

## Multi-voxel pattern analysis of fMRI data predicts clinical symptom severity

Marc N. Coutanche<sup>a,\*</sup>, Sharon L. Thompson-Schill<sup>a</sup>, Robert T. Schultz<sup>b</sup>

<sup>a</sup> Department of Psychology, University of Pennsylvania, 3720 Walnut Street, Philadelphia, PA 19104, USA

<sup>b</sup> Center for Autism Research, Children's Hospital of Philadelphia, 3535 Market Street, Philadelphia, PA 19104, USA

### ARTICLE INFO

#### Article history:

Received 1 November 2010

Revised 29 March 2011

Accepted 7 April 2011

Available online 13 April 2011

#### Keywords:

MVPA

fMRI

Multivariate

Pattern

Patient

Autism

### ABSTRACT

Multi-voxel pattern analysis (MVPA) has been applied successfully to a variety of fMRI research questions in healthy participants. The full potential of applying MVPA to functional data from patient groups has yet to be fully explored. Our goal in this study was to investigate whether MVPA might yield a sensitive predictor of patient symptoms. We also sought to demonstrate that this benefit can be realized from existing datasets, even when they were not designed with MVPA in mind. We analyzed data from an fMRI study of the neural basis for face processing in individuals with an Autism Spectrum Disorder (ASD), who often show fusiform gyrus hypoactivation when presented with unfamiliar faces, compared to controls. We found reliable correlations between MVPA classification performance and standardized measures of symptom severity that exceeded those observed using a univariate measure; a relation that was robust across variations in ROI definition. A searchlight analysis across the ventral temporal lobes identified regions with relationships between classification performance and symptom severity that were not detected using mean activation. These analyses illustrate that MVPA has the potential to act as a sensitive functional biomarker of patient severity.

© 2011 Elsevier Inc. All rights reserved.

### Introduction

In 2001, Haxby and colleagues demonstrated that information that was not evident from univariate analyses could be decoded from patterns of fMRI activation across voxels (Haxby et al., 2001). Since the publication of this seminal study, a class of techniques referred to as multi-voxel pattern analysis (MVPA) has been applied to a variety of questions (e.g., Cox and Savoy, 2003; O'Toole et al., 2005; Spiridon and Kanwisher, 2002; for reviews see Haynes and Rees, 2006; Norman et al., 2006; O'Toole et al., 2007). These techniques analyze recorded activity patterns using tools such as machine learning classifiers, to measure the information present within populations of voxels.

Despite the success of applying MVPA in a wide range of contexts, few studies have extended the method to investigations of atypical neural activity. Several functional studies have used multivariate approaches to classify individuals into different groups (in contrast to classifying trials into conditions) for depression (Fu et al., 2008) and drug addiction (Zhang et al., 2005). Similarly, functional models have predicted future responses to Cognitive Behavioral Therapy (full vs. partial) in depressed patients (Costafreda et al., 2009) and estimated years-to-onset of Huntington's disease symptoms (Rizk-Jackson et al., 2010). Only a small number of clinical studies have conducted within-subject between-condition MVPA: abnormal activity patterns have

been reported during object representation and working memory processes in schizophrenia (Kim et al., 2010; Yoon et al., 2008), and unusual patterns have been detected in the medial prefrontal cortex of participants on the autism spectrum during mental state reflections (Gilbert et al., 2009). None of these patient studies used MVPA measures to predict individual differences in clinical symptom severity.

The primary goal of this paper is to report the potential for MVPA to give high levels of sensitivity in relating fMRI data to patient symptoms. By incorporating the unique contributions of individual voxels, subtleties within activation patterns are reflected in MVPA outcome measures, such as classification performance. Such subtleties are often ignored in univariate analyses, where the levels of voxel activation, or mean activation of a region, are evaluated. In a region that is functionally relevant to a disorder with atypical cognitive or behavioral symptoms, this multivariate characterization could act as a sensitive measure of variation among affected individuals. To investigate this possibility here, we use MVPA to examine a dataset from a study of fusiform gyrus activation in individuals with an Autism Spectrum Disorder (ASD; Schultz et al., 2008).

Investigating the face processing differences in people with autism has been a very active area of research, not least because of the importance of face processing to successful social functioning. Among other deficits, ASD patients show large impairments in recognizing facial identity across changes in viewing conditions (Wolf et al., 2008), despite typical performance at processing complex objects (Boucher and Lewis, 1992; Wolf et al., 2008). A large number of functional neuroimaging investigations have studied the neural substrates of

\* Corresponding author. Fax: +1 215 898 1982.

E-mail address: [coumarc@psych.upenn.edu](mailto:coumarc@psych.upenn.edu) (M.N. Coutanche).

these behavioral abnormalities, particularly in the fusiform gyrus, a highly face-selective brain region, which has come to be known as the fusiform face area (FFA; Kanwisher et al., 1997; Kanwisher and Yovel, 2006; Winston et al., 2004), although the specificity of fusiform computations is much debated (Gauthier et al., 1999; Kanwisher and Yovel, 2006; Schultz et al., 2003). The FFA is strongly activated when typically-developing individuals view faces, but is frequently hypoactive when individuals with an ASD view unfamiliar faces (Critchley et al., 2000; Deeley et al., 2007; Grelotti et al., 2005; Hall et al., 2003; Hubl et al., 2003; Koshino et al., 2008; Pierce et al., 2001; Piggot et al., 2004; Schultz et al., 2000, 2008; Wang et al., 2004). The processes responsible for this relative hypoactivation are an area of ongoing debate.<sup>1</sup> Although it is a related and important question, the present study is neutral on the proximate causes of hypoactivation, focusing instead on the potential for MVPA to give a sensitive functional biomarker.

A secondary aim of the paper is to provide an example of the way MVPA can be used to realize this benefit in studies designed without MVPA in mind. This suggestion may resonate with patient-group investigators looking to make the most of existing datasets, which are often expensive, time-consuming and logistically difficult to obtain. In this study, we illustrate this point with an extreme case; analyzing a dataset from an fMRI study that was not planned with MVPA in mind and that was, in many ways, sub-optimal for this purpose. Designing a study for subsequent MVPA typically involves a number of considerations: The established sensitivity of MVPA to subtle visual differences (e.g., Kamitani and Tong, 2005) makes controlling visual properties, such as luminance and the visual angle of presented images, particularly important. Additionally, in order to draw conclusions about a specific category, an appropriate number of classes are required. Any two-way classification is affected by the activity patterns of both classes, so multiple comparisons are required for drawing conclusions about one condition of interest. For example, above-chance classification between class A and class B could be due to encoded class A information (where the classifier succeeds based on 'A vs. not-A') or class B information ('B vs. not-B'). The successful separation of A vs. B, C and D, but not among B, C and D, however, gives some confidence that A, or at least certain features of A, are central to an area's encoded information (as applied in O'Toole et al., 2005; Spiridon and Kanwisher, 2002).

The above design considerations are recommended where possible; however, it is still possible to benefit from the MVPA approach when a dataset has been designed for univariate analyses. The dataset analyzed here was collected to examine how fusiform activation varies during face tasks that differ in their attentional and perceptual loads (Schultz et al., 2008), in individuals with autism. Designed specifically for univariate analyses, the study was organized in a way contrary to the optimal design considerations reviewed above: the stimuli in each condition were not individually matched for luminance, and two categories of stimuli, faces and houses, were presented to participants. Additionally, the house condition required participants to make a 'same' vs. 'different' judgment about two side-by-side houses, while the face condition was intentionally varied along several dimensions by run, including the perceptual judgments required, number of face stimuli, and presence or absence of emotional expression. A constant house condition was included in the study to act as a common baseline for between-run comparisons of the different face tasks. Finally, each

face condition was allocated a relatively short amount of fMRI time (five 20-second blocks each) giving a small number of trials for each of the face conditions.

In this study, we applied MVPA to four of the six runs in this fMRI dataset by classifying the activity patterns for viewing faces and houses within the participants. By grouping together the different face trials into one class, we were able to increase the number of trials to a suitable level for performing MVPA (see Pereira et al., 2009 for a discussion of the factors relevant to selecting classifier exemplars), while allowing us to investigate underlying commonalities in face activity patterns between the different face tasks. Our results showed that classification performance was more strongly related to symptom severity than a univariate measure of mean activation. The greater sensitivity of MVPA was consistent across a variety of approaches to defining the regions of interest, including an anatomical definition, face-responsive voxels defined in the control group, and even in an area defined based on the mean activation difference itself. Furthermore, using a roaming searchlight analysis across the ventral temporal (VT) lobe, we found a symptom severity relationship with MVPA in regions of cortex that were not highlighted using a univariate measure. This is the first study, to our knowledge, that reports a link between functional MVPA results and standardized measures of symptom severity: an important target for many patient-based investigations. We also hope this study will be encouraging to patient-group researchers who are looking to maximize the utility of existing fMRI datasets, and to those searching for functional techniques that are sensitive to patient symptoms.

