).

To further understand what effects were driving this three-way interaction we conducted a separate 3 (ROI: rpcSTS, rpSTS, raSTS) \times 2 (motion: dynamic, static) repeated-measures ANOVA (Greenhouse–Geisser corrected) on the face data only. Results showed a main effect of motion ($F(1, 8) = 15.1, p = 0.005$) but not of ROI ($F(2, 16) = 1.2, p = 0.33$). Crucially there was a significant two-way interaction ($F(2, 16) = 10.5, p = 0.002$). Bonferroni corrected post-hoc tests demonstrated that dynamic faces produced a greater response than static faces in the rpSTS ($t(8) = 5.0, p = 0.001$) and in the raSTS ($t(8) = 3.6, p = 0.006$) but not in the rpcSTS ($p = 0.6$). This analysis reveals a significant functional dissociation between the rpcSTS (which does not show different selectivity for dynamic and static faces) and the two other face-selective STS regions (which respond more selectively to dynamic than static faces).

Discussion

Extensive prior evidence has led to the hypothesis that face-selective regions in the STS preferentially represent the dynamic aspects of a face while face-selective regions in the fusiform gyrus preferentially represent the static or invariant aspects of a face (Puce

et al., 1998; Allison et al., 2000; Haxby et al., 2000; Grill-Spector et al., 2004; Andrews and Ewbank, 2004). In the present study, we used fMRI to quantitatively test this claim by systematically examining how face-selective ROIs responded to movies of faces, bodies, scenes, objects and scrambled objects and to static images taken from these same movies. Results demonstrated that while dynamic faces did not produce a significantly greater response than static faces in the rFFA and rOFA, the rpSTS and raSTS regions showed a substantially greater response to dynamic faces more than static faces. This preference for dynamic faces over static faces was most striking in the raSTS where dynamic faces were the only stimuli that produced any significant response above the fixation baseline condition. This functional dissociation between face-selective regions, some strongly selective for dynamic face information and others which respond equally to dynamic and static faces, is consistent with cognitive and neural models of face perception, which propose that different aspects of face perception (such as identity or expression discrimination) are preferentially processed in different cortical regions (Bruce and Young, 1986; Haxby et al., 2000; Ishai, 2008; but see Calder and Young, 2005). Indeed, given the demonstrable role of the STS in social perception (Allison et al., 2000), and of the FFA for identity discrimination (Grill-Spector et al., 2004; Yovel and Kanwisher, 2004), it seems plausible that face-selective face regions on the ventral cortical surface compute who the face is, while face-selective regions in the STS compute what the face is doing.

The face-selective ROI we report in the posterior STS is in an area of cortex thought to be involved in a variety of perceptual and social cognitive operations (Allison et al., 2000; Hein and Knight, 2008). Consistent with the strong preference for dynamic faces reported here, the STS has been implicated in face processing functions that depend on the moving components of a face (e.g., the eyes and the mouth), including facial expression processing (Winston et al., 2004; Andrews and Ewbank, 2004; Engell and Haxby, 2007; Said et al., 2010), gaze discrimination (Hoffman and Haxby, 2000; Pelphrey et al., 2003a, 2003b, 2005; Engell and Haxby, 2007), and the perception of eye and mouth movements (Puce et al., 1998). It therefore seems likely that the facial expressions and gaze shifts combined with the movement displayed by the actors in our dynamic face stimuli contributed to the elevated response we observed for dynamic faces in the rpSTS region. Note however that our results show more than just a response to gaze and emotion information, which is also present in the static images in our study. Rather, we show that if gaze and expression information is extracted in these regions, it is specifically dynamic changes in gaze or expression that the region cares most about. Consistent with this hypothesis are psychophysical results showing that emotional expression information is better extracted from dynamic than static face images (Ambadar et al., 2005).

Despite the converging evidence for a cortical region specialized for dynamic face processing in the STS it is important to note that other studies report that this general area of cortex is engaged in other types of cognitive operations that do not explicitly involve face perception. These include biological motion perception (Grossman and Blake, 2002; Beauchamp et al., 2003; Pelphrey et al., 2003a, 2003b), the perception of goal directed actions (Saxe et al., 2004; Pelphrey et al., 2004; Brass et al., 2007; Vander Wyk et al., 2009), and body perception (Kontaris et al., 2009). The diversity of these tasks has led to claims that the STS does not contain areas specialized for particular cognitive operations, but rather that the area is engaged in generalized processing dependent on particular task requirements (Hein and Knight, 2008). Our data suggest instead that there may be discrete cortical regions along the STS that are specialized for particular cognitive operations. The strong preference we find in the rpSTS and raSTS for dynamic faces compared to nine other stimulus categories (including dynamic bodies and static faces), and the contrasting response profile of the rpcSTS, demonstrates that functional dissociations can be found in this region. Future work will address

whether these ROIs, selective for dynamic faces in the STS, are distinct from regions previously implicated with biological motion perception (Grossman and Blake, 2002), action understanding (Saxe et al., 2004) and theory of mind tasks (Saxe and Kanwisher, 2003).

