BRAIN
A JOURNAL OF NEUROLOGY

Abnormal functional specialization within medial prefrontal cortex in high-functioning autism: a multi-voxel similarity analysis

Sam J. Gilbert,[1] Julia D.I. Meuwese,[1] Karren J. Towgood,[1] Christopher D. Frith[2,3] and Paul W. Burgess[1]

1 Institute of Cognitive Neuroscience and Department of Psychology, University College London, London, UK
2 Wellcome Department of Imaging Neuroscience, Institute of Neurology, University College London, London, UK
3 Center for Functional Integrative Neuroscience, University of Aarhus, Denmark

Correspondence to: Dr Sam Gilbert,
Institute of Cognitive Neuroscience,
17 Queen Square,
London WC1N 3AR, UK
E-mail: sam.gilbert@ucl.ac.uk

Multi-voxel pattern analyses have proved successful in 'decoding' mental states from fMRI data, but have not been used to examine brain differences associated with atypical populations. We investigated a group of 16 (14 males) high-functioning participants with autism spectrum disorder (ASD) and 16 non-autistic control participants (12 males) performing two tasks (spatial/verbal) previously shown to activate medial rostral prefrontal cortex (mrPFC). Each task manipulated: (i) attention towards perceptual versus self-generated information and (ii) reflection on another person's mental state ('mentalizing' versus 'non-mentalizing') in a 2×2 design. Behavioral performance and group-level fMRI results were similar between groups. However, multi-voxel similarity analyses revealed strong differences. In control participants, the spatial distribution of activity generalized significantly between task contexts (spatial/verbal) when examining the same function (attention/mentalizing) but not when comparing different functions. This pattern was disrupted in the ASD group, indicating abnormal functional specialization within mrPFC, and demonstrating the applicability of multi-voxel pattern analysis to investigations of atypical populations.

Introduction

Functional magnetic resonance imaging (fMRI) can potentially contribute at least three types of evidence to shed light on brain differences associated with atypical populations: (i) evidence for under- or over-activation of specialized brain regions in response to experimental manipulations; (ii) evidence for atypical interactions between two or more brain regions; or (iii) evidence for atypical organization of the brain into functionally distinct regions.

Recent years have seen an increase in studies using fMRI to investigate autism spectrum disorder (ASD), a neurodevelopmental condition affecting ∼1% of the population (Baird et al., 2006), characterized by impaired communication and social interaction, along with repetitive behaviors and restricted interests. In this population, evidence has accumulated for both atypical activation of specific brain regions (Castelli et al., 2002; Müller et al., 2004; Schmitz et al., 2006) and atypical interactions between distinct brain regions (Bird et al., 2006; Kana et al., 2006; Just et al., 2007). However, although some studies have investigated functional organization of the brain in ASD (Pierce et al., 2001; Hadjikhani et al., 2004; Chiu et al., 2008; Kleinhans et al., 2008), this has been a less common approach. This is perhaps surprising because ASD, as a developmental disorder, is likely to affect the processes by which subregions of the brain become specialized for different functions during development. Moreover, behavioral studies indicate that ASD is unlikely to involve just a single primary processing deficit (Minshew et al., 1997), suggesting that it is inadequate to hypothesize that ASD involves disruption to any one brain system in the context of otherwise typical brain organization. In this study, we demonstrate that multi-voxel pattern analysis of fMRI data can be used to provide direct neuroimaging evidence for atypical functional organization of the brain in ASD.

Neuroimaging studies investigating functional specialization in the brain have used two broad approaches. The first, most common approach is to investigate two or more tasks (along with baseline conditions), presumed to depend on separable cognitive processes, and show that they are associated with spatially distinct activation peaks, after averaging across a group of participants. For instance, Gilbert et al. (2007) investigated two functions that have been associated with activity in mrPFC: (i) mentalizing, i.e. reflecting on one's own mental states or those of other people (Frith and Frith, 2003, 2006) and (ii) stimulus-oriented (SO) versus stimulus-independent (SI) attention, i.e. performing a task whilst attending to task-relevant perceptual information versus doing the same task 'in one's head' (Gilbert et al., 2005; Burgess et al., 2007). An example of such a task would be navigating around a visually presented shape (SO condition), versus imagining the same shape and continuing to navigate around it (SI condition; Gilbert et al., 2005, 2007). At the group level, activation peaks in mrPFC relating to mentalizing were significantly posterior and superior to those associated with attention (SO versus SI).

