

A quantitative link between face discrimination deficits and neuronal selectivity for faces in autism[☆]



Xiong Jiang^a, Angela Bollich^{c,d}, Patrick Cox^a, Eric Hyder^a, Joette James^{c,d}, Saqib Ali Gowani^{e,g}, Nouchine Hadjikhani^{f,h}, Volker Blanzⁱ, Dara S. Manoach^e, Jason J.S. Barton^j, William D. Gaillard^{b,d}, Maximilian Riesenhuber^{a,*}

^a Department of Neuroscience, Georgetown University Medical Center, Washington, DC 20007, USA

^b Department of Neurology, Georgetown University Medical Center, Washington, DC 20007, USA

^c Center for Autism Spectrum Disorder, Children's National Medical Center, Washington, DC 20010, USA

^d Center for Neuroscience, Children's National Medical Center, Washington, DC 20010, USA

^e Department of Psychiatry, Massachusetts General Hospital, Boston, MA 02114, USA

^f Department of Radiology, Massachusetts General Hospital, Boston, MA 02114, USA

^g The Aga Khan University, Karachi 74800, Pakistan

^h Swiss Federal Institute of Technology, CH-1015 Lausanne, Switzerland

ⁱ University of Siegen, 57076 Siegen, Germany

^j University of British Columbia, Vancouver, Canada BC V6T 1Z4

ARTICLE INFO

Article history:

Received 26 September 2012

Received in revised form 8 February 2013

Accepted 13 February 2013

Available online 26 February 2013

Keywords:

Face
Autism
ASD
fMRI
fMRI-RA
Local correlation

ABSTRACT

Individuals with Autism Spectrum Disorder (ASD) appear to show a general face discrimination deficit across a range of tasks including social-emotional judgments as well as identification and discrimination. However, functional magnetic resonance imaging (fMRI) studies probing the neural bases of these behavioral differences have produced conflicting results: while some studies have reported reduced or no activity to faces in ASD in the Fusiform Face Area (FFA), a key region in human face processing, others have suggested more typical activation levels, possibly reflecting limitations of conventional fMRI techniques to characterize neuron-level processing. Here, we test the hypotheses that face discrimination abilities are highly heterogeneous in ASD and are mediated by FFA neurons, with differences in face discrimination abilities being quantitatively linked to variations in the estimated selectivity of face neurons in the FFA. Behavioral results revealed a wide distribution of face discrimination performance in ASD, ranging from typical performance to chance level performance. Despite this heterogeneity in perceptual abilities, individual face discrimination performance was well predicted by neural selectivity to faces in the FFA, estimated via both a novel analysis of local voxel-wise correlations, and the more commonly used fMRI rapid adaptation technique. Thus, face processing in ASD appears to rely on the FFA as in typical individuals, differing quantitatively but not qualitatively. These results for the first time mechanistically link variations in the ASD phenotype to specific differences in the typical face processing circuit, identifying promising targets for interventions.

© 2013 The Authors. Published by Elsevier Inc. All rights reserved.

1. Introduction

Effective processing of faces is essential for social interaction. While adults typically show very high accuracy in recognizing and

discriminating faces (Diamond and Carey, 1986), a number of studies have shown that individuals with Autism Spectrum Disorder (ASD) appear to show a general face discrimination deficit (Weeks and Hobson, 1987; Klin et al., 1999; Critchley et al., 2000; Adolphs et al., 2001; Snow et al., 2011). It has even been suggested (Dawson et al., 2005; Schultz, 2005) that deficits in face processing are at the root of social dysfunction in autism, and that these might cause the emergence of other aspects of autism. However, fMRI studies probing the neural bases of face processing deficits in ASD have produced conflicting results: While several studies (Schultz et al., 2000; Pierce et al., 2001; Grelotti et al., 2005; Humphreys et al., 2008) have reported reduced or no activation in the human "Fusiform Face Area"

[☆] This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial-ShareAlike License, which permits non-commercial use, distribution, and reproduction in any medium, provided the original author and source are credited.

* Corresponding author at: Department of Neuroscience, Georgetown University Medical Center, Research Building Room WP-12, 3970 Reservoir Rd. NW, Washington, DC 20007, USA. Tel.: +1 202 687 9198; fax: +1 202 784 3562.

E-mail address: mr287@georgetown.edu (M. Riesenhuber).

(FFA) (Kanwisher et al., 1997), a central brain region in human face processing (Grill-Spector et al., 2004; Jiang et al., 2006; Kanwisher and Yovel, 2006), others have suggested more typical activation levels (Pierce et al., 2004; Hadjikhani et al., 2007; Perlman et al., 2011). One reason for this disagreement in the literature might be the large phenotypic heterogeneity in autism (Geschwind, 2009), that has been shown to also extend to face processing abilities (Barton et al., 2004; Hedley et al., 2011). In addition, a fundamental problem limiting the ability of conventional fMRI to probe the neural bases of behavioral differences in ASD is that the average activation level observed in fMRI experiments may not necessarily correlate with behavioral performance. FMRI studies of face processing have commonly focused on comparing the average BOLD-contrast response to different classes of stimuli (e.g., faces vs. houses). However, because of the limited spatial resolution of fMRI, relating BOLD-contrast signal change, behavioral performance, and assessments of neural population activity are complicated as the density of selective neurons as well as the broadness of their tuning contribute to the average activity level in a voxel: A particular mean voxel response could be obtained by a few neurons in that voxel which each respond unselectively to many different faces, or by a large number of highly selective neurons which each respond only to a few faces, yet these two scenarios have very different implications for behavioral discrimination ability (Jiang et al., 2006). The average stimulus-driven BOLD-contrast response is therefore insufficient to estimate neuronal tuning specificity and behavioral performance, as shown by a number of recent studies (Bölte et al., 2006; DeGutis et al., 2007; Mahon et al., 2007; Riesenhuber, 2007).

In contrast, fMRI adaptation techniques have been shown to be able to more directly probe neuronal selectivity than conventional methods relying on average BOLD-contrast stimulus responses (Grill-Spector et al., 2006). The fMRI rapid adaptation technique (fMRI-RA) is motivated by findings from monkey electrophysiology experiments in ventral temporal cortex, reporting that the second presentation of a stimulus (within a short time period) evokes a smaller neural response than the first (Miller et al., 1993). It has been shown that this adaptation can be measured using fMRI, and that the degree of adaptation depends on stimulus similarity, with repetitions of the same stimulus causing the greatest suppression. Several studies (Fang et al., 2007; Gilaie-Dotan and Malach, 2007; Jiang et al., 2006, 2007; Murray and Wojciulik, 2004) have provided evidence that parametric variations in visual object parameters (shape, orientation, or viewpoint) are reflected in systematic modulations of the fMRI-RA signal, and can be used as an indirect measure of neural population tuning (Grill-Spector et al., 2006). Crucially, we (Jiang et al., 2006) and others (Gilaie-Dotan and Malach, 2007; Goh et al., 2010) have provided quantitative experimental evidence that behavioral face discrimination performance and neural tuning selectivity in the FFA as estimated with fMRI-RA are tightly linked, as predicted by a computational model of human face processing (Jiang et al., 2006).

In the present study, we tested the hypothesis that, as in neurotypical individuals, the FFA is the central region for face processing in ASD, and variations in face discrimination performance in ASD are related to selectivity differences of face neurons in the FFA. Specifically, within the ASD population, higher face discrimination performance is predicted to be associated with more selective neural tuning in the FFA, whereas lower performance is associated with broader tuning (Goh et al., 2010; Jiang et al., 2006). We tested this hypothesis using two independent methods: a novel analysis of voxel-wise correlations to probe sparseness of neural activations, and fMRI rapid adaptation, an established technique that is considered to be a more direct probe of neuronal selectivity than conventional methods relying on average BOLD-contrast stimulus responses (Grill-Spector et al., 2006).

2. Materials and methods

2.1. Participants

Twenty-seven subjects (age range 19–58, ten female) diagnosed with an ASD (see Table 1) participated in this study. ASD diagnosis was determined using the Autism Diagnostic Observation Schedule (ADOS), Module 4. Twelve of the subjects were recruited at Massachusetts General Hospital (MGH) and only participated in the behavioral experiments, and the other fifteen were recruited at Georgetown University Medical Center and the Center for Autism Spectrum Disorders at Children's National Medical Center, in Washington, DC, and participated in the behavioral and fMRI Experiments. These participants with autism also completed the Autism-Spectrum Quotient (AQ) assessment instrument, and scored 32 or above, a cutoff found to reliably identify adult individuals with autism, with a low false alarm rate (Baron-Cohen et al., 2001). Fifteen typical adults recruited from the local community served as behavioral controls (age range 19–59) and only participated in the face discrimination experiment described in Section 2.3. The controls were age- and gender-matched to the fifteen ASD subjects who participated in the fMRI experiments. In our ASD sample, we did not find a significant correlation of face discrimination ability (performance in the M6 condition, see below) and age ($r = -0.15$, $p > 0.4$), nor in the age- and gender-matched controls ($r = 0.01$, $p > 0.95$). Experimental procedures were approved by Georgetown University's Institutional Review Board and the Partners Human Research Committee at Massachusetts General Hospital, and written informed consent was obtained from all subjects prior to the experiment. Subjects participating in the fMRI portion of the study were familiarized with the scanning environment in a mock scanner as part of their first study visit, and trained to minimize head motion. On the second visit, subjects underwent fMRI scanning followed by the out-of-scanner face discrimination testing.

Table 1

Age, gender, IQ, and diagnoses of the 27 participants with autism. Participants listed in bold font are those who participated in the fMRI experiment (see Materials and methods). ASD (autism spectrum disorder), HFA (high-functioning autism), RHL (right-hemisphere learning disability), SEPD (social-emotional processing disorder).