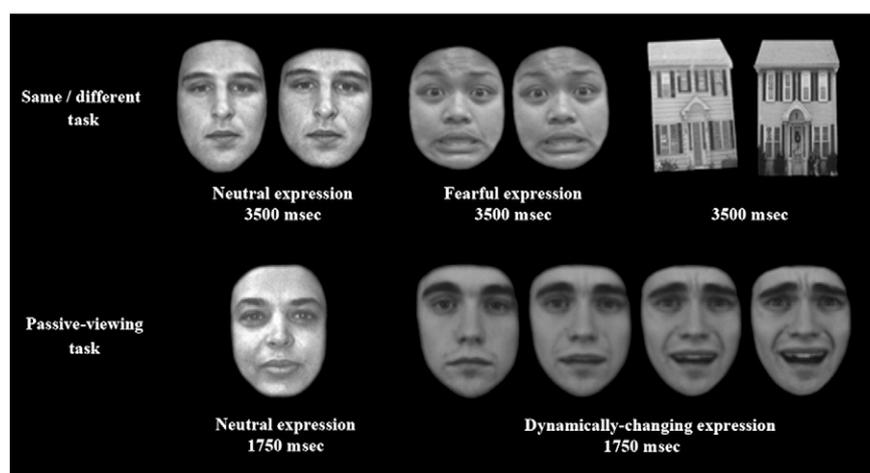
## Materials and methods

### Participants

Twelve males on the autism spectrum (ages 9.3–24.2, mean (M) = 13.9 years) and twelve typically-developing male controls (ages 9.4–23.3, M = 13.6 years) were selected for these analyses from a total sample of more than twenty in each group on the basis of having the lowest scanner movement, while matching the groups by age. All participants were recruited and studied at the Yale Child Study Center. All participants or their legal guardians gave informed written consent and were compensated for their participation, in accordance with procedures and protocols approved by the Institutional Review Board of the Yale University School of Medicine. Each ASD diagnosis was confirmed by a consensus diagnosis process involving two Ph.D level clinicians experienced with ASD differential diagnoses, using results from the Autism Diagnostic Observation Schedule (ADOS; Lord et al., 1989, 2000) and Autism Diagnostic Interview-Revised (ADI-R; Le Couteur et al., 1989; Lord et al., 1994). Total ADOS scores ranged between 9 and 22 (M = 15.9, standard deviation (s.d.) = 4.3). The ADOS is a set of standardized semi-structured interactions between a clinician and the relevant individual. The participant's behaviors and responses are recorded and scored by the clinician against a series of standardized categories. An algorithm is subsequently used to combine these scores, where a cut-off helps determine each diagnosis. The ADI is an extensive structured interview, conducted by a clinician, with a parent of the patient. The interviewee is questioned extensively about his or her knowledge and experiences of the patient's current and prior behaviors, and developmental trajectory. Responses are scored according to an established list of criteria, which are then combined through a standardized algorithm. These forms of assessment are suitable for participants of all ages: the clinician administering the ADOS selects one of four different modules of interactions, based on the mental age of the individual being assessed.

Control participants were screened for personal and family histories of psychiatric disorders and neurological trauma, and current axis I disorders with standardized symptom inventories that cover all DSM-IV axis I disorders: the parent report Childhood

<sup>1</sup> For example, recent evidence has suggested that hypoactivation may be less robust for familiar faces (Pierce et al., 2004; Pierce and Redcay, 2008), but this is not always found (Grelotti et al., 2005). A further debate concerns the source of fusiform signal variance in ASD patients: while some studies have suggested that individual differences in looking behavior may drive signal variance (Bookheimer et al., 2008; Dalton et al., 2005; Hadjikhani et al., 2004, 2007; Perlman et al., 2010), the paradigms used in several of these studies have been critiqued (Klin, 2008; Schultz et al., 2008) and not all investigations have been able to confirm this (Humphreys et al., 2008; Schultz et al., 2008).



**Fig. 1.** Examples of stimuli with associated presentation times. Two of the analyzed runs featured face stimuli in a same vs. different task (top row) and two featured a passive-viewing task (bottom row).

Symptom Inventory for children aged 5–11 years (Gadow and Sprafkin, 1994), the Adolescent Screening Inventory for individuals aged 12–18 years (Gadow et al., 2002), and the Adult Self-Report Inventory for individuals older than 18 years (Gadow et al., 1999).

Participants' IQs were assessed using the WASI, WISC-IV or WAIS-III. The ASD group's mean IQ ( $M = 101.6$ ,  $s.d. = 21.8$ ) was lower than the control group's ( $M = 112.8$ ,  $s.d. = 10.1$ ), although not significantly so ( $t_{22} = 1.60$ ,  $p = 0.12$ ). The groups did not differ in chronological age ( $t_{22} = 0.19$ ,  $p = 0.85$ ). All participants were administered the Benton Test of Facial Recognition (Benton et al., 1994), the Edinburgh Handedness Inventory (Oldfield, 1971) and a computerized battery of face recognition, perception and memory (Wolf et al., 2008) as part of the original study. All participants, except for one in each group, were right handed. Scores on the Benton Facial Recognition Task were significantly higher ( $t_{22} = 3.45$ ,  $p = 0.002$ ) in controls ( $M = 43.1$ ,  $s.d. = 2.8$ ) than in ASD individuals ( $M = 38.1$ ,  $s.d. = 4.2$ ), as expected (Wolf et al., 2008).

### Stimuli and design

Participants were presented with gray-scale face and house images. The facial expressions and behavioral task differed among the four runs (see Fig. 1). Images were presented in alternating 20-second blocks separated by 12 s of rest, giving a total of five face and five house blocks in each run. In the first face task, participants viewed two neutral expression faces side-by-side and indicated with a button press whether the images showed the same or different people. The pictures were presented for 3500 ms, followed by a 500 ms inter-stimulus interval (ISI). Participants were asked to answer as quickly and confidently as possible. The second task followed the same format as the first but with fearful instead of neutral faces. In the third task, participants passively-viewed single neutral faces for 1750 ms, followed by a 250 ms ISI. The fourth task had participants passively view dynamic movies of faces changing their expression from neutral to fearful. The neutral face appeared for 1250 ms, followed by 500 ms of an emotion morph and then 1750 ms of a fearful expression. The house stimuli in each run were presented side-by-side for 3500 ms and participants were asked to indicate with a button press whether the images showed the same or different houses. Run order was counterbalanced across participants by reversing the task order in half the sample. The block order varied across runs within participants.<sup>2</sup>

<sup>2</sup> The original experiment also contained a fifth run investigating image fixation on a cross hair presented between the eyes on each face, and a sixth run with overlapping faces and houses, which were not analyzed here.

All participants practiced the tasks for 20 to 40 min before the scanning session until it was clear that they understood the tasks. All participants also underwent mock scanning to habituate them to the scanning environment and to train them to stay still.

Neutral face stimuli were taken from a single standardized set (Endl et al., 1998). Fearful face images were from multiple sources (Gur et al., 2002; Karolinska directed emotional faces image set, Lundqvist and Litton, 1998; NimStim Face Stimulus Set, Tottenham et al., 2002; Japanese and Caucasian Facial Expressions of Emotion; Matsumoto and Ekman, 1988, California Facial Expressions, Dailey et al., 2001; Pictures of Facial Affect, Ekman and Friesen, 1976; in-house database). The dynamic faces were from a custom stimuli collection created with MorphMan 2000 Software (STOIK Imaging, Moscow). All stimuli were presented once during the experiment. Hair, ears and any other peripheral identifying features were cropped from all faces. The house stimuli were custom photographs of houses in New Haven, CT. All stimuli were converted to grayscale and resized to  $150 \times 210$  pixels. Stimuli were presented using E-Prime (Psychology Software Tools Inc., Pittsburgh, PA) and PsyScope 1.2.5 for the dynamic faces (Cohen et al., 1993).

### Scanning

Functional T2-weighted images were acquired using a Siemens Trio 3-T scanner with a standard quadrature head coil. 40 axial slices were acquired parallel to the AC–PC plane, with whole-brain coverage. In-plane voxel size =  $3.516 \times 3.516$  mm, slice thickness = 3.5 mm with no gap, TR = 2320 ms, TE = 25, flip angle =  $60^\circ$ . T1-weighted anatomical images with the same thickness were acquired in the same session (TR = 300, TE = 2.43, flip angle =  $60^\circ$ ), as was a T1-weighted 3D anatomical data set (MP-RAGE, TR = 2530, TE = 3.66, TI = 1100, flip angle =  $7^\circ$ , resulting in  $1 \text{ mm}^3$  voxels). Stimuli were presented on a translucent screen at the rear of the scanner and viewed through a periscopic prism system mounted on the head coil. Singly-presented faces subtended approximately  $8^\circ$  of visual angle horizontally in the middle of the screen. Side-by-side stimuli subtended  $19^\circ$  of visual angle horizontally. During trials requiring a response, participants used fiber-optic button boxes in each hand to indicate their choice.

### Imaging preprocessing

Imaging data were preprocessed using the Analysis of Functional NeuroImages (AFNI) software package (Cox, 1996). All functional images were slice time corrected and deobliques to bring them in line with the axial plane. The first four volumes of each functional run

were removed to allow the signal to homogenize, and signal spikes were removed (using AFNI's 3dDespike). A motion correction algorithm was applied to register all volumes to the closest functional volume to the anatomical scan (Cox and Jesmanowicz, 1999) and the anatomical image was aligned to this same functional volume. Linear and quadratic trends were removed from each run and low frequency patterns were removed using a high-pass filter threshold of 0.0125 Hz. Voxel activation was scaled to have a mean of 100, with a maximum limit of 200. The skull was removed from the anatomical images, and all images were transformed into a standardized space (Talairach and Tournoux, 1988). Voxels in the functional datasets were resampled in the process to 3.5 mm × 3.5 mm × 3.5 mm. All participants displayed less than 3.6 mm of movement. The two groups did not differ in mean ( $t_{22} = 0.87$ ,  $p = 0.39$ ) or maximum ( $t_{22} = 1.30$ ,  $p = 0.21$ ) movement in the scanner.

### Regions

All analyses were conducted within the boundaries of the VT cortex, which was manually defined for each participant in the same manner as in previous studies (Haxby et al., 2001), extending 70 to 20 mm posterior to the anterior commissure in Talairach brain atlas coordinates, consisting of lingual, parahippocampal, fusiform and inferior temporal gyri. To examine the area of the fusiform gyrus typically activated by faces, we defined regions of interest in three distinct ways. Although a face localizer is a frequent method for locating the fusiform area responsive to faces, this is not suitable for a patient group known to show reduced fusiform activation to faces. In the first approach for sampling the relevant fusiform area, we placed three spheres (radius 0.5 cm each) at the average Talairach coordinates for the centers of right hemisphere FFA activation from face vs. object comparisons in previous studies (+40x, -55y, -10z from Kanwisher et al., 1997; +38x, -58y, -10z and +36x, -50y, -10z from Schultz et al.'s (2000) control groups 1 and 2). The overlapping spheres covered a total region of 29 voxels. Although the FFA's exact location varies between individuals, this approach was one way to approximate the area typically involved in face processing. We focused on the right hemisphere in this approach, as left hemisphere coordinates were not given in one of the published studies (Schultz et al., 2000), and only a small proportion of participants contributed to the average left hemisphere coordinates in the other (Kanwisher et al., 1997).