Similar techniques have been used to establish abnormal functional brain organization in ASD. Gilbert et al. (2008) investigated a task alternating between SO and SI conditions in a group of high-functioning participants with ASD, along with a control group. Activation peaks within mrPFC in participants with ASD were significantly posterior to those in the control group. Thus, these studies have established (i) a considerable degree of functional specialization within mrPFC (consistent with similar results from a meta-analysis; Gilbert et al., 2006), and (ii) evidence for abnormal functional specialization in ASD.

The second principal technique for investigating functional specialization using fMRI involves analysing data from each participant individually, to investigate the distribution of activation elicited by particular stimuli or task demands across a set of voxels (Norman et al., 2006; O'Toole et al., 2007). For instance, Haxby et al. (2001) showed that by examining the pattern of activity across voxels in ventral temporal cortex, it was possible to 'decode' which category of objects participants were viewing (e.g. faces, houses, chairs, shoes, etc.) Similar techniques have been used to demonstrate that it is possible to decode the orientation of visual stimuli viewed by participants, by examining the distribution of activity across voxels in early visual cortex (Kamitani and Tong, 2005). Within mrPFC, Haynes et al. (2007) showed that it was possible to predict which of two tasks participants were about to voluntarily perform, by examining the distribution of activation across individual voxels within this region, even though the overall level of mrPFC activity did not distinguish the two tasks. It is thought that multi-voxel pattern analysis techniques are effective because, although each individual voxel may contain a very large number of neurons, the distribution of different types of neuron may not be even from one voxel to the next, allowing functional differences to be observed (Haynes and Rees, 2006).

Gilbert et al. (2007) used a similar technique to examine their data comparing mrPFC activation related to mentalizing with that related to SO versus SI attention. Participants performed tasks in two separate contexts (spatial/verbal). Within each task context, the two orthogonal contrasts of Attention (SO versus SI) and Mentalizing (Mentalizing versus Non-mentalizing) were manipulated in a 2×2 factorial design. The spatial distribution of activity across mrPFC voxels related to mentalizing generalized significantly from one task context to the other, when the data were analysed on a participant-by-participant basis. Likewise, activity in the same voxels related to SO versus SI attention also generalized significantly between the two task contexts. However, the spatial distribution of activity related to mentalizing did not correlate significantly with the distribution related to attention. In other words, knowledge of how strongly a particular voxel responded to a particular contrast was significantly predictive of how well that same voxel would respond to the contrast examining the analogous function (attention or mentalizing) in the other task context. However, this knowledge did not predict how well that voxel would respond to the other function, even within the same task context. Thus, by analyzing the spatial distribution of activity across mrPFC voxels, as well as by analysing the location of activation peaks at the group level, this study established that there is relatively little overlap between neural populations within mrPFC involved in mentalizing and attention.