Age	Gender	WASI voc.	WASI perf.	WASI full	ADOS
45	M	118	129	127	HFA
52	F	118	104	112	ASD
35	F	105	104	105	ASD
24	M	112	138	128	ASD
54	F	132	126	133	HFA
22	M	118	124	124	HFA
19	F	119	121	123	ASD
27	F	109	118	115	ASD
59	F	116	112	116	ASD
23	M	116	119	120	ASD
20	M	102	112	108	HFA
28	M	93	104	99	ASD
30	M	62	115	85	HFA
25	M	119	101	111	ASD
58	M	134	134	139	ASD
23	M	117	119	120	SEPD + ASD
40	F	105	86	114	RHL + ASD
44	F	128	95	119	RHL
21	M	115	69	112	RHL + ASD
47	M	107	90	112	RHL
36	M	131	128	133	SEPD + ASD
23	F	139	125	136	SEPD + ASD
48	M	132	109	124	RHL
25	M	112	111	114	ASD
23	M	141	127	140	SEPD + ASD
42	F	131	117	127	ASD
48	M	119	106	114	ASD

2.2. Famous faces recognition test

Face pictures of fifty celebrities and fifty ordinary people (unknown to the subjects) were presented on the screen in random order, and subjects indicated whether they knew the face on the screen or not by pressing one of two buttons (Barton et al., 2004). Pictures were kept visible until subjects pressed a button, which also served to initiate the next trial. After subjects finished the test, a questionnaire with the names of all fifty celebrities was presented to subjects, and subjects needed to indicate whether they knew each celebrity. Based on responses to the questionnaire, responses from four subjects were excluded from analysis due to insufficient familiarity with the famous faces used (>5 unfamiliar). To avoid response bias, d-prime (d') was used to quantify subjects' ability to identify famous faces.

2.3. Face discrimination test

Using face stimuli generated by a photorealistic face morphing system (Blanz and Vetter, 1999; Jiang et al., 2006) along twenty-five within-gender morph lines based on fifty individual prototype faces (200 by 256 pixels, twenty-six females, see Fig. 1A), we tested subjects' face discrimination abilities using a two-alternative forced choice (2AFC) paradigm (Fig. 1B). For each trial, after a 500 ms fixation, the target face was presented for 200 ms, followed by a mask image for 400 ms, followed by two choice faces presented side-by-side for 4000 ms or when subjects responded (whichever came first), and subjects were asked to judge which one of the two faces was the same as the target face. The next trial would automatically start 1000 ms after subjects made a response. If subjects failed to respond within 4000 ms, an auditory alarm ("beep") would be presented, and the next trial would start 1000 ms after the beep.

One of the choice faces was always the same as the sample face, while the other choice face differed from the first one by one of four possible different levels of similarity. This was done by creating "morph lines" interpolated between two prototype faces, and then choosing face images separated by a specified distance along this continuum (Fig. 1A). We tested four different levels of intra-pair similarity, with shape differences of 30%, 60%, 90%, and 120% (conditions M3/6/9/12, respectively, see Fig. 1A), with 100% corresponding to the distance between two prototype faces, and 120% difference created by extrapolation (Blanz and Vetter, 1999). Within each trial, the three faces were either all upright or all inverted (data from inverted faces will be

reported elsewhere). Stimuli were presented to participants on an LCD monitor on a dark background, 1024 × 768 resolution, 60 Hz refresh rate, at a distance of 60 cm. An in-house software package was used to present the stimuli and to record the responses. Participants completed a total of 800 trials (80 per condition) in sixteen blocks.

Our previous study (Jiang et al., 2006) showed that, for typical adults, an asymptote is reached in the M6 condition, in both behavioral performance and releases from fMRI adaptation, i.e., a significant increase from M3 to M6 condition in discriminating faces, but not from the M6 to M9 condition (Jiang et al., 2006). In the present study, we therefore used performance in the M6 condition as a measure of subjects' face discrimination ability to correlate with the fMRI data; nearly identical results were obtained when we fitted subjects' performance to a sigmoid function, and then used the fitted parameters as the indicators of behavioral performance (see Fig. S5 in Jiang et al. (2006)).

2.4. Functional localizer scans

To locate the FFA regions, a block design was used to collect MRI images from two localizer scans for each subject (Haxby et al., 1999; Jiang et al., 2006; Kanwisher et al., 1997). During each run, following an initial 10.2 s fixation period, 50 grayscale images of faces, houses, and scrambled faces were presented to participants in blocks of 30.6 s (each image was displayed for 512 ms and followed by a 100 ms blank screen), and were separated by a 20.4 s fixation block (Fig. 2). Each block was repeated twice in each run, which lasted for 316.2 s, and participants were asked to passively view these images while keeping their fixation at the center of the screen. The face and house images used in the localizer scans were purchased from <http://www.hemera.com> and post-processed using programs written in MATLAB (The Mathworks, MA) to eliminate background variations, and to adjust image size, luminance, and contrast. The final size of all images was scaled to 200 by 200 pixels, and half of the faces were scrambled using a grid of 20 by 20 pixel elements while the outlines of the faces were kept intact. The fMRI data of functional localizer scans from one subject were discarded due to lack of activation across visual cortex (for this subject, the FFA region of interest (ROI) was defined by stimuli > baseline from the event-related scans, see below), and only one run of images was collected for two other subjects. The data from the localizer scans were also used in the local regional heterogeneity analysis (see below) to probe the sparsity of FFA activations.

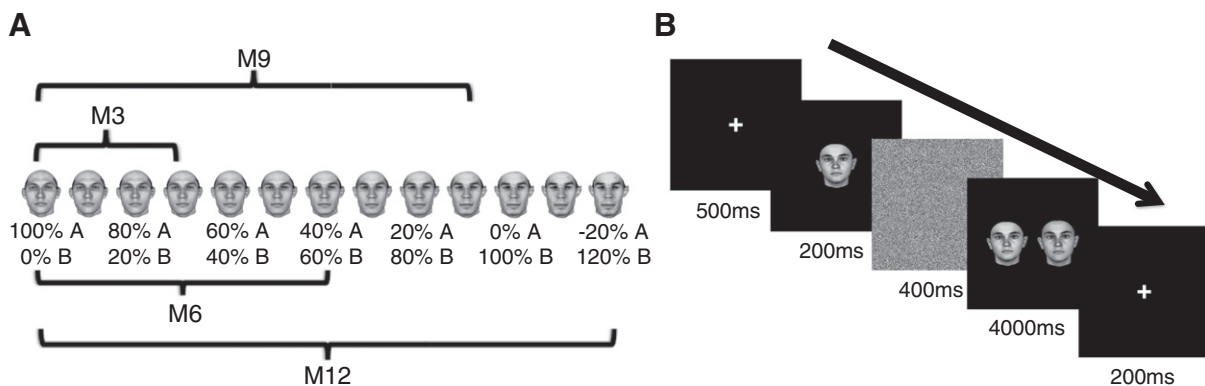


Fig. 1. Face stimuli and experimental design of the psychophysics experiment. (A) Shows an example "morph line" between a pair of face prototypes (shown at the far left and second from the right), created using the photorealistic morphing system (Blanz and Vetter, 1999). The relative contribution of each face prototype changes smoothly along the line. For example, the fourth face from the left (the third morph) is a mixture of 70% of the face on the far left and 30% of the face on the far right. Four conditions, M3/6/9/12, which correspond to shape differences of 30%, 60%, 90%, and 120%, were examined in the present study. The difference between two prototype faces was defined as 100%; 120% difference was achieved by extrapolation along the morph line beyond face prototypes (Blanz and Vetter, 1999). (B) Design of the 2AFC face discrimination experiment. Subjects viewed a target face, followed by a mask, followed by two test faces, presented side-by-side, and had to indicate which of the test faces was identical to the previously presented target face. There were four levels of shape difference, M3/6/9/12, between the two test faces as explained in (A).

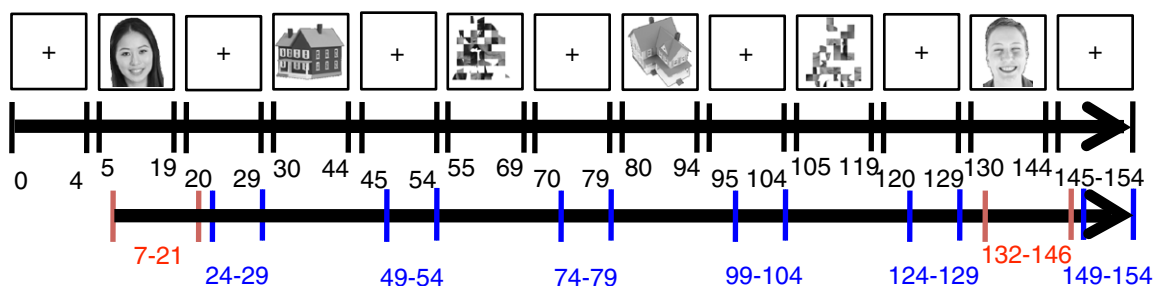


Fig. 2. Functional localizer scans. Participants were asked to passively view blocks of face, house, and scrambled face images that were separated by fixation blocks. The numbers (in black font) indicate the first and last MRI acquisition of each block. MRI images that were chosen for the local regional heterogeneity analysis are shown in red (face blocks) and blue (fixation blocks). Due to the lag of hemodynamic responses, for face blocks, the images from $n0 + 2$ to $nm + 2$ were used ($n0$ is the time point when face blocks started, and nm is the time point when face blocks ended), and the images from $n0 + 4$ to nm were used for fixation blocks to avoid signal overlap from the blocks before and after it ($n0$ is the time point when a fixation block started, and nm is the time point when it ended).

2.5. Rapid event-related (ER) scans

MRI images from four ER scans were collected for each subject. Each run lasted 538.56 s and had two 10.2 s fixation periods, one at the beginning and the other at the end. Between the two fixation periods, a total of 127 trials were presented to participants at a rate of one every 4.08 s. During each trial (except null trials), two faces were displayed sequentially (300 ms each with a 400 ms blank screen in-between), and followed by a 3080ms blank screen (Jiang et al., 2006). For each run, the data from the first two trials were discarded, and analyses were performed on the data of the other 125 trials – 25 each of the five different conditions: three conditions of interest of varying intra-pair stimulus similarity (M3/M6/M9, see Fig. 1A above) (Jiang et al., 2006), task trials, in which an ‘oddball’ target face, which participants needed to identify, could appear as either the first or the second one of the pair of faces, and null trials (Jiang et al., 2006). Performance of subjects in the ASD group on the oddball task inside the scanner was nearly perfect and did not differ from controls (Jiang et al., 2006). Trial order was randomized and counterbalanced using M-sequences (Buracas and Boynton, 2002). While inside the scanner, participants were asked to watch all the faces but only respond to instances of the target face by pressing a button with the right hand. Morphed faces (200 by 200 pixels) along ten within-gender morph lines of twenty individual prototype faces (ten females) were used, along with one additional oddball target face, different from the face prototypes used to generate the morphed stimuli. The stimuli of both localizer and ER scans were presented on black background using E-Prime software (<http://www.psnet.com/products/e-prime/>), back-projected on a translucent screen located at the rear of the scanner, and viewed by participants through a mirror mounted on the head coil. The fMRI data from the ER scans from one subject were discarded due to excessive head movements, and only two runs of fMRI data were collected from two other subjects. One subject needed to be rescanned on the ER paradigm as her data from the first run indicated that she had erroneously performed a face discrimination task on the face pairs. Subjects were familiarized with the scanning environment and experimental task in a mock scanner prior to scanning, and one or two practice blocks of trials were administered to subjects to confirm that they were able to do the oddball task inside the scanner.