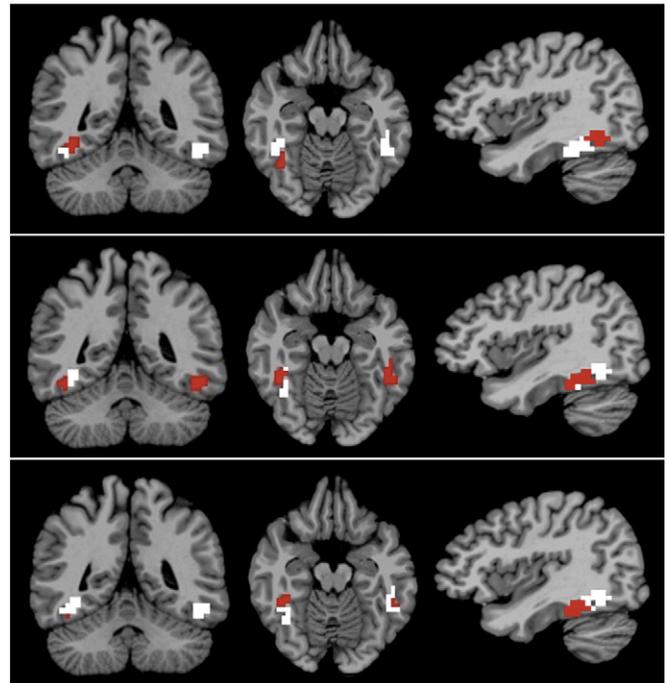
In the second approach, we isolated the area of significantly greater activation to faces than to houses in the right and left VT cortices of the control group. For this purpose we performed a traditional univariate analysis on images smoothed with a Gaussian filter (full-width, half-maximum = 8 mm) and extracted beta values from a face vs. house contrast using the four runs (with six nuisance vectors for translation and rotation movement), in the VT lobe of control participants. Corrected significance was established with a Monte Carlo voxel-cluster threshold technique (program AlphaSim) in the intersection of participants' VT regions, giving an overall corrected alpha level of 0.05 (voxelwise  $p < 0.001$ ; cluster size  $\geq 3$  voxels). This resulted in a 25-voxel right fusiform gyrus cluster (center of mass at +41x, -39y, -17z) and a 29-voxel left fusiform gyrus cluster (center of mass at -47x, -41y, -16z) that were significantly more active to faces than houses in the control group. Conducting MVPA within these clusters would be circular for the controls, so they were only used for assessing the relationship between classification accuracies and symptom severity in patients (an independent group from the control participants).

In the third approach to defining relevant regions, we directly examined the area of face hypoactivation, by isolating the cluster of voxels that showed significantly less face activation in the ASD group than in the control group. We used the univariate approach described above, with the same corrected threshold, to identify the fusiform

hypoactive area, by applying a two-tailed between-groups t-test to the face vs. house coefficients in the intersection of participants' VT regions. A 23-voxel cluster of face hypoactivation was detected in the right fusiform gyrus of the ASD group (center of mass at +41x, -38y, -18z). A 6-voxel cluster was identified in the left fusiform gyrus (center of mass at -49x, -41y, -16z). We believe the area of fusiform hypoactivation is a very suitable area to test the sensitivity of MVPA, given that there is theoretical interest in the hypoactivation, and considering that defining the region based on univariate activation should only help the univariate relationship in a comparison to MVPA. The fusiform regions of interest are shown in Fig. 2.

Although ASD individuals frequently show reduced fusiform face activity, their parahippocampal place area (PPA; Epstein and Kanwisher, 1998) demonstrates robust activity to images of places (Humphreys et al., 2008). We assessed the right and left PPA regions for a relationship between classification performance and symptom severity. We expected that faces and houses would be successfully classified within the PPA, however by examining the relationship with symptom severity, we could determine if any such link extended to other relevant areas of VT cortex. We localized PPA clusters in the right and left hemispheres by first identifying the peak voxel in a house > face activation contrast (in the manner described above for the second fusiform approach) in the right and left medial ventral temporal lobes of each participant. We then centered a 3 × 3 × 3 voxel cluster at each peak voxel, creating a 27-voxel cluster in each hemisphere: a volume similar in size to the fusiform regions.

Finally, to verify that any relationship to symptoms was specific to this latter stage of visual processing, we analyzed two areas in the occipital lobe sensitive to basic visual features: a voxel cluster with the same volume as one of the fusiform regions of interest (ROIs), and a larger area based on the approximate location of Brodmann's area 17 (BA17). We defined the volume-matched cluster by localizing the cluster of 23 voxels (the same volume as the right hypoactive fusiform region) with the greatest response to face and house trials compared



**Fig. 2.** Fusiform regions of interest. Regions for each hemisphere (red) are shown on the total area of all regions (white) at  $x = 39$ ,  $y = -46$ ,  $z = -15$ . The three approaches to defining the regions of interest included: placing spheres at three sets of FFA coordinates from published studies (top), isolating voxel clusters significantly active to faces in the control group (middle), using the clusters of hypoactivation in the ASD group (bottom). Right and left are reversed by convention.

to the fixation period, near to the calcarine sulcus in each participant. This anatomical restriction helped localize the cluster to the typical region of V1. The BA17 region was defined through the AFNI implementation of the Talairach daemon database (Lancaster et al., 2000). This area was included to ensure that we checked for symptom relationships with classification performance in activity patterns that may be distributed across a larger volume.

#### Multi-voxel pattern analysis

All pattern analyses were implemented in MATLAB using custom scripts and the framework provided by the Princeton Multi-Voxel Pattern Analysis toolbox (Dretle et al., 2006). MVPA was conducted on spatially unsmoothed data. All time points were first convolved with a model of the hemodynamic response and thresholded at 0.8 for each slice repetition time (TR), giving eight TRs in each block. The multi-voxel analyses were performed in a 5-fold cross-validation procedure: the data were separated into five across-run folds so that each fold included data from every run and, therefore, examples from all face conditions. A classifier was then trained on four folds and tested on the independent fifth, where the testing set was alternated for each of five iterations. We performed z-scoring within each cross-validation fold to preserve fold independence. To reduce the risk of the classifier model overfitting the training data in the large BA17 region, an ANOVA feature selection method was employed for this area, where voxels are selected that differ significantly between face and house conditions with a liberal threshold of  $p < 0.05$ . Crucially, the ANOVA was only conducted on training data for each fold, therefore ensuring that feature selection was not peaking. This particular selection approach preserves a reasonable number of features and has been used successfully in a number of MVPA studies (e.g., Diana et al., 2008; McDuff et al., 2009; Polyn et al., 2005). This feature selection procedure yielded a mean of 58 voxels (s.d. = 12) for controls and 66 voxels (s.d. = 19) for patients.

A ridge regression classifier was used for the classification procedure; an attractive choice for its ability to compensate for multicollinearity among features (a property of fMRI data). The method has also achieved success in previous machine learning applications (Zhang and Yang, 2003). Classification performance was recorded as the proportion of correct guesses from all iterations of the cross-validation procedure. Ridge regression requires a value to be selected for its penalty parameter. A custom script was developed to select this parameter using an embedded cross-validation technique, where the classifier was trained on part of the training data and tested on the remaining training time-points using a range of penalty values. By restricting the penalty-search process to training data, we ensured that the testing data remained independent. The process was performed for a broad range of penalty values (0, 0.01, 0.1, 1, 10, 100, 1000, and 10,000) and then for a narrower search of ten penalties around the broad-search penalty that gave the highest classification performance. The penalty value giving the highest overall classification performance on the training data was selected for the test set classification.

To compare classification performance of the control and patient groups, we used two-tailed between-group t-tests on classification accuracy. We employed permutation testing (Golland and Fischl, 2003) to assess whether each individual's classification performance was greater than expected by chance, by randomly permuting class labels one thousand times to simulate the null distribution. All permutations included the same number of face and house blocks in each fold to avoid biasing the classifier, and the complete cross-validation process was conducted for each permutation, including the described penalty-search procedure. Classification performance in the top 5% of the random permutations indicated above-chance accuracy (at  $p < 0.05$ ).

We also examined within- and between-category pattern correlations, in a similar manner to Haxby et al. (2001), to explore the basis for the significant MVPA-symptom relationships we report. Specifically, we calculated mean face and house patterns for each of the four runs and then correlated every face and house trial's activity pattern with the other runs' average patterns, in each participant. The resulting correlations were averaged to give a measure of within-category reliability for faces (face trials correlated with mean face patterns) and houses (house trials correlated with mean house patterns), and a measure of between-category similarity (face trials correlated with mean house patterns and house trials correlated with mean face patterns).