While multi-voxel pattern analysis techniques have proved effective in previous studies investigating typical participants, it is not yet known whether such techniques present advantages for comparing atypical populations with control groups. In the present study, we apply the protocol developed by Gilbert et al. (2007) to the investigation of functional organization of mrPFC in high-functioning ASD. Participants performed tasks in two separate contexts ('Alphabet' and 'Spatial') that alternated between SO phases (where on-screen information was task-relevant) and SI phases (where participants were required to perform the same task 'in their head', in the absence of task-relevant visual stimuli). The transitions between these phases were cued by
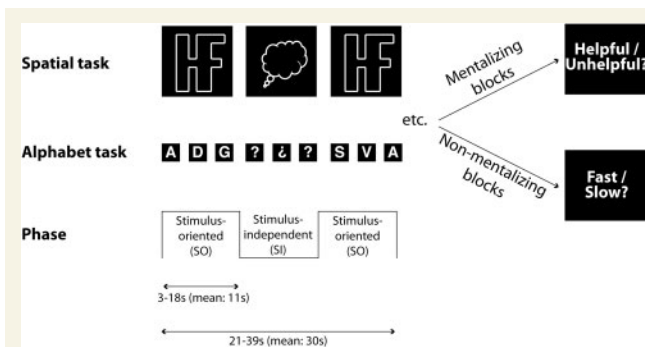
Figure 1 Schematic illustration of the behavioral tasks in Gilbert *et al*. (2007) and this study. Three orthogonal factors were manipulated in a 2 × 2 × 2 design. There were two separate task contexts (Alphabet/Spatial), presented in separate runs. Within these task contexts, participants alternated between two types of block (Mentalizing/Non-mentalizing). Within each block, participants alternated between two phases (Stimulus-Oriented; SO/Stimulus-Independent; SI). This allowed four orthogonal contrasts to be examined: (i) Alphabet Mentalizing (i.e. Mentalizing versus Non-mentalizing blocks in the Alphabet task context, collapsing over SO and SI phases); (ii) Alphabet Attention (i.e. SO versus SI phases of the Alphabet task context, collapsing over Mentalizing and Non-mentalizing blocks); (iii) Spatial Mentalizing; and (iv) Spatial Attention. In the Spatial task context (SO phase), participants repeatedly pressed one of two response buttons, as if navigating around the edge of a complex shape in a clockwise direction, to indicate whether the next corner would require a left or a right turn. During the SI phase this shape was replaced by a 'thought-bubble' shape and participants were required to imagine the shape that was presented in the SO phase and continue navigating as before. In the Alphabet task context (SO phase), participants classified upper-case letters of the alphabet according to whether they were composed of straight lines or curves. The stimuli cycled through the alphabet, skipping two letters between each stimulus and the next. In the SI phase, the letters were replaced with question marks. Participants mentally continued the sequence and continued classifying as before.

changes in the appearance of the visual stimuli, and occurred at unpredictable times. In half of the blocks ('Mentalizing blocks'), participants were told that they were performing the tasks in collaboration with an experimenter (cf. Gallagher *et al*., 2002), who was able to control the timing of transitions between the SO and SI phases with a button-press. At the end of these blocks (*M*: 30 s) participants made a judgment as to whether the experimenter was trying to be helpful or unhelpful in her timing of the transitions in that block. In fact, the timing of transitions between SO and SI phases was random. In non-mentalizing blocks, participants were told that the timing of these transitions was chosen randomly by the computer. At the end of these blocks, participants judged whether the transitions between phases occurred faster or slower than usual. Thus, the design of this study crossed two orthogonal factors: Attention (SO/SI)

and Mentalizing (Mentalizing/Non-mentalizing). In addition, these two factors were manipulated within two separate task contexts, performed in separate scanning sessions (see Fig. 1).

Methods

Participants

Thirty-two individuals participated in the study: 16 participants with Autism Spectrum Disorder (14 males) and 16 non-autistic control participants (12 males). Groups were matched on age [ASD *M*: 32 years, SD: 7.7; control *M*: 31 years, SD: 5.7; $t(30) = 0.6$, $P = 0.6$], verbal IQ [ASD *M*: 117, SD: 13.7; control *M*: 119, SD: 10.4; $t(30) = 0.6$, $P = 0.5$] and performance IQ [ASD *M*: 115, SD: 14.3; control *M*: 117, SD: 13.7; $t(30) = 0.4$, $P = 0.7$], measured using the Wechsler Adult Intelligence Scale (Wechsler, 1999). All participants in the ASD group had previously received a diagnosis from an independent clinician according to standard criteria. None of the participants had previously taken part in the study of Gilbert *et al*. (2007).

The Autism Diagnostic Observational Schedule—Generic (ADOS-G; Lord *et al*., 2000) was used to characterize the current level of functioning for all participants in the ASD group. On this measure, eight participants met criteria for autism, while four participants met criteria for autistic spectrum disorder. A further four participants did not meet either of these criteria and were included on the basis of their clinical diagnosis alone. All participants were right-handed, had normal or corrected-to-normal vision and were naïve with respect to the purpose of the experiment. The experiment was performed with local ethical committee approval from the UCL Research Ethics Committee and in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki. Written informed consent was obtained from all participants before their inclusion in the study.