2.6. MRI data acquisition and analysis

MRI data were acquired at Georgetown University's Center for Functional and Molecular Imaging using an echo-planar imaging (EPI) sequence on a 3T Siemens Trio scanner (Flip angle = 90° , TR = 2.04 s, TE = 29 ms, FOV = 205, 64×64 matrix) with an eight-channel head coil ($n = 5$) or a twelve-channel head coil ($n = 10$) (note that since contrast measures used in the analyses are all within-subject, these

differences in coils do not present a problem in the current study; in addition, internal quality tests showed that the two head-coils are comparable in terms of signal-to-noise ratio). Thirty-five interleaved axial slices (thickness = 4.0 mm, no gap; in-plane resolution = $3.2 \times 3.2 \text{ mm}^2$) were acquired for the two functional localizer and four functional runs. At the end, three-dimensional T1-weighted MPRAGE images (resolution $1 \times 1 \times 1 \text{ mm}^3$) were acquired from each subject.

After discarding the images acquired during the first five acquisitions of each run, the EPI images were temporally corrected to the middle slice, spatially realigned and unwrapped together with the images from the localizer scans using the SPM2 software package (<http://www.fil.ion.ucl.ac.uk/spm/software/spm2/>), then all images were resliced to $2 \times 2 \times 2 \text{ mm}^3$, normalized to a standard MNI reference brain in Talairach space, and smoothed with 6mm Gaussian kernel using SPM2.

The FFA regions were identified for each individual subject independently with the data from the localizer scans (except for one subject, who did not show any activation throughout visual cortex during the localizer scans even though activation in the ER scans was normal, and whose FFA was defined by the contrast of stimuli versus baseline in the ER scans). We first modeled the hemodynamic activity for each condition (face, scrambled face, and house) in the localizer scans with the standard canonical hemodynamic response function (Friston et al., 1995), then identified the FFA ROI with the contrast of face versus house masked by the contrast of face versus baseline ($p < 0.0001$). To obtain comparably-sized FFAs across subjects, we defined the FFA as the ROI consisting of the 45 contiguous voxels with the highest statistical threshold for each subject (Jiang et al., 2006). The right occipital face area (OFA) (Haxby et al., 2000) was defined in the same way in thirteen subjects (45 voxels, $n = 13$, as the OFA could not be reliably identified in two subjects), as was the left FFA (45 voxels, $n = 14$). We also identified a V1/2 ROI (a cuboid ROI centered at the local maximal voxel (defined by the contrast of face versus baseline) in the right V1/2 region, size: 48 voxels, shape: $4 \times 4 \times 3$ voxels, $n = 15$), and the right parahippocampal place area (PPA) ROI (Epstein and Kanwisher, 1998) in each subject using the contrast of house versus face masked by house versus baseline ($p < 0.00001$), of same size (45 contiguous voxels, $n = 14$ as no activations in visual cortex during the localizers was found in one subject, see above). We were able to identify a right STS ROI (Haxby et al., 2000) in nine subjects using the contrast of face versus house masked by face versus baseline ($p < 0.001$, uncorrected), but activity in this area did not correlate significantly with face discrimination ability on any of the measures used (FS_L , FS_{RA} , FS_{369} , and H_{corr} ; see below) (data not shown). To test whether results were affected by which localizer scan was used to define the FFA ROI, we also identified another set of independent right FFA using the data from the second localizer scan, and then analyzed the data from the first localizer scan. Results obtained from these FFA

were similar to those obtained with FFAs defined from both localizer scans.

After removing low frequency temporal noise from the EPI runs with a high pass filter (1/128 Hz), fMRI responses were modeled with a design matrix comprising the onset of each non-null trial and movement parameters as regressors using SPM2, and proportional scaling was applied to remove the effects of global variations (Aguirre et al., 1998), as we expected differences between the three adaptation conditions (M3/6/9) to be small and limited to local face-selective regions based on previous findings (Gilaie-Dotan and Malach, 2007; Jiang et al., 2006). We then extracted the hemodynamic response for each subject in the right and left FFA and OFA, using a standard canonical hemodynamic response function (Friston et al., 1995) with the MarsBar toolbox (Brett et al., 2002) and in-house software written in Matlab.

2.7. Local regional heterogeneity analysis

In the present study, we introduce a novel technique to probe the sparseness of neural activation patterns and provide an indication of neural selectivity. The technique, called *local regional heterogeneity analysis*, is based on an analysis of voxel-wise correlations of fMRI activation patterns. It is motivated by our earlier computational, behavioral and imaging studies (Jiang et al., 2006; Riesenhuber and Wolff, 2009) that provided quantitative evidence that human face perception is based on the discrimination of sparse activation patterns over highly selective face neurons. In this account, the sparseness of activation patterns in the FFA is related to face neuron selectivity: The high selectivity of face neurons in typical subjects produces a sparse neural code, as each face neuron only responds to a small set of faces highly similar to its preferred face. In contrast, less selective face neurons respond to a greater number of dissimilar face stimuli, leading to greater overlap in responses among face neurons and less sparse representations. Indeed, training studies in monkeys have shown that learning produces sparser codes, with neurons after training responding to fewer stimuli (Freedman et al., 2006; Kobatake et al., 1998), and we have found compatible results using fMRI-RA in humans, indicating training-induced sharpening of neural tuning, which were paralleled by improvements in behavioral discrimination abilities (Jiang et al., 2007). For a selective face representation based on sharply tuned face neurons, activation patterns should thus be sparser than for a population with more broadly tuned neurons. This then predicts that the sharply tuned population should have a higher degree of heterogeneity in neural correlations (as small groups of neurons preferring similar faces should show high correlation but low correlation with other neurons) than the population with more broadly tuned neurons (in which more neurons respond in a similar way, increasing the uniformity of correlations). This hypothesis is supported by single-unit studies that have found that neurons with similar tuning tend to show more correlated firing than neurons with dissimilar tuning (Bair et al., 2001; Jermakowicz et al., 2009). Thus, when measured with fMRI, tuning specificity in the FFA should correlate with the degree of sparseness, which in turn should correlate with the heterogeneity of correlations in the FFA, i.e., the standard deviation or standard error of correlations between voxels. In other words, across subjects, a lower local regional heterogeneity of correlations in the FFA should be associated with lower behavioral performance in discriminating faces, and greater heterogeneity with better behavioral performance.

For the local regional heterogeneity analysis, we first extracted the raw time series data in the FFA from the first localizer scan (see Fig. 2) for each subject (with normalization but without additional smoothing), followed by removal of the mean, any linear trends, and low frequency variations (Friston et al., 1995). The fMRI data from every time point in the face blocks and fixation blocks (after compensating for delays caused by the slow hemodynamic response) were

used in a pair-wise correlation analysis between each voxel, which resulted in a set of pairwise correlation coefficients (for n voxels), r_{ij} ,

$$r_{ij} = \text{corr}(Vox_i, Vox_j), \quad i, j \in 1..n \quad (1)$$

We then calculated a measure of local heterogeneity, H_{corr} , as the standard error of the mean (SEM) of those correlation coefficients (r_{ij} , $i < j$, because $r_{ij} = r_{ji}$, and $r_{ii} = 1$).

$$H_{corr} = \sqrt{\frac{\sum_{i=1}^{n-1} \sum_{j=i+1}^n (r_{ij} - u)^2}{N \times (N-1)}}, \quad \text{where } N = \sum_{i=1}^{n-1} i, u = \frac{1}{N} \sum_{i=1}^{n-1} \sum_{j=i+1}^n r_{ij} \quad (2)$$

H_{corr} was then used for a correlation analysis with subjects' behavioral performance in the M6 condition. We also calculated H_{corr} in the right OFA, PPA, and V1/2 ROIs of each individual subject to probe whether a correlation between behavioral performance and H_{corr} is limited to FFA (as predicted by the model), or is a general global effect. For comparison, we also ran the local regional heterogeneity analysis on the data from neurotypical subjects who were part of our previous study (Jiang et al., 2006) (FFA ROI size, range 42–49, mean 45.0 ± 0.9). Finally, to investigate whether spatial resampling might affect the regional heterogeneity measurements, we re-analyzed the data using different voxel sizes ($1.5 \times 1.5 \times 1.5\text{mm}^3$, yielding an FFA of 100 voxels, and $3 \times 3 \times 3\text{mm}^3$, yielding an FFA of 15 voxels).

2.8. Face selectivity indices

In addition to H_{corr} , we defined two face selectivity indices (one for the functional localizer scans, and the other for the fMRI-RA scans) to investigate the relationship between fMRI response and behavioral performance.

For the data from the functional localizer scans, we extracted the fMRI responses to faces and houses in the right FFA from localizer scans and defined a *Face Selectivity Index*, FS_L , (Kourtzi et al., 2003) as

$$FS_L = \frac{\text{Face} - \text{House}}{\text{Face} + \text{House}} \quad (3)$$

For the data from the fMRI-RA scans, we defined a *Face Selectivity index*, FS_{RA} , as

$$FS_{RA} = \frac{(M6 - M3) - (M9 - M6)}{(M3 + M6 + M9)/3} \quad (4)$$

FS_{RA} is high for signal asymptotes at M6 (indicating high specificity as in typical controls), i.e., for an increase in release from fMRI adaptation from M3 to M6, but not from M6 to M9, as found in typical controls (Jiang et al., 2006). Conversely, FS_{RA} is low if release from adaptation increases beyond M6 and extends to M9 (indicating less specific face neuron tuning).

Furthermore, to provide an additional index to show the link between responses in the fMRI-RA Experiment and behavioral face discrimination performance, we defined a new face selectivity index, FS_{369} , as,

$$FS_{369} = \frac{M6}{M3 + M9} \quad (5)$$

A high FS_{369} would point to a strong release from adaptation at M6, which would in turn suggest high face selectivity. Conversely, a low FS_{369} suggests weak release from adaptation at M6, which would in turn suggest low face selectivity.