#### Searchlight analysis

We employed a spherical searchlight analysis (Kriegeskorte et al., 2006) to explore how the relationship to symptoms varied for the univariate and multivariate results across the VT lobe. Three-dimensional searchlight clusters were first mapped onto the VT area of each participant, creating a series of voxel clusters that covers the VT cortex. For the voxels within each searchlight, activity patterns for the face and house trials were classified with a Gaussian Naive Bayes classifier in a 5-fold cross-validation procedure, with all types of face stimuli in each fold, as described above. Classification performance was allocated to the central voxel of each searchlight, giving a map of accuracies. This analysis was performed three times for each participant, using radii of 2, 3 and 4 voxels, producing clusters with volumes of 33, 123 and 257 voxels respectively, when not restricted by the VT region's boundaries. To directly compare the multivariate searchlights with a univariate approach, we also recorded the mean activation to faces and to houses within each searchlight. We compared the methods' sensitivities to symptom severity by correlating the ASD individuals' multivariate result (classification performance) and then univariate result (mean activation to faces minus mean activation to houses) from each searchlight, with the ADOS social scores, where higher scores indicate a greater number and severity of social symptoms indicative of autism. It is not appropriate to report the maximum correlation values from an extensive analysis such as this (Vul et al., 2009), however it is possible to compare the sensitivities of the univariate and multivariate measures to symptom severity. To achieve this aim, we permuted the clinical scores 10,000 times and then correlated the MVPA and univariate results with the permuted clinical scores. This created a null distribution for testing the significance of the correlation with the actual clinical scores.

## Results and discussion

### MVPA results

Before relating MVPA results to symptom severity, we tested for significant classification performance in the regions of interest. Three approaches were taken to define these regions, as described in the *Methods*, in part because of the difficulty of using a traditional face localizer in a group characterized by face hypoactivation. In the first approach, three overlapping spheres were placed at coordinates from previous FFA studies, giving a 29-voxel cluster in the right fusiform gyrus. Both the typically-developing ( $M = 0.74$ , s.d. = 0.09) and ASD ( $M = 0.78$ , s.d. = 0.07) groups showed high face vs. house classification performance, where chance was 0.50. Permutation tests revealed that classification performance was significantly above chance for all typically-developing and ASD participants ( $p < 0.05$  for all but one control who had a trend at  $p = 0.06$ ). There were no significant group differences in classification performance ( $t_{22} = 1.24$ ,  $p = 0.23$ ).

The second approach employed a 25-voxel cluster in the right fusiform gyrus and a 29-voxel cluster in the left fusiform gyrus,

reflecting significant face activation in the control group. These were not suitable regions of analysis for the control participants, but ASD participants demonstrated high classification performance in the right ( $M = 0.69$ ,  $s.d. = 0.11$ ) and left ( $M = 0.64$ ,  $s.d. = 0.06$ ) regions. Permutation testing revealed significant above-chance performance for all but one ASD individual in the right cluster (ten participants at  $p < 0.02$ , one at  $p = 0.05$ , and one at  $p = 0.38$ ) and all but two in the left cluster (ten participants at  $p < 0.02$ , one at  $p = 0.09$ , and one at  $p = 0.39$ ). The participant with the least significant result in the right cluster ( $p = 0.38$ ) and a trend in the left cluster ( $p = 0.09$ ) corresponded to the left-handed ASD individual in the sample.

The final fusiform region of interest was the fusiform hypoactivation detected in the ASD group through a univariate group comparison: a 23-voxel cluster in the right fusiform gyrus and a 6-voxel cluster in the left fusiform gyrus. The control group showed high classification performance in the right ( $M = 0.71$ ,  $s.d. = 0.08$ ) and left ( $M = 0.71$ ,  $s.d. = 0.08$ ) clusters. The ASD group had lower performance in the right hypoactive cluster ( $M = 0.66$ ,  $s.d. = 0.09$ ) than controls, although not significantly so ( $t_{22} = 1.53$ ,  $p = 0.14$ ). The ASD group's classification performance in the smaller left fusiform cluster ( $M = 0.54$ ,  $s.d. = 0.04$ ) was significantly lower than the control group's ( $t_{22} = 6.51$ ,  $p < 0.001$ ). Permutation testing showed that classification accuracies were significantly above chance in the right and left regions for all controls ( $p < 0.04$ ). In the ASD group, classification performance was significant or approaching significant for the right cluster in all but one participant (nine at  $p < 0.05$ , two at  $p < 0.08$  – including the left-handed ASD participant, one at  $p = 0.18$ ), however for the left hypoactive cluster, ten of the ASD participants' classification accuracies were not significantly above chance (two at  $p < 0.03$ , one at  $p = 0.06$ , and nine at  $p > 0.13$ ). As performance in the left hypoactive region was not significant for the majority of the ASD group, possibly because of the small size of the cluster (6 voxels), we did not analyze this area further.

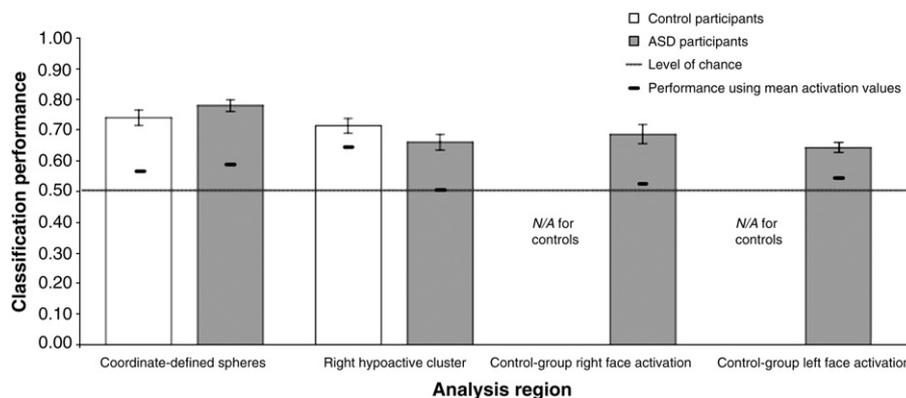
To assess the importance of the voxel *patterns*, we replaced the voxel responses with the regions' mean activation levels at each time point, and repeated the above classifications. This replacement produced a substantial reduction in ASD classification performance for the coordinate-defined spheres ( $M = 0.59$ ,  $s.d. = 0.07$ ; two-tailed paired comparison:  $t_{11} = 9.43$ ,  $p < 0.001$ ), the area of control-group right face activation ( $M = 0.52$ ,  $s.d. = 0.07$ ;  $t_{11} = 6.99$ ,  $p < 0.001$ ), control-group left face activation ( $M = 0.54$ ,  $s.d. = 0.06$ ;  $t_{11} = 6.61$ ,  $p < 0.001$ ), and the right hypoactive cluster ( $M = 0.50$ ,  $s.d. = 0.07$ ;  $t_{11} = 5.56$ ,  $p < 0.001$ ). Control participants also experienced significant reductions in classification performance for the coordinate-defined spheres ( $M = 0.56$ ,  $s.d. = 0.07$ ;  $t_{11} = 5.62$ ,  $p < 0.001$ ) and right hypoactive cluster ( $M = 0.64$ ,  $s.d. = 0.05$ ;  $t_{11} = 3.59$ ,  $p = 0.004$ ). It is noteworthy that in the region of right hypoactivation, using mean

activation values instead of the voxel patterns was particularly detrimental for ASD classification performance: the mean reduction in performance was 0.07 ( $s.d. = 0.07$ ) for controls and 0.16 ( $s.d. = 0.10$ ) for ASD participants, giving a significant group-difference in the size of the decrease ( $t_{10} = 2.83$ ,  $p = 0.02$ ). Using mean activation values here is conceptually similar to a typical univariate analysis, but has an advantage of producing results on the same scale as MVPA. The MVPA and mean-replaced classification results are shown in Fig. 3.

We conducted face vs. house classifications in PPA clusters for each participant, for our subsequent analysis of the relationship between MVPA results and symptom severity. In a 27-voxel cluster centered in the right PPA of each individual, control ( $M = 0.84$ ,  $s.d. = 0.07$ ) and ASD ( $M = 0.87$ ,  $s.d. = 0.05$ ) participants showed high classification performance. Similarly, a 27-voxel cluster in the left PPA gave high classification performance in the control ( $M = 0.82$ ,  $s.d. = 0.08$ ) and ASD ( $M = 0.85$ ,  $s.d. = 0.06$ ) individuals. Permutation testing revealed greater-than-chance accuracy for both regions in every participant ( $p < 0.04$ ). We also conducted the face vs. house classification within two visually-responsive occipital areas, a 23-voxel cluster near the calcarine sulcus of each participant, and a larger approximation of BA17, for subsequently relating performance to symptom severity. In the 23-voxel cluster, classification performance was high in control ( $M = 0.67$ ,  $s.d. = 0.04$ ) and ASD ( $M = 0.69$ ,  $s.d. = 0.06$ ) participants, with no significant difference between the groups ( $t_{22} = 0.90$ ,  $p = 0.38$ ). Permutation testing showed greater-than-chance accuracy in all participants ( $p < 0.05$ ). Similarly for the BA17 region, classification performance was high in control ( $M = 0.77$ ,  $s.d. = 0.04$ ) and ASD ( $M = 0.76$ ,  $s.d. = 0.04$ ) participants, with no significant difference between the groups ( $t_{22} = 0.48$ ,  $p = 0.64$ ). Permutation testing revealed greater-than-chance accuracy in all participants ( $p < 0.002$ ). We had predicted above-chance performance in these regions in advance, because of the visual differences in the stimuli, as described in the introduction. Verifying above-chance classification performance is important for the later link to symptom severity: as the two stimuli classes can be distinguished in these visually-responsive areas, any lack of a relationship with symptom severity in the next stage of the investigation cannot be because there is no relevant information in these regions.