Tasks and procedure

The present study followed the procedure described by Gilbert *et al*. (2007). In SO phases of the spatial task context, participants repeatedly pressed one of two buttons, as if navigating around the edge of a complex shape in a clockwise direction, to indicate whether the next corner would require a left or a right turn. During SI phases, the shape was replaced by a similarly sized white 'thought-bubble'; participants were required to imagine the shape that was presented in the SO phase and continue navigating from their current position.

In SO phases of the Alphabet task context, participants classified capital letters by pressing one of two buttons, according to whether the letter was composed entirely of straight lines or whether it had any curves. Subsequent letters were presented immediately following each button press, forming a regular sequence that cycled through the alphabet, skipping two letters between each stimulus and the next. During the SI phase, these letters were replaced with alternating question marks and upside-down question marks. Participants were required to mentally continue the sequence from their current position in the alphabet, performing the same classification task for each self-generated letter. The first letter to be presented in each SO phase was the appropriate continuation of the sequence, assuming that the sequence had been correctly maintained during the preceding SI phase.

Each task context (Alphabet/Spatial) was performed in two out of four runs in an AABB order counterbalanced across participants. Within each run, participants performed a total of eight blocks,

which alternated between mentalizing and non-mentalizing conditions. A different screen background (dark blue or dark red) was used for each condition, counterbalanced across participants. The length of each block varied randomly between 21 and 39 s (*M*: 30 s). In a randomly selected half of blocks ('fast blocks') transitions between the SO and SI phases occurred with a mean rate of every 7.6 s (range: 3–18 s). In other blocks ('slow blocks') transitions occurred at a mean rate of every 13.5 s (range: 3–18 s). At the end of each block, there was a 1-s pause, followed by a 5-s period during which participants indicated with a button press whether they believed the experimenter was trying to be helpful or unhelpful (in mentalizing blocks) or whether they believed the SI/SO transitions were faster or slower than average (in non-mentalizing blocks). Following this judgment, there was a 5-s reminder whether transitions were to be controlled by the computer or the experimenter in the following block. There was then a variable pause between 5 and 11 s (*M*: 8 s) before the next block began.

Pre-scan training

Participants took part in a pre-scan training session lasting ~40 min. They were first read a cover story explaining that the experiment would sometimes involve collaboration with the experimenter (see Supplementary material). They were then trained on each of the two task contexts. Following this, they performed one run of eight blocks of each task context. These runs were identical to the tasks performed in the experimental session, except that transitions between SO and SI phases in mentalizing blocks were controlled by button presses of the experimenter, who sat next to the participant (in accordance with the cover story). By contrast, in the experimental session transitions between SO and SI phases were always controlled by the computer, regardless of whether it was a mentalizing or non-mentalizing block.

Scanning procedure and data analysis

Behavioral data were analysed as in the previous study of Gilbert *et al.* (2007). Functional imaging data were acquired using a 3T Siemens Allegra head-only system (36 axial slices, 2-mm thick, separated by 1.7 mm; in-plane resolution 64×64; $3\,mm \times 3\,mm$ pixels). Functional scans were acquired during four sessions, each comprising 174 volumes (lasting ~7 min). Data were preprocessed and analysed using SPM5 software (http://www.fil.ion.ucl.ac.uk/spm/software/spm5/) using standard procedures (i.e. realigned, corrected for different slice acquisition times, normalized into 2-mm cubic voxels and smoothed with an 8-mm full-width half-maximum Gaussian kernel before random-effects analysis; see Gilbert *et al.*, 2007 for full details). For one participant in the ASD group, only three of the four sessions were acquired due to technical problems, meaning that the Alphabet task was only performed in one session for this participant.