2.9. Comparison to Jiang et al. (2006)

To link the results of the present study to those of our earlier study (Jiang et al., 2006) and further validate the local regional heterogeneity analysis method, we re-analyzed the data from (Jiang et al., 2006) using the same methods as in the current study (with FFA ROI based on the 45 most significant contiguous voxels, see Materials and methods). The designs of fMRI Experiment 1 in (Jiang et al., 2006) and the fMRI Experiment in the present study are identical except for minor changes in trial duration (4 s in (Jiang et al., 2006), 4.08 s in the present study) and MRI acquisition (TR = 2 s, TE = 30 ms, 44 slices, thickness = 3.2 mm, and a single-channel head coil were used in Jiang et al. (2006), and TR = 2.04 s, TE = 29 ms, 35 slices, thickness = 4.0 mm, and an eight- or twelve-channel head coil were used in the present study). In addition, the design of the 2AFC face discrimination experiment was slightly different between the present study and Jiang et al. (2006), with choice faces presented sequentially in Jiang et al. (2006) and side-by-side in the present study, but this difference did not affect face discrimination performance, ($p > 0.35$, two-tailed t-test, M6 condition).

3. Results

3.1. Heterogeneity of face discrimination abilities in ASD

Twenty-seven adults diagnosed with an ASD (see Materials and methods) and fifteen typical controls participated in the 2AFC face discrimination behavioral experiment (Fig. 1B). Based on results from our

earlier study (Jiang et al., 2006) that had shown that by the M6 condition (i.e., for face pairs separated by 60% shape difference) typical controls approach an asymptote for both release from adaptation in the fMRI-RA experiment as well as discrimination ability (thick blue line with error bars in Fig. 3A), we used behavioral performance in the M6 condition to define whether a subject in the ASD group had typical or impaired face discrimination abilities, depending on whether his/her performance at the M6 condition was above or below the mean minus one standard deviation of the performance of typical controls at the M6 condition. At the group level, a two-sample t-test revealed a significant difference between the ASD and control groups in the M6 condition ($p < 0.03$, Fig. 3B). In line with a previous report (Barton et al., 2004), a high heterogeneity in face discrimination performance was observed in the ASD group – about half of the subjects in that group ($n = 13$, green lines in Fig. 3A) showed performance levels approximately comparable to typical controls, while the other half ($n = 14$, red lines in Fig. 3A) exhibited more substantial impairments in discrimination performance. However, note that this separation of the ASD cohort into groups with and without face processing deficits is only meant to illustrate the substantial heterogeneity of face processing abilities in ASD – the data themselves indicate a continuum of performance levels, ranging from severe impairment to typical performance levels, rather than the existence of discrete subgroups. To directly compare the heterogeneity of behavioral performance between the ASD and control groups, we conducted an F-test on performance in the M6 condition, and found that the variance in ASD group was significantly higher than control group ($p < 0.015$), further supporting that there is a high heterogeneity in face discrimination in individuals with ASD.

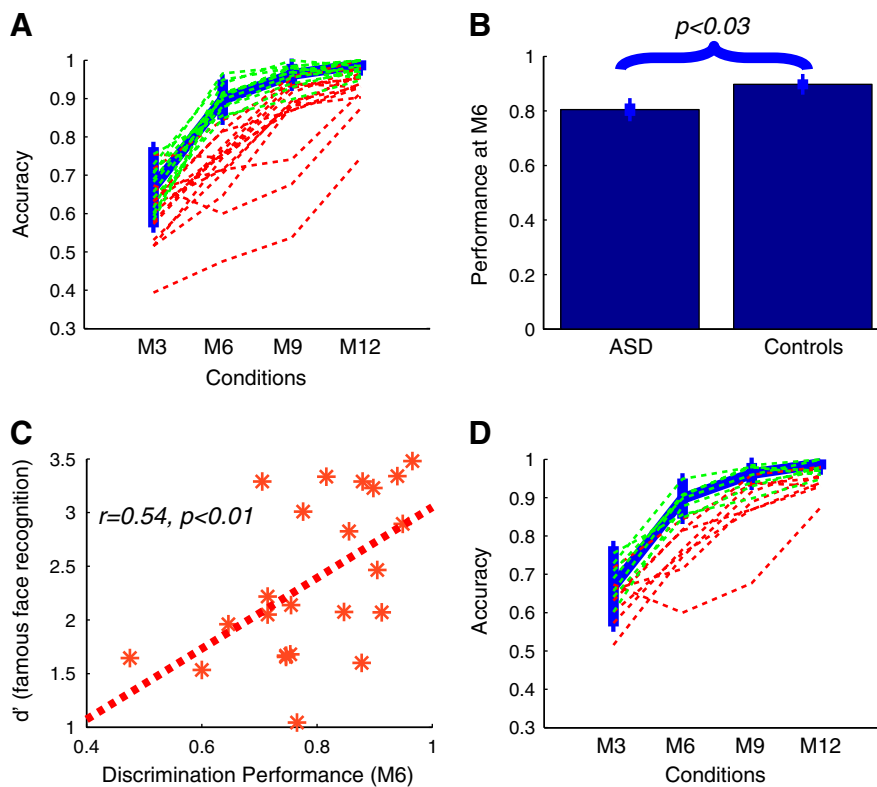


Fig. 3. Heterogeneity of face discrimination abilities in autism and correlation with famous face recognition performance. (A) Individual behavioral face discrimination performance of twenty-seven participants with ASD (thin green and red lines), and the mean and standard deviation (SD) of fifteen typical controls (thick blue lines). Green and red lines show subjects in the ASD cohort with performance within and below one standard deviation from the mean performance of control subjects, respectively (see text). (B) Mean accuracy from nine typical controls and twenty-seven ASD subjects. (C) Correlation between face discrimination ability in the M6 condition vs. performance (d') in the famous face recognition experiment for participants in the ASD group ($N = 23$, as four participants with ASD were not sufficiently familiar with the famous individuals used in the test, see Materials and methods). (D) Individual behavioral face discrimination performance of the fifteen subjects with ASD who also participated in the fMRI experiments (thin lines), and the mean and standard deviation (SD) of the controls (thick blue lines). Green and red lines show normal and below-normal face discrimination performance as in (A). Error bars indicate SD.

Subjects in the ASD group also participated in the famous face recognition experiment (Barton et al., 2004) (see Materials and methods). Results showed a significant correlation between accuracy in discriminating faces in the M6 condition and recognition performance in the famous faces test ($r = 0.54$, $p < 0.008$), suggesting that the ability to discriminate the fine shape differences in the computer-generated morphed faces was related to the ability to recognize real faces (Fig. 3C). This parallels a previous report showing that ASD subjects with better famous face recognition have better ability to discriminate changes in facial structure (Barton et al., 2004). In line with previous reports (Barton et al., 2004), we did not observe significant correlations between face discrimination performance in the M6 condition and the demographic and neuropsychological measurements, including age ($p > 0.46$), IQ ($p > 0.98$), ADOS social score ($p > 0.82$), and ADOS communication scores ($p > 0.83$). The data from the fifteen subjects who also participated in the fMRI component of the study are shown in Fig. 3D.

3.2. Face-selective response amplitude in the FFA does not predict behavioral face discrimination ability

After identifying the right FFA using the contrast of faces versus houses masked by faces versus baseline in each individual subject (Grill-Spector et al., 2004; Jiang et al., 2006) (see Fig. 4 for the FFA in one representative subject), we extracted the fMRI responses to faces and houses in the right FFA from localizer scans and calculated the FS_L for fourteen subjects. Pearson's correlation analysis showed no correlation between FS_L and behavioral performance in the M6 condition ($r = 0.30$, $p > 0.30$, see Fig. 5A). Similar results were obtained when correlating behavioral performance in the M6 condition with the fMRI response to faces ($r = 0.07$, $p > 0.82$), or with the differences

in fMRI responses to faces and to houses ($r = 0.17$, $p > 0.56$), or with the ratio of fMRI responses to faces divided by fMRI responses to houses ($r = 0.27$, $p > 0.35$) (Fig. 5B–D).

Together with the conflicting reports on FFA activation levels from previous fMRI studies of face processing study in ASD (reduced or no FFA activity to faces in Humphreys et al. (2008), Grelotti et al. (2005), Pierce et al. (2001), and Schultz et al. (2000) versus typical activation levels at FFA in Hadjikhani et al. (2007), Perlman et al. (2011), and Pierce et al. (2004)), these results provide further support that the commonly used contrast of faces versus houses or other classes of objects in fMRI studies is insufficient to predict behavioral abilities, likely due to the aforementioned confound that measures of average activity in a voxel can be affected by the density of selective neurons as well as the broadness of their tuning (Grill-Spector et al., 2006; Jiang et al., 2006; Riesenhuber, 2007).

3.3. Local heterogeneity of voxel-wise correlations in the FFA provides a direct link between neural selectivity (sparseness of neural activations) and individual behavioral performance

We next tested the hypothesis that the local heterogeneity of voxel-wise correlations, H_{corr} , in the FFA during the localizer scans could predict behavioral performance in the ASD cohort. Using the data from face blocks only, Pearson's correlation analysis revealed a trend of correlation between H_{corr} and behavioral performance ($r = 0.48$, $p < 0.085$, see Fig. 6A). In order to increase the power of our analysis, we next augmented the data set used to calculate pair-wise correlations of activation by including the data from both face blocks and resting blocks (see Fig. 2), following reports from a number of single unit recording studies (Bair et al., 2001; Jermakowicz et al., 2009) that found that pair-wise correlation

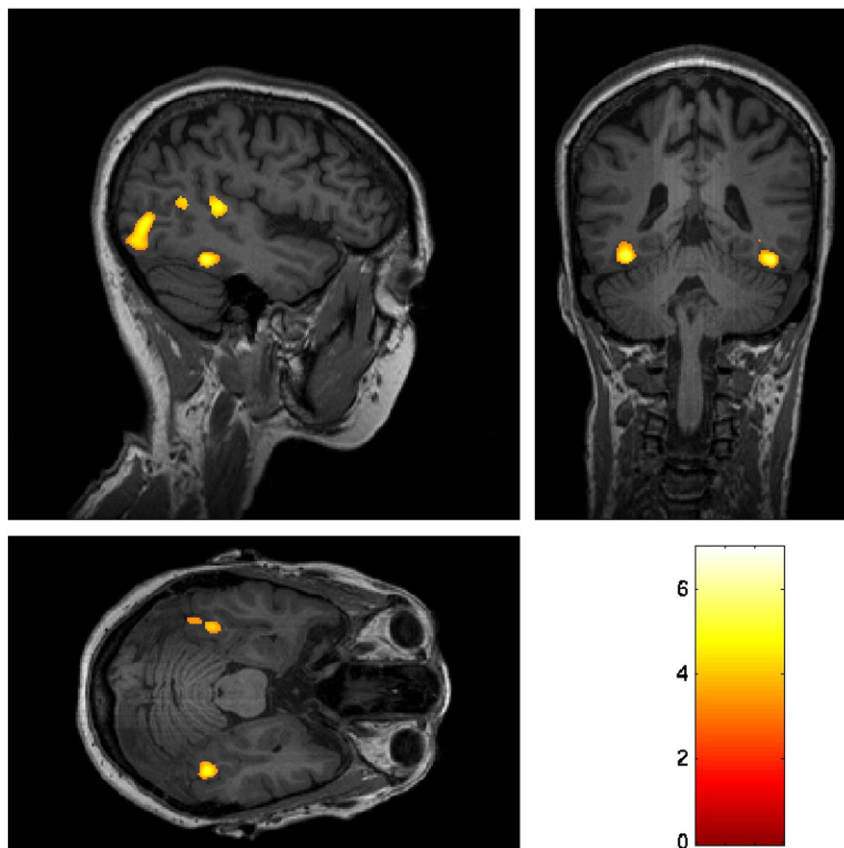


Fig. 4. Left and right FFA and right OFA from one representative subject. For illustration purpose, a less strict threshold ($p < 0.001$, uncorrected, masked by face > baseline, $p < 0.00001$, uncorrected) was used.