#### Relationship to symptoms

We investigated the sensitivity of MVPA to individual variation in patient symptoms by examining the relationship of classification performance to standardized measures of clinical severity. We also assessed the relationship between symptoms and a univariate measure. Face vs. house classification accuracy was significantly negatively correlated with patients' ADOS total scores for all the right



**Fig. 3.** Classification performance within the fusiform regions of interest for control and ASD participants. The bars reflect classification accuracy using the voxel patterns of each region. The dash-marks within each bar show classification performance when the voxel pattern at each timepoint is replaced by the region's mean activation. Note: it was not appropriate to analyze the controls' data in the control-group face activation regions, for reasons discussed in the text.

fusiform regions (Table 1). These scores are a measure of severity from a structured extended interaction with the patient by an experienced clinical professional. Higher ADOS scores indicate greater severity of symptoms, such that lower classification accuracies were found in more severely affected ASD individuals. Significant negative correlations were also found between the ADOS social component sub-scores and classification accuracies in the right hypoactive and control-group right face activation clusters (Fig. 4). The coordinate-defined spheres relationship approached significance. The social component score of the ADI, an additional assessment of social symptom severity from rated interviews with one of the patients' parents, was significantly related to classification performance in all three right fusiform regions. Performance in the left area of control-group face activation was not significantly related to the clinical measures, although it approached significance for the ADOS total score. The regions' mean activation values (face z-scores–house z-scores) were not significantly correlated with any measure of clinical severity. Table 1 lists the statistical values for these results.

Analyses of the PPA clusters showed that classification performance was not significantly correlated with ADOS total or social scores (Table 1). Classification accuracies were, however, significantly correlated with ADI social scores in the right and left PPA clusters (both at  $p=0.02$ ). Despite this latter significant result, the weak relationship between PPA classification accuracy and symptom severity as measured by the ADOS, suggests a degree of specificity for the fusiform areas. Classification performance within the 23-voxel cluster near the calcarine sulcus, and within the approximate BA17 region, was not significantly correlated with ADOS total or social scores (all  $p>0.7$ ), suggesting the significant relationships in the fusiform regions do not result from differences in basic visual processing. This is also evidence against scanner motion acting as a mediating factor in the significant relationships with symptoms: any

systematically-varying scanner motion would affect other brain areas encoding stimuli differences. Although BA17 classification performance was unexpectedly significantly negatively correlated with ADI social scores ( $p=0.05$ ), the very weak correlations with ADOS scores ( $p=0.78$  and  $p=0.83$ ), and a weak ADI relationship with performance in the 23-voxel occipital cluster ( $p=0.68$ ), give confidence that the strong fusiform and ADOS correlations are not because of early visual processing or motion differences. Motion effects are further ruled-out by the very weak correlations between scanner movement and classification performance in all the fusiform regions (coordinate-defined spheres:  $r=-0.06$ ,  $p=0.84$ ; control-group right face activation:  $r=-0.14$ ,  $p=0.66$ ; control-group left face activation:  $r=0.00$ ,  $p>0.99$ ; right hypoactive cluster:  $r=-0.07$ ,  $p=0.84$ ).

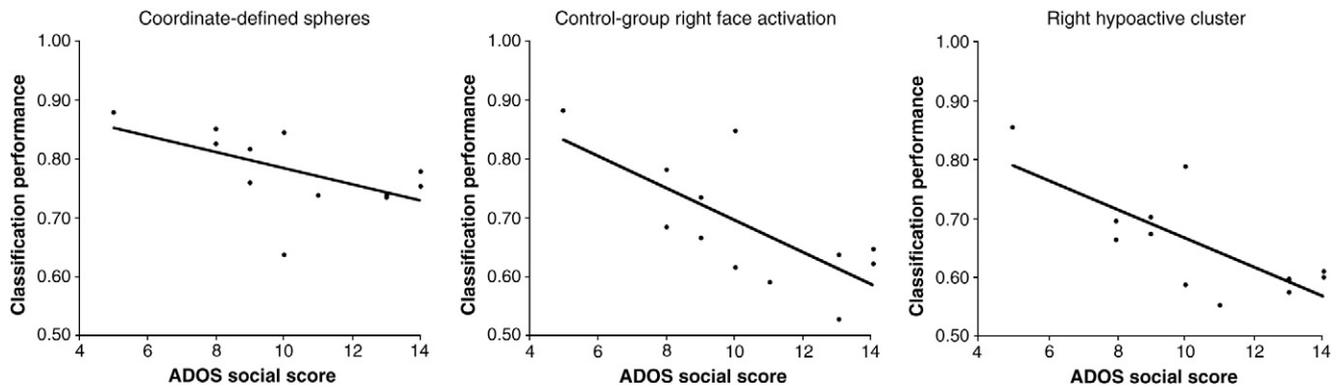
Additionally, neither age (coordinate-defined spheres:  $r=0.04$ ,  $p=0.90$ ; control-group right face activation:  $r=0.13$ ,  $p=0.69$ ; control-group left face activation:  $r=-0.11$ ,  $p=0.74$ ; hypoactive cluster:  $r=0.16$ ,  $p=0.63$ ) nor IQ (coordinate-defined spheres:  $r=0.37$ ,  $p=0.24$ ; control-group right face activation:  $r=0.47$ ,  $p=0.12$ ; control-group left face activation:  $r=0.18$ ,  $p=0.58$ ; hypoactive cluster:  $r=0.42$ ,  $p=0.17$ ) were significantly correlated with performance in the fusiform regions, suggesting these variables were not driving the significant effects. Finally, we examined the signal-to-noise ratio (SNR) to ensure that the relationships between classification performance and symptom severity were not driven by a systematically lower SNR in participants with greater symptom severity. For each ASD individual, we calculated a value for the SNR by dividing the mean baseline (an estimate of the signal) by the standard deviation of the residual time series (an estimate of the noise). There were no significant relationships between symptom severity, measured through the ADOS social scores, and mean SNRs in the VT lobe ( $r=0.03$ ,  $p=0.93$ ), right control-group face activation ( $r=-0.36$ ,  $p=0.25$ ), left control-group face activation

**Table 1**

Correlations between multivariate and univariate results and clinical measures of symptom severity. The right column lists the results of paired two-tailed t-tests for differences between correlations (degrees of freedom = 9).

		MVPA performance		Mean face activation		Correlation difference	
		r	p	r	p	t	p
Coordinate-defined spheres	ADOS total	-0.58	0.05*	0.37	0.23	-2.57	0.03*
	ADOS social score	-0.57	0.055	0.31	0.33	-2.29	0.05*
	ADI social score	-0.59	0.04*	-0.10	0.77	-1.59	0.15
Control-group face activation (right)	ADOS total	-0.70	0.01*	-0.31	0.33	-2.08	0.07
	ADOS social score	-0.72	0.01*	-0.39	0.21	-1.74	0.12
	ADI social score	-0.79	0.002*	-0.50	0.10	-1.72	0.12
Control-group face activation (left)	ADOS total	-0.52	0.08	-0.03	0.92	-3.37	0.008*
	ADOS social score	-0.48	0.11	-0.04	0.91	-2.73	0.02*
	ADI social score	-0.19	0.54	0.18	0.57	-1.99	0.08
Hypoactivation cluster	ADOS total	-0.72	0.008*	-0.30	0.34	-2.30	0.05*
	ADOS social score	-0.75	0.005*	-0.35	0.26	-2.26	0.05*
	ADI social score	-0.77	0.003*	-0.39	0.21	-2.20	0.06
Volume-matched right PPA cluster	ADOS total	-0.38	0.22	-	-	-	-
	ADOS social score	-0.34	0.28	-	-	-	-
	ADI social score	-0.66	0.02*	-	-	-	-
Volume-matched left PPA cluster	ADOS total	-0.46	0.13	-	-	-	-
	ADOS social score	-0.47	0.12	-	-	-	-
	ADI social score	-0.66	0.02*	-	-	-	-
Volume-matched occipital cluster	ADOS total	0.10	0.76	-	-	-	-
	ADOS social score	0.09	0.77	-	-	-	-
	ADI social score	-0.13	0.68	-	-	-	-
Brodmann area 17	ADOS total	-0.09	0.78	-	-	-	-
	ADOS social score	-0.07	0.83	-	-	-	-
	ADI social score	-0.58	0.05*	-	-	-	-

\* statistically significant at  $p<0.05$ .



**Fig. 4.** Scatter plots of face vs. house classification performance against ADOS social scores of the ASD participants. Each plot reflects one of the approaches to defining the fusiform regions: coordinate-defined spheres (left), control-group right face activation (middle) and right hypoactive cluster (right). The y-axis begins at the level of chance.

( $r = -0.46$ ,  $p = 0.13$ ) or right hypoactive cluster ( $r = -0.31$ ,  $p = 0.33$ ). The SNR in the coordinate-defined spheres was close to being significantly related to symptom severity ( $r = -0.57$ ,  $p = 0.055$ ), however the weak relationships for the other regions suggest that systematic differences in SNR cannot account for the MVPA-symptom severity relationships.

Scores on the Benton face recognition task, where higher scores indicate greater face recognition ability, were not significantly related to classification performance in the coordinate-defined spheres (ASD:  $r = 0.29$ ,  $p = 0.35$ ; controls:  $r = -0.32$ ,  $p = 0.31$ ) or the left cluster of control-group face activation (ASD:  $r = 0.46$ ,  $p = 0.13$ ), but approached significance in the right cluster of control-group face activation (ASD:  $r = 0.54$ ,  $p = 0.07$ ) and right hypoactive cluster (ASD:  $r = 0.56$ ,  $p = 0.06$ ; controls:  $r = -0.28$ ,  $p = 0.37$ ) for ASD participants.

Behavioral performance for the in-scan 'same vs. different' task with neutral faces was not significantly correlated with classification performance in the fusiform regions for ASD participants (coordinate-defined spheres:  $r = 0.10$ ,  $p = 0.76$ ; control-group right face activation:  $r = -0.23$ ,  $p = 0.47$ ; control-group left face activation:  $r = -0.33$ ,  $p = 0.30$ ; right hypoactive cluster:  $r = -0.15$ ,  $p = 0.63$ ) or controls (coordinate-defined spheres:  $r = -0.27$ ,  $p = 0.40$ ; right hypoactive cluster:  $r = 0.05$ ,  $p = 0.88$ ). It is possible that these weak correlations are due to behavioral performance approaching ceiling, although controls (face  $M = 0.88$ ,  $s.d. = 0.07$ ; house  $M = 0.96$ ,  $s.d. = 0.03$ ) showed greater task accuracy than ASD participants (face  $M = 0.80$ ,  $s.d. = 0.09$ ; house  $M = 0.92$ ,  $s.d. = 0.05$ ) for faces ( $t_{22} = 2.34$ ,  $p = 0.03$ ) and houses ( $t_{22} = 2.66$ ,  $p = 0.01$ ). It is also possible that the influence of the passive viewing task activation on the classification results dilutes a link between behavioral and classifier performance.