Multi-voxel similarity analyses

These analyses examined the similarity, across voxels in mrPFC, of the spatial distribution of activity elicited by each contrast, regardless of the overall level of activity. Each participant was examined individually, and the signal at each voxel in mrPFC (defined as in Gilbert *et al.* 2007 as $-8 \leqslant x \leqslant 8$; $y \geqslant 40$; $-12 \leqslant z \leqslant 30$) was extracted for each of the four orthogonal contrasts: (i) Alphabet Attention (SO > SI), (ii) Alphabet Mentalizing (Mentalizing > Non-mentalizing), (iii) Spatial Attention (SO > SI) and (iv) Spatial Mentalizing (Mentalizing > Non-mentalizing). In order to investigate the
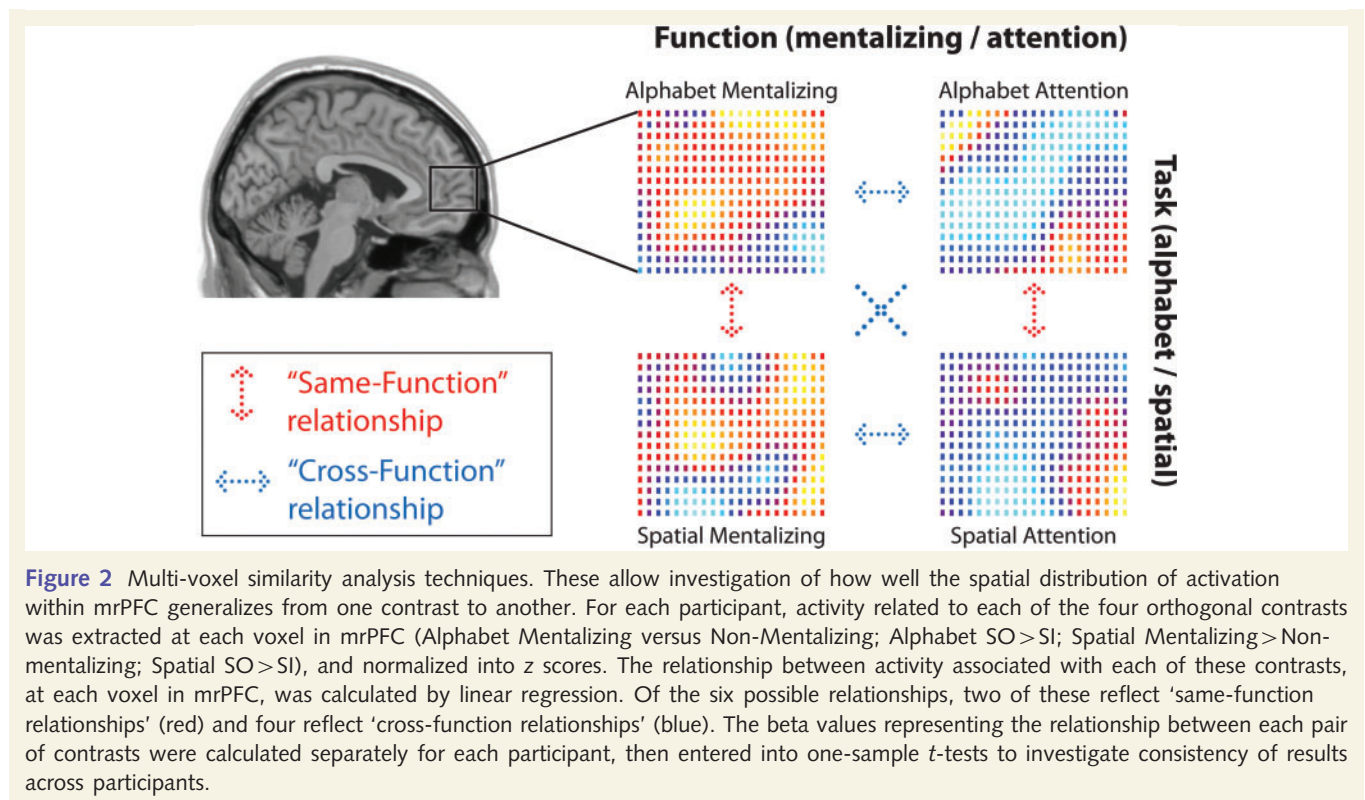
distribution of activity across voxels in mrPFC, activity was first normalized into z scores, separately for each contrast. This yielded a separate vector for each contrast, representing the contrast estimates for each voxel, relative to other voxels within mrPFC. Linear regression analyses were conducted to investigate the similarity between the spatial distribution of activation for each contrast, in comparison with each of the other contrasts (i.e. Alphabet Attention compared with Alphabet Mentalizing, Alphabet Attention compared with Spatial Attention, and so on for all six possible pairwise comparisons of two contrasts). For each comparison of two contrasts, a linear regression was performed to indicate the slope of the regression of the vector for one contrast on the vector for the other contrast. These regression analyses were conducted separately for each participant, and the resulting six beta values, representing each pairwise comparison of two contrasts, were entered into one-sample *t*-tests to test for significant results across participants. A significant positive beta value for a particular comparison between two contrasts would indicate that activity elicited in one contrast predicted the spatial distribution of activity for the other contrast. Alternatively, a significant negative beta values would indicate that the spatial distribution of activity was different in the two contrasts (i.e. inversely related). This analysis was conducted separately for the ASD and control groups, as illustrated in Fig. 2. Note, all results below were similar when Fisher-transformed correlation coefficients were analysed rather than the results of linear regression analyses. Along with these basic analysis, we conducted three specific analyses to compare the two groups: (i) 'Same-Function' versus 'Cross-Function' comparisons; (ii) consistency of functional specialization; and (iii) test–retest reliability. In addition, all multi-voxel pattern analyses were repeated after omitting the spatial smoothing step in preprocessing, in order to investigate the effects of smoothing on the observed results.