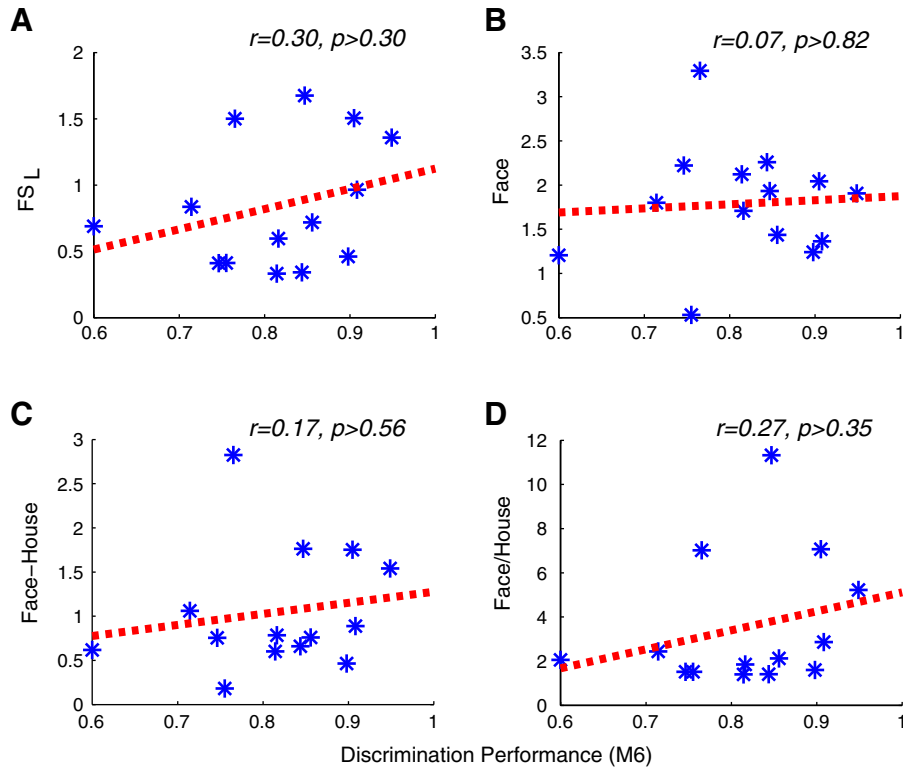


Fig. 5. fMRI response amplitude in the FFA cannot predict face discrimination ability. No significant correlation was found between the behavioral face discrimination performance in the M6 condition and (A) face selectivity FSL defined by fMRI responses in the FFA to faces versus houses ($n = 14$), (B) amplitude of fMRI response to faces ($n = 14$), (C) differences of response amplitude to faces and houses ($n = 14$), and (D) the ratio of fMRI responses to faces versus houses. Dashed lines show the regression line.

between two neurons in the presence of visual stimuli (which in turn are related to similarities in neuronal tuning (Bair et al., 2001; Jermakowicz et al., 2009)) can also be detected in spontaneous activity (i.e., in the absence of stimuli). Indeed, analyzing the joint data set revealed a highly significant correlation between H_{corr} and behavioral performance ($r = 0.75, p < 0.0022$, see Fig. 6B). Using the data from

both face and resting blocks, a trend was observed in the left FFA ($r = 0.55, p < 0.054$, Fig. 6C). In contrast, there was no correlation of behavioral face discrimination performance and H_{corr} in the more posterior OFA ($r = 0.19, p > 0.54$, see Fig. 6D; note that correlations in the right FFA for the same subgroup of subjects were comparable to the full cohort, $r = 0.73, p < 0.0074$), lower visual areas (V1/2,

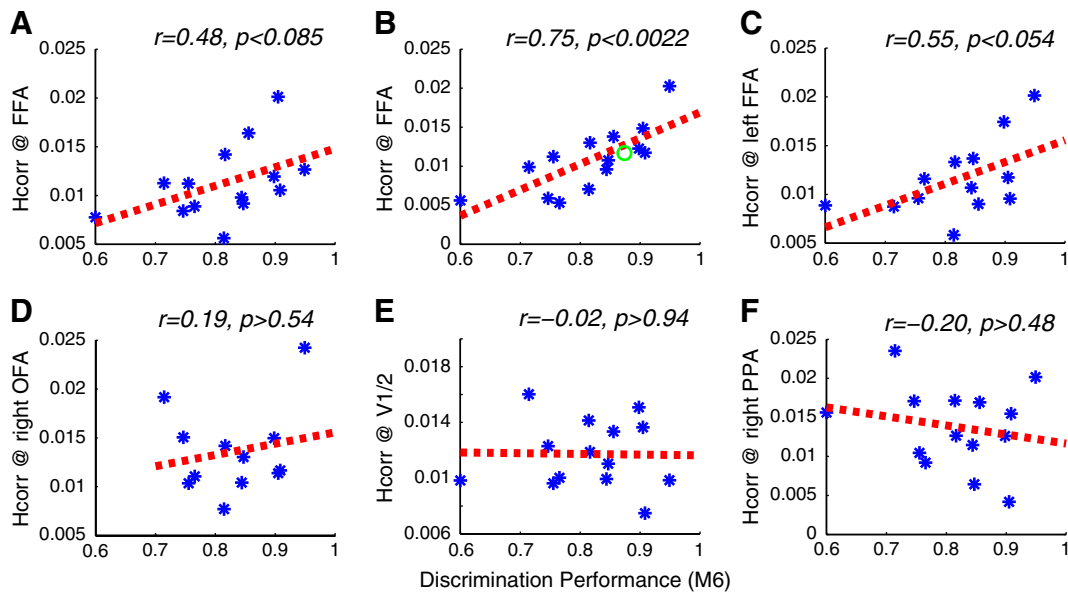


Fig. 6. Local regional heterogeneity, H_{corr} , the standard error of voxel-wise correlations within the FFA, predicts individual face discrimination ability. The plots show the correlation of behavioral face discrimination performance in the M6 condition and local regional heterogeneity H_{corr} in the FFA (A) based on face blocks only, and (B) with face and fixation blocks, and a trend was observed in the left FFA (C), but not in the right OFA (D), lower visual areas (V1/2, right hemisphere) (E), nor the right PPA (F). Green circle in (B) shows data from the typical subjects in (Jiang et al., 2006). Dashed lines show the respective regression lines.

$r = -0.02$, $p > 0.94$, see Fig. 6E), nor in the right PPA ($r = -0.20$, $p > 0.48$, see Fig. 6F), showing that the correlation of behavior and voxel-wise heterogeneity is specific to the FFA, as predicted, and not a general effect. Furthermore, the lack of correlation between behavioral performance and H_{corr} in the more posterior right OFA suggests that the FFA, rather than the OFA, is still the key region mediating face perception in individuals with ASD. Similar results were obtained when using the second localizer scan to identify the right FFA and the first localizer scan for analysis. In contrast, the average voxel-wise correlation within the FFA was not significantly correlated with behavioral performance ($r = -0.43$, $p > 0.12$). Furthermore, we probed the impact of head movement on the local regional heterogeneity analysis. There was no correlation between H_{corr} in the FFA and any of the six head movement parameters obtained during preprocessing (at least $p > 0.31$), and regressing out the six head movement parameters did not change the conclusions ($r = 0.75$, $p < 0.03$).

To further validate the H_{corr} measure, we calculated the H_{corr} values of the typical adult subjects in our earlier study (Jiang et al., 2006), which fit very well with the data from the autistic population (green circle in Fig. 6B), demonstrating the strength of the H_{corr} analysis to estimate neural selectivity and predict behavioral ability, even across studies. Finally, similar correlations between H_{corr} in the FFA and behavioral performance in the M6 condition were observed when the fMRI data were resampled at $1.5 \times 1.5 \times 1.5 \text{ mm}^3$ ($r = 0.707$, $p < 0.005$), or $3 \times 3 \times 3 \text{ mm}^3$ ($r = 0.603$, $p < 0.023$), indicating that the results are not an artifact of voxel resampling.

3.4. fMRI-RA provides another direct link between neural selectivity and individual behavioral performance

As stated above, previous studies (Gilaie-Dotan and Malach, 2007; Goh et al., 2010; Jiang et al., 2006) have established a tight link

between neural selectivity as estimated with fMRI-RA and behavioral face discrimination performance in typical adults. This link is based on a computational model of face processing (Jiang et al., 2006) that predicts that viewing a particular face should be associated with a sparse activation pattern over face neurons tuned to faces similar to the currently viewed face, with little activation of neurons tuned to dissimilar faces. Thus, in an fMRI-RA paradigm that varies the similarity between two face images shown successively in a single trial, the BOLD-contrast signal in the FFA for increasing within-pair face dissimilarity should progressively increase as the two faces activate increasingly disjoint subpopulations of neurons (causing increasingly lower amounts of neuronal adaptation), up to where the two images activate different subpopulations of neurons, at which point the response level should asymptote and not increase for further increases in face dissimilarity. Correspondingly, at the behavioral level, this model predicts that the ability to discriminate specific faces is directly related to the dissimilarity of the neural activation patterns associated with these faces in the FFA. These predictions were confirmed experimentally by us (Jiang et al., 2006) and others (Gilaie-Dotan and Malach, 2007; Goh et al., 2010). We now tested whether this same link between behavioral ability and neural selectivity would hold in the ASD group.