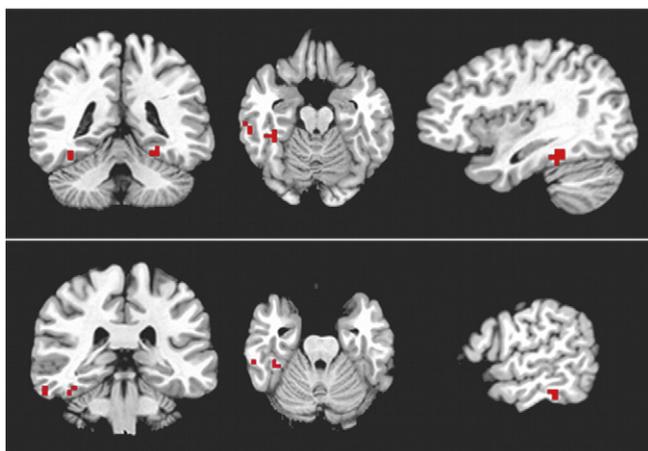
We also explored whether lower classification accuracies in more severely-affected participants result from greater variability in their multi-voxel face patterns, or from their face and house patterns being less discriminable (more positively correlated). To examine this, we performed within- and between-category correlation analyses in the fusiform regions with MVPA-symptom relationships. The right hypoactive region's face patterns were significantly or close-to-significantly less correlated with average face patterns in individuals with increased symptom severity (ADOS social:  $r = -0.63$ ,  $p = 0.03$ ; ADI social:  $r = -0.64$ ,  $p = 0.03$ ; ADOS total:  $r = -0.54$ ,  $p = 0.07$ ). There were no significant relationships between house correlations and ADOS total or social scores, with just a trend for ADI social scores ( $r = -0.55$ ,  $p = 0.07$ ). The between-category correlations were more positive (reflecting less discriminable patterns) in individuals with higher ADI social scores ( $r = 0.71$ ,  $p = 0.01$ ;  $r = 0.69$ ,  $p = 0.01$ ), but this did not reach significance for the ADOS total or social scores. In the right cluster of control-group face activation, only the ADI scores gave significant relationships: increased symptom severity was associated with less correlated within-category face patterns

( $r = -0.63$ ,  $p = 0.03$ ) and higher between-category correlations ( $r = 0.78$ ,  $p = 0.003$ ;  $r = 0.73$ ,  $p = 0.01$ ), with a weak trend for house patterns ( $r = -0.52$ ,  $p = 0.09$ ). The coordinate-defined spheres had no significant symptom relationships for within- or between-category correlations, excepting between house correlations and ADOS total scores ( $r = -0.59$ ,  $p = 0.04$ ), although this was not significant for the ADOS or ADI social scores.

#### Searchlight analysis

We examined regional variation of multivariate and univariate sensitivities to symptom severity in the VT lobes using a spherical searchlight analysis. We ran the searchlight technique with multivariate (face vs. house classification accuracy) and univariate (mean activation to faces minus mean activation to houses) measures for each ASD participant. The recorded values for each sphere were correlated with the ADOS social scores. The searchlight procedure was conducted using a radius of 2, 3 and then 4 voxels to examine how the results vary with searchlight size. The strength of the relationships between the searchlight measures and symptom severity was greater overall for the multivariate measure than for mean activation, indicated by two-tailed paired t-tests on the absolute correlation coefficients of searchlights with a 2-voxel, ( $t_{970} = 3.96$ ,  $p < 0.001$ ), 3-voxel ( $t_{970} = 3.06$ ,  $p = 0.002$ ) and 4-voxel radius ( $t_{970} = 2.47$ ,  $p = 0.01$ ).

We also examined which searchlights were significantly related to symptom severity by permuting the participants' ADOS social scores 10,000 times and computing the searchlights' correlations for each permutation. Comparing the correlation from using the correct ADOS scores gave a map of  $p$ -values for each searchlight size, for the univariate and multivariate measures. No clusters of more than one searchlight were significantly correlated with ADOS social scores when mean activation was employed as the searchlight dependent variable, for any of the three radii (at a liberal threshold of  $p < 0.01$ ). In contrast, when classification performance was employed as the searchlight measure, using a 2-voxel radius detected three searchlight clusters (of at least 2 contiguous central voxels) that were significantly related to symptom severity: a 10-voxel cluster centered in the right fusiform gyrus (center of mass:  $x = 39$ ,  $y = -36$ ,  $z = -18$ ), a 4-voxel cluster centered in the left parahippocampal gyrus (center of mass:  $x = -25$ ,  $y = -39$ ,  $z = -13$ ) and a 3-voxel cluster centered in the right inferior temporal gyrus (center of mass:  $x = 58$ ,  $y = -31$ ,  $z = -19$ ). These results are shown in Fig. 5. Using a 3-voxel searchlight radius revealed a 4-voxel cluster in the right fusiform gyrus (center of mass:  $x = 41$ ,  $y = -35$ ,  $z = -20$ ) that partially overlapped with the 10-voxel cluster reported in the 2-voxel radius analysis. No significant searchlights were detected with a 4-voxel radius. We also applied a cluster-based correction for multiple



**Fig. 5.** Searchlights with significant correlations between face vs. house classification performance and ADOS social scores. Each red voxel represents the center of one searchlight with a radius of 2 voxels. Significance ( $p < 0.01$ ) was determined by permuting the clinical scores. Top row:  $x = 36$ ,  $y = -40$ ,  $z = -17$ . Bottom row:  $x = 57$ ,  $y = -31$ ,  $z = -21$ . Right and left are reversed by convention.

comparisons, although this did not produce significant results for any of the searchlight radii. Despite this null finding, the detection of these regions at a liberal threshold, using MVPA results but not the univariate measure, suggests the presence of stronger relationships between symptoms and the MVPA results, than the univariate measure employed here.

## Discussion

### Review of aims

The primary aim of this study was to examine if MVPA measures can be sensitive to patient symptom severity. We found that classification performance, a multivariate measure of separability for the face and house fMRI patterns, was strongly related to standard measures of clinical severity in ASD participants. Specifically, in both anatomically and functionally defined clusters of right fusiform voxels, classification performance was significantly negatively correlated with symptom severity, while mean activation levels were not. This greater sensitivity of pattern analyses extended to voxels that were defined using differences in univariate measures. Assessments of the PPA region showed that this sensitivity is not a general property of VT cortex. Analyzing two occipital areas additionally confirmed that the finding did not generalize to activity patterns involved in early visual processing. A searchlight analysis across the ventral temporal lobes detected regions where classification performance was significantly related to symptom severity, which were not detected using the searchlights' mean activation levels, although only when a liberal threshold was employed.

We have provided an example of obtaining these benefits from a functional dataset not designed with MVPA in mind. By combining multiple face conditions into one face class, we were able to utilize a large number of trials for classifier training and testing. Although the design of the experiment placed limitations on the conclusions that can be drawn from the results (discussed below), multivariate classification performance was still more sensitively related to symptom severity than the univariate measures we employed.

### MVPA and patient groups

The findings in this study provide, to the best of our knowledge, the first evidence that MVPA functional results can reliably predict clinical symptoms. The sensitivity to clinical severity obtained using MVPA supports the idea that subtle variations in activity patterns, reflected in MVPA results, can in some cases more sensitively reflect individual variation in an area's functional characteristics, than certain measures of

mean activation. It is noteworthy that this stronger link to symptom severity was also found in a set of voxels that was defined based on a univariate statistic (i.e., a significant difference in activation values between groups). This latter finding gives confidence that the greater sensitivity does not reflect the activity patterns of voxels that are separate from those demonstrating univariate differences.

The greater sensitivity to individual differences reported here will be of interest to researchers who are involved in characterizing variation across participants in a wide range of fields. Among clinical investigators, this interest may even extend to those looking to select individuals for future interventions. As Scherf et al. (2010) noted when discussing the failure to find a significant relationship between (univariate) fusiform gyrus activation and ADOS scores, "such predictability could have substantial implications for identifying individuals who might benefit from a behavioral intervention designed to improve face processing" (p.13). Future research will be required to establish if the sensitivity reported here extends to other patient groups, and to other regions with reduced univariate activation.

The findings reported in this study are also relevant to clinical researchers looking to make the most of existing functional datasets. The detection of regions with patterns of activity that reflect variations in patient symptoms, without a corresponding significant univariate relationship, suggests the encouraging possibility that additional regions of interest may be identifiable in previously-collected datasets. In this dataset, we found that symptom severity was related to face vs. house classification performance in a coordinate-defined area (based on face activity coordinates in prior literature), which neighbors hypoactivation in this particular group of ASD participants. A significant relationship here suggests that areas of the fusiform gyrus without a significant difference in univariate activation, may nevertheless show activity patterns that vary systematically with symptom severity. This may be expected from an activity pattern perspective, where a significant group difference in univariate activation can be conceptualized as two very distinct activity patterns. Variations in face perception-related activity may still be present in nearby areas of cortex, even if undetectable with univariate techniques.

Linking multivariate searchlight results to individual differences gives further potential for revealing new regions of interest. In the context of this dataset, our searchlight finding of a symptom relationship in the inferior temporal gyrus (ITG) fits with several previous studies that have suggested the region may play a role in face processing in this patient group (e.g., Koshino et al., 2008; Schultz et al., 2000). Some univariate studies have not detected ITG involvement (e.g., Pierce et al., 2001), giving the possibility that systematic differences in ITG activity – differences that may not always be detectable with univariate analyses – could have been present, undetected, in the functional data of such studies. Although the identified searchlight locations have backing from prior literature, the failure to detect these regions at a more stringent threshold means these results should be interpreted with some caution. Despite this caveat, the multivariate searchlight approach has the potential to highlight new regions that are functionally related to patient symptoms.