'Same-Function' versus 'Cross-Function' comparisons

For these analyses, we collapsed the six comparisons between contrasts into two categories: (i) Same-Function comparisons, where the same function (Attention or Mentalizing) was compared between the two tasks (e.g. Alphabet Attention compared with Spatial Attention) and (ii) Cross-Function comparisons, where different functions were compared (e.g. Alphabet Attention compared with Alphabet Mentalizing or Alphabet Attention compared with Spatial Mentalizing). In the study of Gilbert *et al.* (2007), the Same-Function relationships were significantly positive, suggesting that the spatial distribution of activity was reproducible between tasks when either Attention or Mentalizing functions were examined alone. However, the Cross-Function relationships were not significantly different from zero in this study, suggesting that the spatial distribution of activity was unrelated in comparisons between Attention and Mentalizing functions. In the present study, we investigated whether these results could be replicated, by collapsing results into a single pair of figures representing Same-Function and Cross-Function relationships.

Consistency of functional specialization

The results of Gilbert *et al.* (2007) suggested that some mrPFC voxels showed relatively high activity for the Attention contrast, regardless of the task, whereas other voxels showed relatively high activity for the Mentalizing contrast, regardless of the task, but these two sets of voxels were unrelated. There are two implications of these findings.

Figure 2 Multi-voxel similarity analysis techniques. These allow investigation of how well the spatial distribution of activation within mrPFC generalizes from one contrast to another. For each participant, activity related to each of the four orthogonal contrasts was extracted at each voxel in mrPFC (Alphabet Mentalizing versus Non-Mentalizing; Alphabet SO > SI; Spatial Mentalizing > Non-mentalizing; Spatial SO > SI), and normalized into z scores. The relationship between activity associated with each of these contrasts, at each voxel in mrPFC, was calculated by linear regression. Of the six possible relationships, two of these reflect 'same-function relationships' (red) and four reflect 'cross-function relationships' (blue). The beta values representing the relationship between each pair of contrasts were calculated separately for each participant, then entered into one-sample t-tests to investigate consistency of results across participants.

First, this would imply that some voxels showed preferential activity for the Attention contrast and some voxels showed preferential activity for the Mentalizing contrast (while others showed relatively high activity for both contrasts or neither). Second, the preferential activity of individual voxels for Attention versus Mentalizing contrasts would be expected to generalize between the Alphabet and Spatial task contexts, because voxels showing relatively high activity for Attention or Mentalizing in one task tended to do so in the other. In order to quantify this, we calculated a single figure for each participant representing the consistency of functional specialization for Attention versus Mentalizing across the two tasks. In other words, this analysis examined whether the preference of individual voxels for Attention versus Mentalizing functions in one task tended to generalize to the other.