Between the two functional localizer runs, MRI images from four runs of event-related (ER) scans using an fMRI-RA paradigm were collected ($n = 15$), while subjects performed an oddball target face detection task that was orthogonal to the conditions of interest, i.e., the similarity of the two faces in a pair was irrelevant for the oddball task (Jiang et al., 2006), thereby avoiding differential task difficulty-related modulations of the conditions of interest (Grady et al., 1996; Riesenhuber, 2007; Sunaert et al., 2000). Three conditions of interests, M3/6/9, were tested (see Fig. 1A). Within each trial, two faces of varying dissimilarity (either 30%, 60%, or 90% of shape change, corresponding to

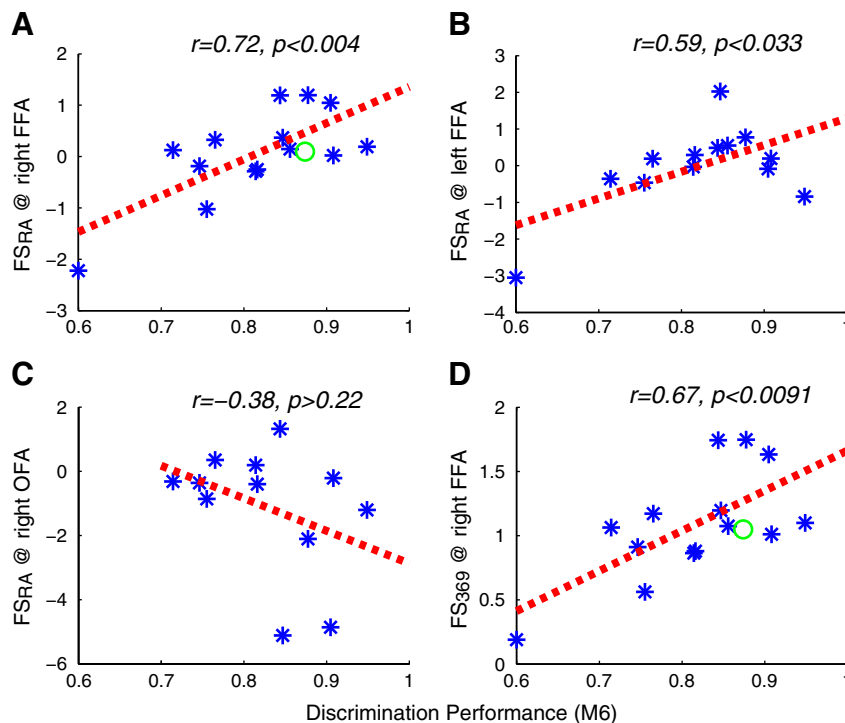


Fig. 7. Neural selectivity as estimated with fMRI-RA correlates with face discrimination ability. (A) Mean fMRI responses in the FFA during the fMRI-RA scans revealed only a weak release from adaptation for small shape differences, suggesting a reduced selectivity of face neurons in the FFA in autism. Correlation of behavioral face discrimination performance in the M6 condition and face selectivity FS_{RA} revealed a significant correlation in the right FFA (B) and left FFA (C), but not in the right OFA and other brain regions (see text, not shown). Another index of face selectivity based on activity in the fMRI-RA experiment, FS_{369} , likewise showed strong correlation with behavioral face discrimination ability (D). Green circles in (B) and (D) indicates the average data from adults with typical face perception (Jiang et al., 2006). Dashed lines show the respective regression lines. Error bars show within-subject SEM.

M3/6/9) were presented in rapid succession, as in our previous studies (Jiang et al., 2006) (Fig. 1A).

The fMRI responses to pairs of faces in the M3/6/9 conditions were extracted from the independently defined right FFA ROI of each individual (same ROI as above). Over the whole ASD group, there was no significant difference between signal levels in the M3 and M6 conditions (Fig. 7A, paired t-test, $p > 0.29$), in contrast to the controls of Jiang et al. (2006), indicating decreased release from adaptation for small shape differences that is compatible with generally reduced selectivity of face neurons in the FFA in autism. We then used responses in the M3/6/9 conditions to calculate the FS_{RA} face selectivity index for each subject (see Materials and methods) to probe whether it predicted behavioral face discrimination ability. Pearson's correlation analysis revealed a significant correlation between FS_{RA} and behavioral performance in the M6 condition ($r = 0.72$, $p < 0.004$, see Fig. 7B), suggesting that the heterogeneity of face perception abilities in individuals with autism might be the consequence of varying degrees of neuronal selectivity to faces in the FFA region, in agreement with the H_{corr} analysis results. Further strengthening the validity of the FS_{RA} measure to relate estimated neural selectivity to behavior, the data from our prior study with typical adults (Jiang et al., 2006) (green circle in Fig. 7B) fit well with the data from this study (note that the current study used a 12-channel coil with a signal-to-noise ratio, SNR, twice as high as that of the previous study (Jiang et al., 2006); the high SNR in the current study permitted the individual-subject correlation analysis, in contrast to the 2006 study, where it was necessary to average over the whole population to reduce variability). In line with the H_{corr} results, a significant correlation was observed in the left FFA ($r = 0.59$, $p < 0.033$, Fig. 7C), but not in the right OFA ($r = -0.38$, $p > 0.22$), nor in lower visual cortex (V1/2) ($r = 0.13$, $p > 0.65$).

Furthermore, Pearson's correlation analysis also revealed a significant correlation between behavioral performance in the M6 condition and another measure of face selectivity, FS_{369} (see Materials and methods), $r = 0.67$, $p < 0.0091$ (Fig. 7D) providing additional evidence of a direct link between neural tuning to faces in the FFA and face discrimination performance in adults with autism. In contrast, as in the functional localizer scans, the amplitude of fMRI responses in the FFA during the event-related scans again did not correlate with behavioral performance ($r = -0.066$, $p > 0.82$).

4. Discussion and conclusions

One of the main challenges in autism research and the development of effective therapies is the identification of differences at the neural level that can be mechanistically and quantitatively linked to the observed behavioral differences that are often highly variable in individuals with autism. In the present study, substantial heterogeneity in face discrimination abilities, linked to real-life face identification skills, was found in adults with autism spectrum disorders. Specifically, despite the significant variability of face discrimination abilities in the ASD cohort, behavioral performance levels were shown to be quantitatively linked to levels of tuning specificity in the right FFA – estimated via a novel analysis of voxel-wise correlations as well as with fMRI-RA, suggesting that the phenotypic heterogeneity in face processing in ASD is mediated by neuronal selectivity to faces in the FFA, similar to findings in typical adults (Gilaie-Dotan and Malach, 2007; Goh et al., 2010; Jiang et al., 2006).

These data are compatible with the theory that different human face processing tasks depend on a common representation for face shape, located in the human FFA, and that decreased selectivity of face neurons in the FFA would therefore result in broad face processing impairments across tasks (Jiang et al., 2006), in both neurotypical individuals (Gilaie-Dotan and Malach, 2007; Goh et al., 2010) and those with ASD (Weigelt et al., 2012). Previous brain imaging studies examining average BOLD contrast response amplitudes to faces have reported conflicting results regarding processing differences in the

FFA in autism, however, probably due to the fact that response amplitude may not be a good index of neuronal tuning specificity. Indeed, in our study we confirm that FFA activation level is not a predictor of face processing ability in individuals with ASD (see Fig. 5 for results with four different face indices, all based on average BOLD-contrast response), creating the need for more advanced MRI techniques to explore the neural bases of face processing deficits in autism.

In contrast, the hypothesis that variations in the selectivity of the face representation in the FFA are responsible for the heterogeneity of face processing abilities in autism is supported by our novel analysis of local voxel-wise correlations within the FFA. While fMRI correlation analyses usually focus on long-distance relationships between different brain regions, few studies have examined local correlations between voxels within a local region (Deshpande et al., 2009; Zang et al., 2004). In those previous approaches, a coherence value was usually assigned to each voxel based on correlations with its nearest neighbors, and this quantity (termed regional homogeneity, or ReHo) was then used for comparisons. Using this approach, two recent studies (Paakki et al., 2010; Shukla et al., 2010) have found altered ReHo in individuals with ASD, including both increased and decreased ReHo in several different brain regions, such as frontal and temporal cortices. For instance, Shukla et al. (2010) found increased ReHo in right fusiform gyrus in children with ASD. However, these previous approaches did not take into account the functional roles of voxels – i.e., a voxel and its neighbors may, or may not, be involved in the same cognitive function or be selective for similar stimuli – nor the variations in correlation with neighborhood voxels. To develop a better measure of voxel-wise response heterogeneity that takes the functional role of each voxel into account, we proposed a novel approach in which an ROI was functionally defined, followed by the computation of all pair-wise correlations between all voxel pairs within this ROI. The variability of these pair-wise correlation coefficients was calculated to define an indicator of local regional heterogeneity, H_{corr} , with a larger value indicating a higher degree of local regional heterogeneity. The rationale behind this approach is that voxel-wise activity correlations should be modulated by correlations in activity of the neurons in the voxels of interest. Crucially, these neuron-level correlations have been shown to be higher for neurons with similar tuning, and are also found in the absence of visual stimuli (Bair et al., 2001; Jermakowicz et al., 2009). H_{corr} therefore probes the variability of BOLD response correlations as an estimator of neuronal selectivity, rather than differences in activation maps for particular stimuli. Based on the sparse coding model for face representations in the FFA, we hypothesized that tuning differences at the neuronal level might be reflected in different levels of local regional heterogeneity. That is, a higher regional heterogeneity in the FFA would indicate sparser representations, which in turn would indicate higher selectivity of face neurons, which in turn should be reflected by better behavioral performance in face processing. Our data confirm this hypothesis, showing that, indeed, local regional heterogeneity is a strong predictor of behavioral performance. As predicted, this correlation was limited to brain regions thought to mediate face perception behavior, in particular the right FFA, but was not found in other brain regions, such as V1/2, nor the PPA, and not even in the more posterior face selective area, OFA. In contrast, there was no significant correlation between behavioral performance and the averaged local voxel-wise correlations in the FFA ($r = -0.43$, $p > 0.12$), suggesting H_{corr} is a more sensitive measure of neural selectivity, possibly because it is less affected by global, unspecific correlations between voxels activation levels, but rather focuses on local differences in correlations. To our knowledge, this is the first time that behavioral performance has been linked to local voxel-wise correlation measures in a functionally-defined ROI. Due to its simplicity, sensitivity, and robustness (in this case relying on less than 5.5 min of scan time), we expect that this method will be of interest for a wide range of applications in which differences in neural

selectivity can be linked to variations in task performance. A particular potential strength of the technique is the possibility to calculate H_{corr} from resting state data (or while the subject is fixating), which could make it possible to estimate neural selectivity (and thus predict behavioral performance on skills known to rely on neural selectivities in particular brain areas) throughout the brain from resting state scans. Indeed, we have already found that subject performance on a standard test of reading ability can be predicted from H_{corr} values in the Visual Word Form Area (VWFA) obtained from fixation blocks (Eden et al., 2011).