Our examination of within- and between-category correlations may be of interest to investigators exploring the basis for MVPA-symptom relationships. In this dataset, the within-participant consistency of multi-voxel face patterns was lower in individuals with greater symptom severity in the hypoactive fusiform region. This may be an underlying factor in the MVPA-symptom relationships, although our finding that greater symptom severity, when measured by the ADI, is accompanied by less discriminable face and house patterns is also suggestive. Future studies may wish to examine further the relative contributions of face pattern consistency and face/non-face discriminability, including whether the variations in multi-voxel face patterns reflect larger differences across face types (which varied by run here), or within face types.

Although in this particular context we found multivariate measures to be a strong predictor of symptoms, there may be other contexts and questions where univariate measures are more sensitive. In a recent MVPA study, Quamme et al. (2010) reported that univariate measures were more sensitive than MVPA to task behavior, at the group level, in several of the regions they examined. As Quamme et al. (2010) described, the complexity of MVPA is accompanied by a vulnerability to over-fitting noise, which is less likely to occur with across-voxel averages. It is therefore very possible that univariate measures could provide a more sensitive measure of individual differences than MVPA in some circumstances. For this reason, it should not be concluded from this paper that MVPA will always be a more sensitive measure for tracking functionally-relevant individual differences, but that in certain circumstances it can be. We further note that a variety of univariate measures are available for fMRI analyses. Although we found that MVPA results were a stronger predictor of symptom severity than the univariate measure we employed (mean face activation–mean house activation), this may not apply to all univariate measures. This is additionally relevant as the MVPA and univariate measures employed here differ in the number of free parameters. Overall, we view the univariate and multivariate approaches as complementary, with each adding its own value.

Despite the findings we report here, it must be acknowledged that the design of the original fMRI study places a limit on interpretations. Specifically, the inclusion of two stimulus categories, faces and houses, limits the condition-specific conclusions that can be drawn, as discussed in the Introduction. Future studies of neural differences in face processing may consider including additional stimulus classes, such as scrambled images and object categories (as in Haxby et al., 2001; Spiridon and Kanwisher, 2002). Assessing classifications of faces vs. non-faces, alongside the classification of different non-face categories, would further investigations of face-specific activity patterns in ASD individuals. Employing alternative processing and classification methods may also contribute additional insights (see O'Toole et al., 2007), for a discussion of different classification approaches). Another approach for future research would be to perform MVPA within anatomically-defined ROIs. For example, ASD symptom severity should be correlated with face-related classification performance in the fusiform, but not parahippocampal, gyrus. As PPA activity can sometimes extend into the fusiform gyrus (Epstein and Kanwisher, 1998), it may be desirable to restrict such anatomically-defined fusiform ROIs to certain sub-sections of the gyrus for face vs. house classifications, or expand the non-face classes as discussed above. Future studies may also wish to use MVPA to study activity pattern variations for different face identities in individuals with an ASD. Investigating the nature of the multi-voxel patterns generated by different faces in this group could advance our understanding of their face processing differences.

### Conclusions

We have shown that MVPA can act as a sensitive fMRI predictor of patient symptoms. We believe this study highlights an important use of MVPA techniques for the study of autism and other clinical conditions. The application of pattern analysis techniques to patient differences is still in its infancy, but this investigation shows that the approach has the potential to measure clinically relevant patterns. Furthermore, MVPA combined with mapping techniques can identify brain regions that may not be revealed with certain univariate approaches.

### Acknowledgments

This work was supported by NIMH grant R01MH073084 (R.T. Schultz), with further support from NIH grant R01MH070850 (S.L. Thompson-Schill). We thank members of the Thompson-Schill lab and CHOP Center for Autism Research for helpful discussions. We

thank Lauren Hallion for valuable comments on an earlier version of the manuscript, and the anonymous reviewers for their insightful suggestions.

### References

- Benton, A.L., Sivan, A.B., Hamsher, K., Varney, N.R., Spreen, O., 1994. Contributions to Neuropsychological Assessment. Oxford University Press, New York.
- 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. *J. Int. Neuropsychol. Soc.* 14 (6), 922–932.
- Boucher, J., Lewis, V., 1992. Unfamiliar face recognition in relatively able autistic children. *J. Child Psychol. Psychiatry* 33, 843–859.
- Cohen, J.D., MacWhinney, B., Flatt, M., Provost, J., 1993. PsyScope: a new graphic interactive environment for designing psychology experiments. *Behav. Res. Methods* 25, 257–271.
- Costafreda, S.G., Khanna, A., Mourao-Miranda, J., Fu, C.H.Y., 2009. Neural correlates of sad faces predict clinical remission to cognitive behavioural therapy in depression. *Neuroreport* 20 (7), 637–641.
- Cox, R.W., 1996. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Comput. Biomed. Res.* 29, 162–173.
- Cox, R.W., Jesmanowicz, A., 1999. Real-time 3D image registration for functional MRI. *Magn. Reson. Med.* 42, 1014–1018.
- Cox, D.D., Savoy, R.L., 2003. Functional magnetic resonance imaging (fMRI) “brain reading”: detecting and classifying distributed patterns of fMRI activity in human visual cortex. *Neuroimage* 19 (2 Pt 1), 261–270.
- 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., et al., 2000. The functional neuroanatomy of social behaviour: changes in cerebral blood flow when people with autistic disorder process facial expressions. *Brain* 123, 2203–2212.
- Dailey, M.N., Cottrell, G.W., Reilly, J., 2001. California Facial Expressions (CAFE). Computer Science and Engineering Department, UCSD, La Jolla, CA.
- Dalton, K.M., Nacewicz, B.M., Johnstone, T., Schaefer, H.S., Gernsbacher, M.A., Goldsmith, H.H., Alexander, A.L., et al., 2005. Gaze fixation and the neural circuitry of face processing in autism. *Nat. Neurosci.* 8 (4), 519–526.
- Deeley, Q., Daly, E.M., Surguladze, S., Page, L., Toal, F., Robertson, D., Curran, S., Giampietro, V., Seal, M., Brammer, M.J., et al., 2007. An event related functional magnetic resonance imaging study of facial emotion processing in Asperger syndrome. *Biol. Psychiatry* 62, 207–217.
- Detre, G.J., Polyn, S.M., Moore, C.D., Natu, V.S., Singer, B.D., Cohen, J.D., Haxby, J.V., Norman, K.A., 2006. The Multi-Voxel Pattern Analysis (MVPA) toolbox. Poster presented at the Annual Meeting of the Organization for Human Brain Mapping (Florence, Italy).
- Diana, R.A., Yonelinas, A.P., Ranganath, C., 2008. High-resolution multi-voxel pattern analysis of category selectivity in the medial temporal lobes. *Hippocampus* 18, 536–541.
- Ekman, P., Friesen, W.V., 1976. Pictures of Facial Affect. Consulting Psychologists Press, Palo Alto, CA.
- Endl, W., Walla, P., Lindinger, G., Laluschek, W., Barth, F.G., Deecke, L., Lang, W., 1998. Early cortical activation indicates preparation for retrieval of memory for faces: an event-related potential study. *Neurosci. Lett.* 240, 58–60.
- Epstein, R., Kanwisher, N., 1998. A cortical representation of the local visual environment. *Nature* 392, 598–601.
- Fu, C.H.Y., Mourao-Miranda, J., Costafreda, S.G., Khanna, A., Marquand, A.F., Williams, S.C.R., Brammer, M.J., 2008. Pattern classification of sad facial processing: toward the development of neurobiological markers in depression. *Biol. Psychiatry* 63 (7), 656–662.
- Gadow, K.D., Sprafkin, J., 1994. Child Symptom Inventories Manual. Checkmate Plus, Stony Brook, NY.
- Gadow, K.D., Sprafkin, J., Weiss, M.D., 1999. Adult Self-Report Inventory-4 Manual. Stony Brook. Checkmate Plus, NY.
- Gadow, K.D., Sprafkin, J., Carlson, G.A., Schneider, J., Nolan, E.E., Mattison, R.E., Rundberg-Rivera, V.A., 2002. DSM-IV-referenced, adolescent self-report rating scale. *J. Am. Acad. Child Adolesc. Psychiatry* 41 (6), 671–679.
- Gauthier, I., Tarr, M.J., Anderson, A.W., Skudlarski, P., Gore, J.C., 1999. Activation of the middle fusiform ‘face area’ increases with expertise in recognizing novel objects. *Nat. Neurosci.* 2, 568–573.
- Gilbert, S.J., Meuwese, J.D.I., Towgood, K.J., Frith, C.D., Burgess, P.W., 2009. Abnormal functional specialization within medial prefrontal cortex in high-functioning autism: a multi-voxel similarity analysis. *Brain* 132, 869–878.
- Golland, P., Fischl, B., 2003. Permutation tests for classification: towards statistical significance in image-based studies. *Inf. Process. Med. Imaging* 330–341.
- Grelotti, D., Klin, A., Gauthier, I., Skudlarski, P., Cohen, D., Gore, J., Volkmar, F., et al., 2005. fMRI activation of the fusiform gyrus and amygdala to cartoon characters but not to faces in a boy with autism. *Neuropsychologia* 43 (3), 373–385.
- Gur, R.C., Sara, R., Hagendoorn, M., Marom, O., Hughett, O., Macy, L., Turner, T., Bajcsy, R., Posner, A., Gur, R.E., 2002. A method for obtaining 3-dimensional facial expressions and its standardization for use in neurocognitive studies. *J. Neurosci. Methods* 115, 137–143.
- Hadjikhani, N., Joseph, R.M., Snyder, J., Chabris, C.F., Clark, J., Steele, S., McGrath, L., Vangel, M., Aharon, I., Feczko, E., et al., 2004. Activation of the fusiform gyrus when individuals with autism spectrum disorder view faces. *Neuroimage* 22, 1141–1150.
- Hadjikhani, N., Joseph, R.M., Snyder, J., Tager-Flusberg, H., 2007. Abnormal activation of the social brain during face perception in autism. *Hum. Brain Mapp.* 28, 441–449.