The face-selective region in the anterior STS was less reliably identified than the posterior STS region (11/13 versus 13/13 participants, respectively). The raSTS was smaller than the rpSTS, but was anatomically consistent across participants. The raSTS region was highly selective for dynamic faces only, probably explaining why it has not been extensively reported in previous face perception fMRI studies that have predominantly used static face stimuli. Despite the strong preference of the raSTS for dynamic faces, a recent fMRI adaptation study of gaze perception that used static stimuli identified a region similar to the one we report here that was able to code gaze direction (Calder et al., 2007). A similar region was also reported in a recent study that compared face-selective ROIs between humans and macaques (Pinsk et al., 2009).

The face-selective region in the posterior continuation of STS (rpcSTS) was reliably identified in less than half of participants (6/13) when 4 runs were used to define the ROI, but this proportion increased to 9 out of 13 participants when we used 6 runs. The functional profile of this region contrasted sharply with that of the rpSTS and raSTS regions by responding similarly to dynamic and static faces.

The response to dynamic and static faces in the rFFA and rOFA was not significantly different, demonstrating that these regions are not specifically engaged in extracting dynamic information from faces. This result is consistent with prior evidence that the FFA and OFA represent the invariant properties of a face, such as identity, for which motion is unlikely to enhance discrimination (Haxby et al., 2000; Grill-Spector et al., 2004; Yovel and Kanwisher, 2004; Rotshtein et al., 2005; Pitcher et al., 2009).

Other regions worth further investigation in the future are the face-selective regions that we, and others, report in the right IFG (Ishai et al., 2002; Fox et al., 2009) and in motor cortex (Adolphs, 2002; Keysers et al., 2010). Our data show that these regions are strongly selective for dynamic faces compared to static faces as previously argued (Fox et al., 2009). Crucial questions for these regions concern why they are present in only some participants, and what function they perform.

In the present study we tested our prediction that dynamic faces would produce a greater response than static faces in face-selective regions in the STS, but that the difference between dynamic and static faces would be less apparent in other face-selective regions such as the FFA. Our hypothesis was strongly supported by the response profiles observed in the rFFA, rOFA, rpSTS and raSTS regions. Further, an additional face-selective STS region, the rpcSTS, did not differentiate between dynamic and static faces. This surprising dissociation supports the hypothesis that different types of face computations might be performed in each of these face-selective STS regions (Freiwald and Tsao, 2010), and suggests future approaches for establishing the nature of these computations. Another key question for future research concerns the relationship between the face-selective regions in the STS reported here, and other cognitive functions that have been attributed to nearby or possibly overlapping cortical regions, such as biological motion perception, action understanding, and theory of mind.

Supplementary materials related to this article can be found online at doi:10.1016/j.neuroimage.2011.03.067.