In order to conduct this analysis, activity across voxels in mrPFC was first normalized into z scores, separately for each of the four contrasts (Alphabet Attention, Alphabet Mentalizing, Spatial Attention, Spatial Mentalizing). This ensured that results could not reflect different strengths of activation in the different contrasts. For example, a positive value for a particular voxel in the Alphabet Attention contrast would show that the level of activation for this contrast was higher than the mean level of activation for this contrast across all voxels in mrPFC, rather than reflecting the absolute degree of activation. Next, the normalized Alphabet Attention map was subtracted from the normalized Alphabet Mentalizing map, to form a new map representing functional specialization in the Alphabet task context. This map represented, at each voxel, the extent to which that voxel was selective for the Attention versus Mentalizing contrast. For example, a positive value for a particular voxel in this map would indicate that the degree of activation for the Alphabet Mentalizing contrast, relative to

other mrPFC voxels, was greater than the degree of activation for the Alphabet Attention contrast, relative to other mrPFC voxels. A similar procedure was conducted in the Spatial task context to yield an equivalent map of functional specialization. Finally, the similarity of the two maps, each representing functional specialization in one of the two task contexts, was assessed via subject-specific linear regression followed by one-sample t-test as above.

Test–retest reliability

In this analysis, we examined the extent to which the spatial distribution associated with a particular contrast in a particular scanning session was similar to the same contrast in another scanning session (i.e. test–retest reliability). This was possible because each task context was performed for two out of the four sessions. In this analysis, we compared the Alphabet Attention contrast in the first Alphabet session with the second Alphabet session, and so on for each of the four contrasts resulting from the crossing of task contexts (Alphabet, Spatial) with functions (Attention, Mentalizing). One participant in the ASD group could not be included in this analysis because only three out of four sessions were conducted due to technical problems. In all other analyses reported below, results for a particular contrast were averaged over the two sessions in which it was replicated.

Results

None of the results reported below differed significantly between ASD participants with different ADOS scores (autism/autism spectrum/none). Behavioral data were in line with the previous
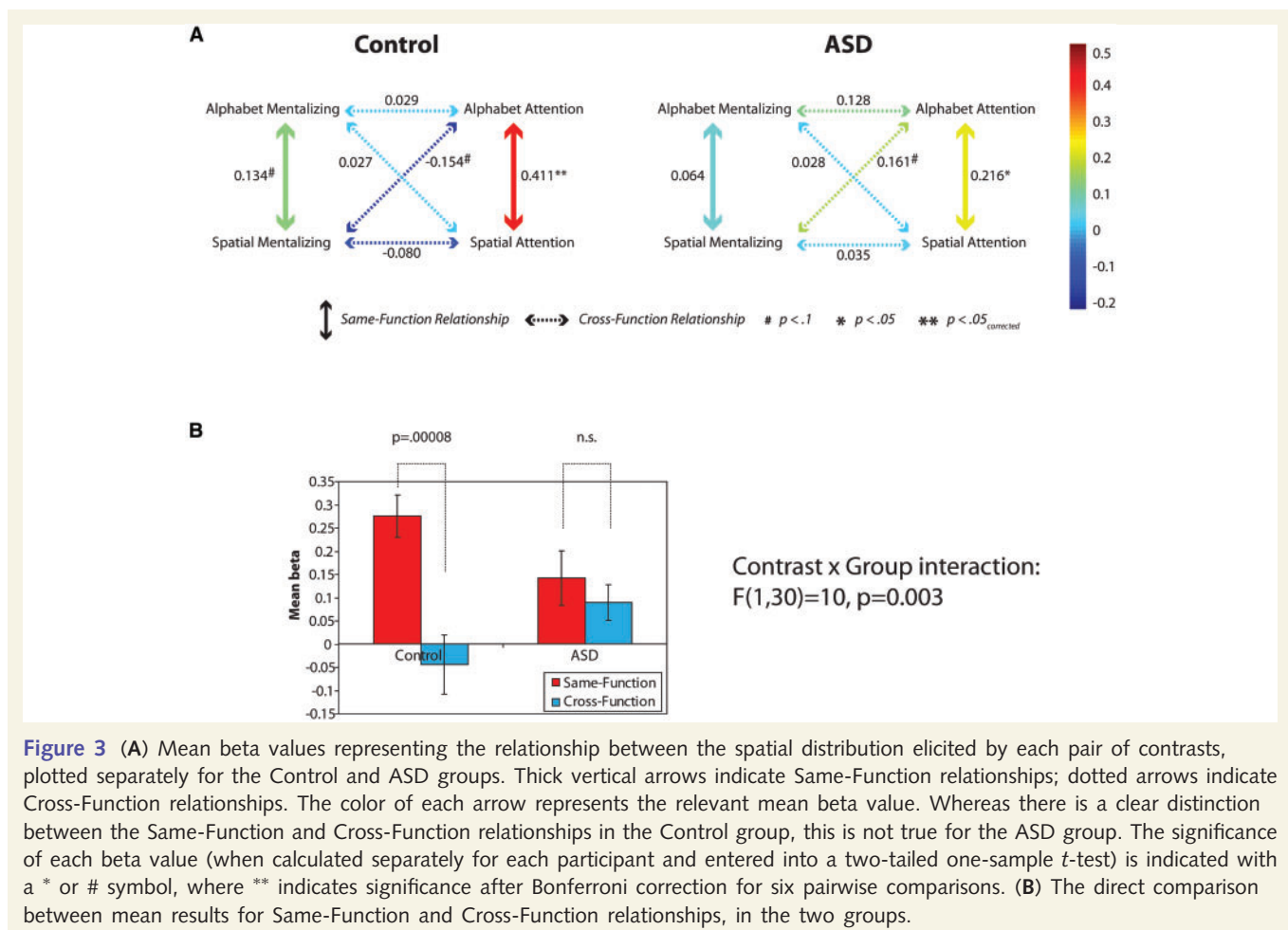
Figure 3 (**A**) Mean beta values representing the relationship between the spatial distribution elicited by each pair of contrasts, plotted separately for the Control and ASD groups. Thick vertical arrows indicate Same-Function relationships; dotted arrows indicate Cross-Function relationships. The color of each arrow represents the relevant mean beta value. Whereas there is a clear distinction between the Same-Function and Cross-Function relationships in the Control group, this is not true for the ASD group. The significance of each beta value (when calculated separately for each participant and entered into a two-tailed one-sample *t*-test) is indicated with a * or # symbol, where ** indicates significance after Bonferroni correction for six pairwise comparisons. (**B**) The direct comparison between mean results for Same-Function and Cross-Function relationships, in the two groups.