- Hall, G.B.C., Szechtman, H., Nahmias, C., 2003. Enhanced salience and emotion recognition in Autism: a PET study. *Am. J. Psychiatry* 160, 1439–1441.
- Haxby, J.V., Gobbini, M.I., Furey, M.L., Ishai, A., Schouten, J.L., Pietrini, P., 2001. Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* 293, 2425–2430.
- Haynes, J., Rees, G., 2006. Decoding mental states from brain activity in humans. *Nat. Rev. Neurosci.* 7 (7), 523–534.
- 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 Res.* 1, 52–63.
- Kamitani, Y., Tong, F., 2005. Decoding the visual and subjective contents of the human brain. *Nat. Neurosci.* 8, 679–685.
- Kanwisher, N., Yovel, G., 2006. The fusiform face area: a cortical region specialized for the perception of faces. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 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. *J. Neurosci.* 17, 4302–4311.
- Kim, M.A., Tura, E., Potkin, S.G., Fallon, J.H., Manoach, D.S., Calhoun, V.D., Turner, J.A., 2010. Working memory circuitry in schizophrenia shows widespread cortical inefficiency and compensation. *Schizophr. Res.* 117, 42–51.
- Klin, A., 2008. Three things to remember if you are a functional magnetic resonance imaging researcher of face processing in autism spectrum disorders. *Biol. Psychiatry* 64 (7), 549–551.
- Koshino, H., Kana, R.K., Keller, T.A., Cherkassky, V.L., Minshew, N.J., Just, M.A., 2008. fMRI investigation of working memory for faces in autism: visual coding and underconnectivity with frontal areas. *Cereb. Cortex* 18, 289–300.
- Kriegeskorte, N., Goebel, R., Bandettini, P., 2006. Information-based functional brain mapping. *Proc. Natl. Acad. Sci. U. S. A.* 103, 3863–3868.
- Lancaster, J.L., Woldorff, M.G., Parsons, L.M., Liotti, M., Freitas, C.S., Rainey, L., Kochunov, P.V., Nickerson, D., Mikiten, S.A., Fox, P.T., 2000. Automated Talairach atlas labels for functional brain mapping. *Hum. Brain Mapp.* 10, 120–131.
- Le Couteur, A., Rutter, M., Lord, C., Rios, P., Robertson, S., Holdgrafer, M., McLennan, J., 1989. Autism diagnostic interview: a standardized investigator-based instrument. *J. Autism Dev. Disord.* 19, 363–387.
- Lord, C., Rutter, M., Goode, S., Heemsbergen, J., Jordan, H., Mawhood, L., Schopler, E., 1989. Autism diagnostic observation schedule: a standardized observation of communicative and social behavior. *J. Autism Dev. Disord.* 19, 185–212.
- Lord, C., Rutter, M., Couteur, A., 1994. Autism Diagnostic Interview-Revised: a revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *J. Autism Dev. Disord.* 24 (5), 659–685.
- Lord, C., Risi, S., Lambrecht, L., Cook, E.H., Leventhal, B.L., DiLavore, P.C., Pickles, A., Rutter, M., 2000. The autism diagnostic observation schedule-generic: a standard measure of social and communication deficits associated with the spectrum of autism. *J. Autism Dev. Disord.* 30, 205–223.
- Lundqvist, D., Litton, J.E., 1998. The Averaged Karolinska Directed Emotional Faces – AKDEF. AKDEF CDROM. Psychology section, Karolinska Institute, Stockholm, USA.
- Matsumoto, D., Ekman, P., 1988. Japanese and Caucasian Facial Expressions of Emotion (JACFEE) [Slides]. Intercultural and Emotion Research Laboratory, Department of Psychology, San Francisco State University, San Francisco, CA.
- McDuff, S.G.R., Frankel, H.C., Norman, K.A., 2009. Multivoxel pattern analysis reveals increased memory targeting and reduced use of retrieved details during single-agenda source monitoring. *J. Neurosci.* 29, 508–516.
- Norman, K.A., Polyn, S.M., Detre, G.J., Haxby, J.V., 2006. Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends Cogn. Sci.* 10, 424–430.
- Oldfield, R.C., 1971. The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113.
- O’Toole, A.J., Jiang, F., Abdi, H., Haxby, J.V., 2005. Partially distributed representations of objects and faces in ventral temporal cortex. *J. Cogn. Neurosci.* 17, 580–590.
- O’Toole, A.J., Jiang, F., Abdi, H., Pénard, N., Dunlop, J.P., Parent, M.A., 2007. Theoretical, statistical, and practical perspectives on pattern-based classification approaches to the analysis of functional neuroimaging data. *J. Cogn. Neurosci.* 19, 1735–1752.
- Pereira, F., Mitchell, T., Botvinick, M., 2009. Machine learning classifiers and fMRI: a tutorial overview. *Neuroimage* 45 (Suppl. 1), S199–S209.
- Perlman, S.B., Hudac, C.M., Pegors, T., Minshew, N.J., Pelphrey, K.A., 2010. Experimental manipulation of face-evoked activity in the fusiform gyrus of individuals with autism. *Soc. Neurosci.* 1–9.
- Pierce, K., Redcay, E., 2008. Fusiform function in children with an autism spectrum disorder is a matter of “who”. *Biol. Psychiatry* 64 (7), 552–560.
- 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 (12), 2703–2716.
- Piggot, J., Kwon, H., Mobbs, D., Blasey, C., Lotspeich, L., Menon, V., Bookheimer, S., Reiss, A.L., 2004. Emotional attribution in high-functioning individuals with autistic spectrum disorder: a functional imaging study. *J. Am. Acad. Child Adolesc. Psychiatry* 43, 473–480.
- Polyn, S.M., Natu, V.S., Cohen, J.D., Norman, K.A., 2005. Category-specific cortical activity precedes retrieval during memory search. *Science* 310, 1963–1966.
- Quamme, J.R., Weiss, D.J., Norman, K.A., 2010. Listening for recollection: a multi-voxel pattern analysis of recognition memory retrieval strategies. *Front. Hum. Neurosci.* 4, 61.
- Rizk-Jackson, A., Stoffers, D., Sheldon, S., Kuperman, J., Dale, A., Goldstein, J., Corey-Bloom, J., et al., 2010. Evaluating imaging biomarkers for neurodegeneration in pre-symptomatic Huntington’s disease using machine learning techniques. *Neuroimage*.
- 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. *Front. Hum. Neurosci.* 4, 26.
- 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. *Arch. Gen. Psychiatry* 57, 331–340.
- Schultz, R.T., Grelotti, D.J., Klin, A., Kleinman, J., Van der Gaag, C., Marois, R., Skudlarski, P., 2003. The role of the fusiform face area in social cognition: implications for the pathobiology of autism. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 358 (1430), 415–427.
- Schultz, R.T., Grupe, D.W., Hunyadi, E., Jones, W., Wolf, J., Hoyt, E.G., Lin, D., Herlihy, L.E., 2008. The effect of task differences on FFA activity in autism spectrum disorders. Paper presented at the International Meeting for Autism Research, London, England, May.
- Spiridon, M., Kanwisher, N., 2002. How distributed is visual category information in human occipito-temporal cortex? An fMRI study. *Neuron* 35, 1157–1165.
- Talairach, J., Tournoux, P., 1988. Co-Planar Stereotaxic Atlas of the Human Brain. Thieme, Stuttgart, Germany.
- Tottenham, N., Borscheid, A., Ellertsen, K., Marcu, D.J., Nelson, C.A., 2002. Categorization of facial expressions in children and adults: establishing a larger stimulus set. *J. Cogn. Neurosci.* 14, S74.
- Vul, E., Harris, C., Winkielman, P., Pashler, H., 2009. Puzzlingly high correlations in fMRI studies of emotion, personality, and social cognition. *Perspect. Psychol. Sci.* 4 (3), 274–290.
- Wang, A.T., Dapretto, M., Hariri, A.R., Sigman, M., Bookheimer, S.Y., 2004. Neural correlates of facial affect processing in children and adolescents with autism spectrum disorder. *J. Am. Acad. Child Adolesc. Psychiatry* 43, 481–490.
- 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.
- Wolf, J.M., Tanaka, J.W., Klaiman, C., Cockburn, J., Herlihy, L., Brown, C., South, M., McPartland, J., Kaiser, M.D., Phillips, R., et al., 2008. Specific impairment of face-processing abilities in children with autism spectrum disorder using the *Let’s Face It!* skills battery. *Autism Res.* 1, 329–340.
- Yoon, J.H., Tamir, D., Minzenberg, M.J., Ragland, J.D., Ursu, S., Carter, C.S., 2008. Multivariate pattern analysis of functional magnetic resonance imaging data reveals deficits in distributed representations in schizophrenia. *Biol. Psychiatry* 64, 1035–1041.
- Zhang, J., Yang, Y., 2003. Robustness of regularized linear classification methods in text categorization. Proceedings of SIGIR-2003, 26st ACM International Conference on Research and Development in Information Retrieval, pp. 190–197.
- Zhang, L., Samaras, D., Tomasi, D., Alia-Klein, N., Cottone, L., Leskova, A., Volkow, N., et al., 2005. Exploiting temporal information in functional magnetic resonance imaging brain data. *Medical Image Computing and Computer-Assisted Intervention: MICCAI... International Conference on Medical Image Computing and Computer-Assisted Intervention*, 8(Pt 1), pp. 679–687.