References

- Adolphs, R., 2002. Recognizing emotion from facial expressions: psychological and neurological mechanisms. *Behav. Cognit. Neurosci. Rev.* 1, 21–61.
- Allison, T., Puce, A., McCarthy, G., 2000. Social perception from visual cues: role of the STS region. *Trends Cogn. Sci.* 4, 267–278.
- Ambadar, Z., Schooler, J.W., Cohn, J.F., 2005. Deciphering the enigmatic face: the importance of facial dynamics in interpreting subtle facial expressions. *Psych. Sci.* 16, 403–410.
- Andrews, T.J., Ewbank, M.P., 2004. Distinct representations for facial identity and changeable aspects of faces in human visual cortex. *Neuroimage* 23, 905–913.
- Barton, J.J., Press, D.Z., Keenan, J.P., O'Connor, M., 2002. Lesions of the fusiform face area impair perception of facial configuration in prosopagnosia. *Neurology* 58, 71–78.
- Beauchamp, M.S., Lee, K.E., Haxby, J.V., Martin, A., 2003. fMRI responses to video and point-light displays of moving humans and manipulable objects. *J. Cogn. Neurosci.* 15, 991–1001.
- Brass, M., Schmitt, R., Spengler, S., Gergely, G., 2007. Investigating action understanding: inferential processes versus motor simulation. *Curr. Biol.* 17 (24), 2117–2121.
- Bruce, V., Young, A., 1986. Understanding face recognition. *Br. J. Psychol.* 77, 305–327.
- Calder, A.J., Young, A.W., 2005. Understanding the recognition of facial identity and facial expression. *Nat. Reviews Neurosci.* 6, 641–651.
- Calder, A.J., Beaver, J.D., Winston, J.S., Dolan, R.J., Jenkins, R., Eger, E., Henson, R.N.A., 2007. Separate coding of different gaze directions in the superior temporal sulcus and inferior parietal lobule. *Curr. Biol.* 17, 20–25.
- Cox, R.W., Jesmanowicz, A., 1999. Real-time 3D image registration for functional MRI. *Magnetic Resonance in Medicine* 42, 1014–1018.
- Dale, A.M., Fischl, B., Sereno, M.I., 1999. Cortical surface-based analysis. I. Segmentation and surface reconstruction. *Neuroimage* 9, 179–194.
- Engell, A., Haxby, J., 2007. Facial expression and gaze-direction in human superior temporal sulcus. *Neuropsychologia* 45, 3234–3241.
- Fischl, B., Sereno, M.I., Dale, A.M., 1999. Cortical surface-based analysis. II. Inflation, flattening, and a surface-based coordinate system. *Neuroimage* 9, 195–207.
- Fox, C.J., Iaria, G., Barton, J., 2009. Defining the face-processing network: optimization of the functional localizer in fMRI. *Hum. Brain Mapp.* 30, 1637–1651.
- Freiwald, W.A., Tsao, D.Y., 2010. Functional compartmentalization and viewpoint generalization within the macaque face-processing system. *Science* 330, 845–851.
- Gauthier, I., Tarr, M.J., Moylan, J., Skudlarski, P., Gore, J.C., Anderson, A.W., 2000. The fusiform “face area” is part of a network that processes faces at the individual level. *J. Cogn. Neurosci.* 12, 495–504.
- Gobbini, M.I., Gentili, C., Ricciardi, E., Bellucci, C., Salvini, P., Laschi, C., Guazzelli, M., Pietrini, P., in press. Distinct neural systems involved in agency and animacy detection. *J. Cogn. Neurosci.* doi:10.1162/jocn.2010.21574.
- Grill-Spector, K., Knouf, N., Kanwisher, N., 2004. The fusiform face area subserves face perception, not generic within-category identification. *Nat. Neurosci.* 7 (5), 555–562.
- Grossman, E., Blake, R., 2002. Brain areas active during visual perception of biological motion. *Neuron* 35 (6), 1157–1165.
- Hasson, U., Malach, R., Heeger, D., 2010. Reliability of cortical activity during natural stimulation. *Trends Cogn. Sci.* 14, 40–48.
- Haxby, J.V., Hoffman, E.A., Gobbini, M.I., 2000. The distributed human neural system for face perception. *Trends Cogn. Sci.* 4, 223–233.
- Hein, G., Knight, R.T., 2008. Superior temporal sulcus—it's my area: or is it? *J. Cogn. Neurosci.* 2125–2136.
- Hoffman, E.A., Haxby, J.V., 2000. Distinct representation of eye gaze and identity in the distributed human neural system for face perception. *Nat. Neurosci.* 3, 80–84.
- Ishai, A., 2008. Let's face it: it's a cortical network. *Neuroimage* 40, 415–419.
- Ishai, A., Haxby, J.V., Ungerleider, L.G., 2002. Visual imagery of famous faces: effects of memory and attention revealed by fMRI. *Neuroimage* 17, 1729–1741.
- Kanwisher, N., Yovel, G., 2006. The fusiform face area: a cortical region specialized for the perception of faces. *Philos. Trans. R Soc. Lond. B* 361, 2109–2128.
- Kanwisher, N., McDermott, J., Chun, M., 1997. The fusiform face area: a module in human extrastriate cortex specialized for the perception of faces. *J. Neurosci.* 17, 4302–4311.
- Keysers, C., Kaas, J., Gazzola, V., 2010. Somatosensation in social perception. *Nat. Rev. Neurosci.* 11, 417–428.
- Kontaris, J., Wiggert, A., Downing, P., 2009. Dissociation of extrastriate body- and biological-motion selective areas by manipulation of visual-motor congruency. *Neuropsychologia* 47, 3118–3124.
- Lee, L.C., Andrews, T.J., Johnson, S.J., Woods, W., Gouws, A., Green, G.G., Young, A.W., 2010. Neural responses to rigidly moving faces displaying shifts in social attention investigated with fMRI and MEG. *Neuropsychologia* 48, 477–490.
- McCarthy, G., Puce, A., Gore, J., Allison, T., 1997. Face-specific processing in the fusiform gyrus. *J. Cogn. Neurosci.* 9, 605–610.
- Pelphrey, K.A., Mitchell, T.V., McKeown, M., Goldstein, J., Allison, T., McCarthy, G., 2003a. Brain activity evoked by perception of human walking: controlling for meaningful coherent motion. *J. Neurosci.* 23, 6819–6825.
- Pelphrey, K.A., Singerman, J.D., Allison, T., McCarthy, G., 2003b. Brain activation evoked by the perception of gaze shifts: the influence of context. *Neuropsychologia* 41, 156–170.
- Pelphrey, K.A., Morris, J.P., McCarthy, G., 2004. Grasping the intentions of others: the perceived intentionality of an action influences activity in the superior temporal sulcus during social perception. *J. Cogn. Neurosci.* 16 (10), 1706–1716.
- Pelphrey, K.A., Morris, J.P., Michelich, C.R., Allison, T., McCarthy, G., 2005. Functional anatomy of biological motion perception in posterior temporal cortex: an fMRI study of eye, mouth, and hand movements. *Cereb. Cortex* 15 (12), 1866–1876.
- Phillips, M.L., Young, A.W., Senior, C., Brammer, M., Andrews, C., Calder, A.J., Bullmore, E.T., Perrett, D.I., Rowland, D., Williams, S.C.R., Gray, J.A., David, A.S., 1997. A specific neural substrate for perceiving facial expressions of disgust. *Nature* 389, 495–498.
- Pinsk, M.A., Arco, M., Weiner, K., Kalkus, J., Inati, S., Gross, C.G., Kastner, S., 2009. Neural representations of faces and body parts in macaque and human cortex: a comparative fMRI study. *J. Neurophysiol.* 101, 2581–2600.
- Pitcher, D., Walsh, V., Yovel, G., Duchaine, B., 2007. TMS evidence for the involvement of the right occipital face area in early face processing. *Curr. Biol.* 17 (18), 1568–1573.