The hypothesis of reduced selectivity of face neurons in the FFA in autism is further supported by our fMRI-RA experiment. We had shown previously that face neuron selectivity as estimated with fMRI-RA was in excellent agreement with theoretical predictions (Jiang et al., 2006), and the results of the present study, providing evidence for delayed release from adaptation in the FFA of individuals with autism and face processing impairments, further strengthen the link between face neuron selectivity in the FFA and behavioral face processing abilities.

Neuroimaging studies have identified several brain regions involved in face processing (Haxby et al., 2000; Kanwisher and Yovel, 2006), and there is broad consensus that, in typical adults, the right FFA is the central region in face processing, as shown in a number of studies (Grill-Spector et al., 2004; Jiang et al., 2006; Yovel and Kanwisher, 2005). However, several fMRI studies have suggested that face processing deficits in autism might be due to abnormalities in other brain regions, such as the amygdala, frontal cortices, or the mirror neuron system rather than in the FFA per se (Bookheimer et al., 2008; Hadjikhani et al., 2007; Humphreys et al., 2008; Pierce et al., 2004), or that face processing in autism might rely on other brain regions, such as the more posterior occipital cortex (Hubl et al., 2003), or general non-face object-related regions (Scherf et al., 2010), or more individualized brain networks (Pierce et al., 2001). In contrast, in the present study, we probed the neural selectivity in the left and right FFA and other face and non-face brain regions, and found that only neural tuning in the right FFA (as indirectly estimated via fMRI-RA and variations in voxel-wise response correlations), but not in any other face-selective area, could reliably predict subjects' face discrimination performance. Thus, our results strongly suggest that face processing abilities in adults with autism are mediated by neural selectivity in the right FFA, but not any other face-selective regions, providing direct support for the theory that face processing is only quantitatively different, but qualitatively similar between individuals with ASD and typical adults (Jiang et al., 2006; Weigelt et al., 2012).

It has been postulated that selectivity differences in the FFA in autism might be due to weaker face-related saliency signals from the amygdala (Schultz, 2005), and other studies have provided evidence for reduced connectivity between the amygdala and the FFA (Conturo et al., 2008; Kleinhans et al., 2008). Such a weaker face-related input from the amygdala could affect experience-driven refinement of face-selective representations in the FFA during development (see Golarai et al. (2007)), following current theories of perceptual learning that have identified a key role of the amplitude of neuronal activity in enabling learning (Seitz and Dinse, 2007). In addition, anatomical differences in the fusiform cortex of individuals with ASD might limit the selectivity of neuronal representations there. Recent anatomical studies have identified cortical minicolumn abnormalities – specifically reduced minicolumn width – in autism (Casanova et al., 2006). The reduction in minicolumn width has been hypothesized to be associated with an increased overlap of afferents across minicolumns (Casanova et al., 2006) as well as decreased inhibition between minicolumns (Rubenstein and Merzenich, 2003). While these effects have previously been hypothesized to give rise to sparser representations (Casanova et al., 2006), both of these anatomical differences would in fact serve to broaden the selectivity of neural tuning (Jiang et al., 2006; Wörgötter

and Koch, 1991), as predicted by the model and borne out in our neuroimaging results. It is also conceivable that the two theories are linked, in that the aforementioned microstructural differences might reflect slowed-down experience-driven maturation of neural selectivity, following animal studies that have found an experience-dependent reduction in plasticity mediated by a progressive reduction in intracortical inhibition during development (e.g., (Dornn et al., 2010)). Yet, given that the brain areas studied in Casanova et al. (2006) did not include the fusiform gyrus, the link of variations in minicolumn width to the differences in neural selectivity in the FFA suggested by our study remains speculative and further studies are needed.

If selectivity differences in the FFA in autism are due to reduced experience-driven refinement, training individuals with autism on tasks that recruit the face representation in the FFA (and compensate for a lack of amygdala-based face saliency by task-dependent saliency, cf. (Adolphs et al., 2005)) might serve to sharpen up its selectivity and remediate face-processing differences. Interestingly, it has been shown in a prosopagnosic individual that training using a face categorization task optimized to engage face representations in the FFA led to broad improvements in behavioral performance on a wide range of face perception tasks (DeGutis et al., 2007), supporting a model in which the face representation in the FFA can serve as a general representation for a wide range of face-related tasks (Riesenhuber and Poggio, 2002; Riesenhuber and Wolff, 2009). Future studies will have to show whether targeted training can indeed refine the selectivity of neural representations in the FFA to broadly improve face processing abilities in autism.

Acknowledgements

We thank Joshua O. Goh for discussions, and Greg Wallace, Alex Martin, and Lauren Kenworthy for help with subject recruitment. Additional subjects were recruited with the assistance of the Interactive Autism Network (IAN) Research Database at the Kennedy Krieger Institute and Johns Hopkins Medicine – Baltimore, sponsored by the Autism Speaks Foundation. The research reported in this paper was supported by NIMH grant R01MH076281 and NSF grant 0449743 (to M.R.), SNF PPOOB-110741 (to N.H.), as well as NIH IDDR P30HD40677, and NIH GCRC M01-RR13297.

References

- Adolphs, R., Sears, L., Piven, J., 2001. Abnormal processing of social information from faces in autism. *Journal of Cognitive Neuroscience* 13, 232–240.
- Adolphs, R., Gosselin, F., Buchanan, T.W., Tranel, D., Schyns, P., Damasio, A.R., 2005. A mechanism for impaired fear recognition after amygdala damage. *Nature* 433, 68–72.
- Aguirre, G.K., Zarahn, E., D'Esposito, M., 1998. The inferential impact of global signal covariates in functional neuroimaging analyses. *NeuroImage* 8, 302–306.
- Bair, W., Zohary, E., Newsome, W.T., 2001. Correlated firing in macaque visual area MT: time scales and relationship to behavior. *Journal of Neuroscience* 21, 1676–1697.
- Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., Clubley, E., 2001. The autism-spectrum quotient (AQ): evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders* 31, 5–17.
- Barton, J.J.S., Cherkasova, M.V., Hefter, R., Cox, T.A., O'Connor, M., Manoach, D.S., 2004. Are patients with social developmental disorders prosopagnosic? Perceptual heterogeneity in the Asperger and socio-emotional processing disorders. *Brain* 127, 1706–1716.
- Blanz, V., Vetter, T., 1999. A morphable model for the synthesis of 3D faces. *Proceedings of the 26th annual conference on Computer graphics and interactive techniques*, pp. 187–194. <http://dx.doi.org/10.1145/311535.311556>.
- Bölte, S., Hubl, D., Feineis-Matthews, S., Prvulovic, D., Dierks, T., Poustka, F., 2006. Facial affect recognition training in autism: can we animate the fusiform gyrus? *Behavioral Neuroscience* 120, 211–216.
- Bookheimer, S.Y., Wang, A.T., Scott, A., Sigman, M., Dapretto, M., 2008. Frontal contributions to face processing differences in autism: evidence from fMRI of inverted face processing. *Journal of International Neuropsychological Society* 14, 922–932.
- Brett, M., Anton, J.-L., Valabregue, R., Poline, J.-B., 2002. Region of interest analysis using an SPM toolbox [abstract]. Presented at the 8th International Conference on Functional Mapping of the Human Brain, June 2–6, 2002, Sendai, Japan (Available on