- Pitcher, D., Charles, L., Devlin, J.T., Walsh, V., Duchaine, B., 2009. Triple dissociation of faces, bodies, and objects in extrastriate cortex. *Curr. Biol.* 19 (4), 319–324.
- Puce, A., Allison, T., Bentin, S., Gore, J.C., McCarthy, G., 1998. Temporal cortex activation in humans viewing eye and mouth movements. *J. Neurosci.* 18, 2188–2199.
- Rotshtein, P., Henson, R.N., Treves, A., Driver, J., Dolan, R.J., 2005. Morphing Marilyn into Maggie dissociates physical and identity face representations in the brain. *Nat. Neuro.* 8, 107–113.
- Said, C.P., Moore, C.D., Engell, A.D., Todorov, A., Haxby, J.V., 2010. Distributed representations of dynamic facial expressions in the superior temporal sulcus. *J. Vis.* 10.
- Saxe, R., Kanwisher, N., 2003. People thinking about thinking people: the role of the temporo-parietal junction in theory of mind. *Neuroimage* 19, 1835–1842.
- Saxe, R., Xiao, D.K., Kovacs, G., Perrett, D.I., Kanwisher, N., 2004. A region of right posterior superior temporal sulcus responds to observed intentional actions. *Neuropsychologia* 42, 1435–1446.
- Saxe, R., Brett, M., Kanwisher, N., 2006. Divide and conquer: a defense of functional localizers. *Neuroimage* 30, 1088–1096.
- Scherf, S., Luna, B., Minshew, N., Behrmann, M., 2010. Location, location, location: alterations in the functional topography of face- but not object- or place-related cortex in adolescents with autism. *Front Hum. Neurosci.* 26.
- Vander Wyk, B.C., Hudac, C.M., Carter, E.J., Sobel, D.M., Pelphrey, K.A., 2009. Action understanding in the superior temporal sulcus region. *Psychol. Sci.* 20, 771–777.
- Vul, E., Kanwisher, N., 2011. Begging the question: the non-independence error in fMRI data analysis. In: Hanson, S., Bunzl, M. (Eds.), *Foundations and Philosophy for Neuroimaging*, pp. 71–91.
- Winston, J.S., Henson, R.N.A., Fine-Goulden, M.R., Dolan, R.J., 2004. fMRI-adaptation reveals dissociable neural representations of identity and expression in face perception. *J. Neurophysiol.* 92, 1830–1839.
- Young, A.W., Hay, D.C., McWeeny, K.H., Ellis, A.W., Barry, C., 1985. Familiarity decisions for faces presented to the left and right cerebral hemispheres. *Brain Cogn.* 4, 439–450.
- Yovel, G., Kanwisher, N., 2004. Face perception; domain specific, not process specific. *Neuron* 44 (5), 889–898.