- CD-ROM in *NeuroImage*, Vol 16, No 2. (2002).at https://cirl.berkeley.edu/mb312/abstracts/Marsbar/marsbar_abs.html).
- Buracas, G.T., Boynton, G.M., 2002. Efficient design of event-related fMRI experiments using M-sequences. *NeuroImage* 16, 801–813.
- Casanova, M.F., van Kooten, I.A.J., Switala, A.E., van Engeland, H., Heinsen, H., Steinbusch, H.W.M., Hof, P.R., Trippe, J., Stone, J., Schmitz, C., 2006. Minicolumnar abnormalities in autism. *Acta Neuropathologica* 112, 287–303.
- Conturo, T.E., Williams, D.L., Smith, C.D., Gultepe, E., Akbudak, E., Minshew, N.J., 2008. Neuronal fiber pathway abnormalities in autism: an initial MRI diffusion tensor tracking study of hippocampo-fusiform and amygdalo-fusiform pathways. *Journal of International Neuropsychological Society* 14, 933–946.
- Critchley, H.D., Daly, E.M., Bullmore, E.T., Williams, S.C., Van Amelsvoort, T., Robertson, D.M., Rowe, A., Phillips, M., McAlonan, G., Howlin, P., Murphy, D.G., 2000. The functional neuroanatomy of social behaviour: changes in cerebral blood flow when people with autistic disorder process facial expressions. *Brain* 123, 2203–2212.
- Dawson, G., Webb, S.J., Wijsman, E., Schellenberg, G., Estes, A., Munson, J., Faja, S., 2005. Neurocognitive and electrophysiological evidence of altered face processing in parents of children with autism: implications for a model of abnormal development of social brain circuitry in autism. *Development and Psychopathology* 17, 679–697.
- DeGutis, J.M., Bentin, S., Robertson, L.C., D'Esposito, M., 2007. Functional plasticity in ventral temporal cortex following cognitive rehabilitation of a congenital prosopagnosic. *Journal of Cognitive Neuroscience* 19, 1790–1802.
- Deshpande, G., LaConte, S., Peltier, S., Hu, X., 2009. Integrated local correlation: a new measure of local coherence in fMRI data. *Human Brain Mapping* 30, 13–23.
- Diamond, R., Carey, S., 1986. Why faces are and are not special: an effect of expertise. *Journal of Experimental Psychology: General* 115, 107–117.
- Dorrn, A.L., Yuan, K., Barker, A.J., Schreiner, C.E., Froemke, R.C., 2010. Developmental sensory experience balances cortical excitation and inhibition. *Nature* 465, 932–936.
- Eden, G.F., Flowers, D.L., Napoliello, E.M., Luetje, M., Jiang, X., 2011. Evidence for a relationship between the heterogeneity of local regional correlations within the VWFA and reading ability. Poster presented at Annual Meeting of the Society for Neuroscience, Washington, DC.
- Epstein, R., Kanwisher, N., 1998. A cortical representation of the local visual environment. *Nature* 392, 598–601.
- Fang, F., Murray, S.O., He, S., 2007. Duration-dependent fMRI adaptation and distributed viewer-centered face representation in human visual cortex. *Cerebral Cortex* 17, 1402–1411.
- Freedman, D.J., Riesenhuber, M., Poggio, T., Miller, E.K., 2006. Experience-dependent sharpening of visual shape selectivity in inferior temporal cortex. *Cerebral Cortex* 16, 1631–1644.
- Friston, K.J., Frith, C.D., Turner, R., Frackowiak, R.S., 1995. Characterizing evoked hemodynamics with fMRI. *NeuroImage* 2, 157–165.
- Geschwind, D.H., 2009. Advances in autism. *Annual Review of Medicine* 60, 367–380.
- Gilaie-Dotan, S., Malach, R., 2007. Sub-exemplar shape tuning in human face-related areas. *Cerebral Cortex* 17, 325–338.
- Goh, J.O., Suzuki, A., Park, D.C., 2010. Reduced neural selectivity increases fMRI adaptation with age during face discrimination. *NeuroImage* 51, 336–344.
- Golarai, G., Ghahremani, D.G., Whitfield-Gabrieli, S., Reiss, A., Eberhardt, J.L., Gabrieli, J.D.E., Grill-Spector, K., 2007. Differential development of high-level visual cortex correlates with category-specific recognition memory. *Nature Neuroscience* 10, 512–522.
- Grady, C.L., Horwitz, B., Pietrini, P., Mentis, M.J., Ungerleider, L.G., Rapoport, S.L., Haxby, J.V., 1996. Effect of task difficulty on cerebral blood flow during perceptual matching of faces. *Human Brain Mapping* 4, 227–239.
- Grelotti, D.J., Klin, A.J., Gauthier, I., Skudlarski, P., Cohen, D.J., Gore, J.C., Volkmar, F.R., Schultz, R.T., 2005. fMRI activation of the fusiform gyrus and amygdala to cartoon characters but not to faces in a boy with autism. *Neuropsychologia* 43, 373–385.
- Grill-Spector, K., Knouf, N., Kanwisher, N., 2004. The fusiform face area subserves face perception, not generic within-category identification. *Nature Neuroscience* 7, 555–562.
- Grill-Spector, K., Henson, R., Martin, A., 2006. Repetition and the brain: neural models of stimulus-specific effects. *Trends in Cognitive Sciences* 10, 14–23 (Regul. Ed.).
- Hadjikhani, N., Joseph, R.M., Snyder, J., Tager-Flusberg, H., 2007. Abnormal activation of the social brain during face perception in autism. *Human Brain Mapping* 28, 441–449.
- Haxby, J.V., Ungerleider, L.G., Clark, V.P., Schouten, J.L., Hoffman, E.A., Martin, A., 1999. The effect of face inversion on activity in human neural systems for face and object perception. *Neuron* 22, 189–199.
- Haxby, J.V., Hoffman, E.A., Gobbini, M.I., 2000. The distributed human neural system for face perception. *Trends in Cognitive Sciences* 4, 223–233.
- Hedley, D., Brewer, N., Young, R., 2011. Face recognition performance of individuals with Asperger syndrome on the Cambridge Face Memory Test. *Autism Research* 4, 449–455.
- Hubl, D., Bölte, S., Feineis-Matthews, S., Lanfermann, H., Federspiel, A., Strik, W., Poustka, F., Dierks, T., 2003. Functional imbalance of visual pathways indicates alternative face processing strategies in autism. *Neurology* 61, 1232–1237.
- Humphreys, K., Hasson, U., Avidan, G., Minshew, N., Behrmann, M., 2008. Cortical patterns of category-selective activation for faces, places and objects in adults with autism. *Autism Research* 1, 52–63.
- Jermakowicz, W.J., Chen, X., Khaytin, I., Bonds, A.B., Casagrande, V.A., 2009. Relationship between spontaneous and evoked spike-time correlations in primate visual cortex. *Journal of Neurophysiology* 101, 2279–2289.
- Jiang, X., Rosen, E., Zeffiro, T., VanMeter, J., Blanz, V., Riesenhuber, M., 2006. Evaluation of a shape-based model of human face discrimination using fMRI and behavioral techniques. *Neuron* 50, 159–172.
- Jiang, X., Bradley, E., Rini, R.A., Zeffiro, T., VanMeter, J., Riesenhuber, M., 2007. Categorization training results in shape- and category-selective human neural plasticity. *Neuron* 53, 891–903.
- 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.
- Kanwisher, N., McDermott, J., Chun, M.M., 1997. The fusiform face area: a module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience* 17, 4302–4311.
- Kleinhans, N.M., Bölte, S., Feineis-Matthews, S., Lanfermann, H., Federspiel, A., Strik, W., Poustka, F., Dierks, T., 2008. Abnormal functional connectivity in autism spectrum disorders during face processing. *Brain* 131, 1000–1012.
- Klin, A., Sparrow, S.S., de Bildt, A., Cicchetti, D.V., Cohen, D.J., Volkmar, F.R., 1999. A normed study of face recognition in autism and related disorders. *Journal of Autism and Developmental Disorders* 29, 499–508.
- Kobatake, E., Wang, G., Tanaka, K., 1998. Effects of shape-discrimination training on the selectivity of inferotemporal cells in adult monkeys. *Journal of Neurophysiology* 80, 324–330.
- Kourtzi, Z., Erb, M., Grodd, W., Bühlhoff, H.H., 2003. Representation of the perceived 3-D object shape in the human lateral occipital complex. *Cerebral Cortex* 13, 911–920.
- Mahon, B.Z., Milleville, S.C., Negri, G.A.L., Rumiati, R.I., Caramazza, A., Martin, A., 2007. Action-related properties shape object representations in the ventral stream. *Neuron* 55, 507–520.
- Miller, E.K., Li, L., Desimone, R., 1993. Activity of neurons in anterior inferior temporal cortex during a short-term memory task. *Journal of Neuroscience* 13, 1460–1478.
- Murray, S.O., Wojciulik, E., 2004. Attention increases neural selectivity in the human lateral occipital complex. *Nature Neuroscience* 7, 70–74.
- Paakkii, J.J., Rahko, J., Long, X., Moilanen, I., Tervonen, O., Nikkinen, J., Starck, T., Remes, J., Hurtig, T., Haapsamo, H., Jussila, L., Kuusikko-Gauffin, S., Mattila, M.L., Zang, Y., Kiviniemi, V., 2010. Alterations in regional homogeneity of resting-state brain activity in autism spectrum disorders. *Brain Research* 1321, 169–179.
- Perlman, S.B., Hudac, C.M., Pegors, T., Minshew, N.J., Pelphrey, K.A., 2011. Experimental manipulation of face-evoked activity in the fusiform gyrus of individuals with autism. *Social Neuroscience* 6, 22–30.
- Pierce, K., Müller, R.A., Ambrose, J., Allen, G., Courchesne, E., 2001. Face processing occurs outside the fusiform “face area” in autism: evidence from functional MRI. *Brain* 124, 2059–2073.
- Pierce, K., Haist, F., Sedaghat, F., Courchesne, E., 2004. The brain response to personally familiar faces in autism: findings of fusiform activity and beyond. *Brain* 127, 2703–2716.
- Riesenhuber, M., 2007. Appearance isn't everything: news on object representation in cortex. *Neuron* 55, 341–344.
- Riesenhuber, M., Poggio, T., 2002. Neural mechanisms of object recognition. *Current Opinion in Neurobiology* 12, 162–168.
- Riesenhuber, M., Wolff, B.S., 2009. Task effects, performance levels, features, configurations, and holistic face processing: a reply to Rossion. *Acta Psychologica* 132, 286–292.
- Rubenstein, J.L.R., Merzenich, M.M., 2003. Model of autism: increased ratio of excitation/inhibition in key neural systems. *Genes, Brain, and Behavior* 2, 255–267.
- Scherf, K.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. *Frontiers in Human Neuroscience* 4, 26.
- Schultz, R.T., 2005. Developmental deficits in social perception in autism: the role of the amygdala and fusiform face area. *International Journal of Developmental Neuroscience* 23, 125–141.
- Schultz, R.T., Gauthier, I., Klin, A., Fulbright, R.K., Anderson, A.W., Volkmar, F., Skudlarski, P., Lacadie, C., Cohen, D.J., Gore, J.C., 2000. Abnormal ventral temporal cortical activity during face discrimination among individuals with autism and Asperger syndrome. *Archives of General Psychiatry* 57, 331–340.
- Seitz, A.R., Dinse, H.R., 2007. A common framework for perceptual learning. *Current Opinion in Neurobiology* 17, 148–153.
- Shukla, K.S., Keehn, B., Muller, R.A., 2010. Regional homogeneity of fMRI time series in autism spectrum disorders. *Neuroscience Letters* 476, 46–51.
- Snow, J., Ingeholm, J.E., Levy, I.F., Caravella, R.A., Case, L.K., Wallace, G.L., Martin, A., 2011. Impaired visual scanning and memory for faces in high-functioning autism spectrum disorders: it's not just the eyes. *Journal of International Neuropsychological Society* 17, 1021–1029.
- Sunaert, S., Van Hecke, P., Marchal, G., Orban, G.A., 2000. Attention to speed of motion, speed discrimination, and task difficulty: an fMRI study. *NeuroImage* 11, 612–623.
- Weeks, S.J., Hobson, R.P., 1987. The salience of facial expression for autistic children. *Journal of Child Psychology and Psychiatry* 28, 137–151.
- Weigelt, S., Koldewyn, K., Kanwisher, N., 2012. Face identity recognition in autism spectrum disorders: a review of behavioral studies. *Neuroscience and Biobehavioral Reviews* 36, 1060–1084.
- Wörgötter, F., Koch, C., 1991. A detailed model of the primary visual pathway in the cat: comparison of afferent excitatory and intracortical inhibitory connection schemes for orientation selectivity. *Journal of Neuroscience* 11, 1959–1979.
- Yovel, G., Kanwisher, N., 2005. The neural basis of the behavioral face-inversion effect. *Current Biology* 15, 2256–2262.
- Zang, Y., Jiang, T., Lu, Y., He, Y., Tian, L., 2004. Regional homogeneity approach to fMRI data analysis. *NeuroImage* 22, 394–400.