study of Gilbert *et al*. (2007) and did not differ significantly between the ASD and control groups. Results from conventional analyses of fMRI data were also consistent with this earlier study, and were comparable between the two groups. Using small-volume corrections centered on mrPFC, there were no significant group differences related to the contrast between SO and SI conditions. The contrast between Mentalizing and Non-mentalizing conditions yielded significantly greater mrPFC activity in the ASD group. At a whole-brain corrected threshold, there were no significant group differences in any analysis. Here, we focus on multi-voxel similarity analyses. For other results, see Supplementary materials.

Multi-voxel similarity analyses were conducted to examine the similarity in the distribution of activity across voxels in mrPFC for each of the four orthogonal contrasts: (i) Alphabet Attention (SO > SI), (ii) Alphabet Mentalizing (Mentalizing > Non-mentalizing), (iii) Spatial Attention (SO > SI), (iv) Spatial Mentalizing (Mentalizing > Non-mentalizing). Here, we report results from analyses using spatially smoothed data, for consistency with the earlier study of Gilbert *et al*. (2007) and conventional fMRI analyses. However, all results were similar when using unsmoothed data. In particular, all non-significant group

differences remained non-significant in both analyses, and all significant group differences remained significant.

The results from the control group were consistent with the findings of Gilbert *et al*. (2007). As in the previous study, the six pairwise comparisons between contrast images revealed just two significant positive relationships: Alphabet Attention versus Spatial Attention, and Alphabet Mentalizing versus Spatial Mentalizing [$t(15) > 1.78$, $P < 0.05$; one-tailed tests were used here because they replicate previous significant results]. All other tests were non-significant [$t(15) < 0.9$, $P > 0.38$], apart from a marginally significant *negative* relationship between the maps for Alphabet Attention and Spatial Mentalizing [$t(15) = 2.0$, $P = 0.07$; there was no prior expectation of a relationship here, so a two-tailed test was used]. In other words, the comparisons between analogous functions in the two task contexts produced significant positive relationships, but no such positive relationships were observed in the comparisons between Mentalizing and Attention functions (which, if anything, were associated with a negative relationship; see Fig. 3A). Below, the relationships between analogous contrasts in different task contexts will be labeled 'Same-Function relationships', and the other relationships will be labeled 'Cross-Function relationships'. Thus, in the
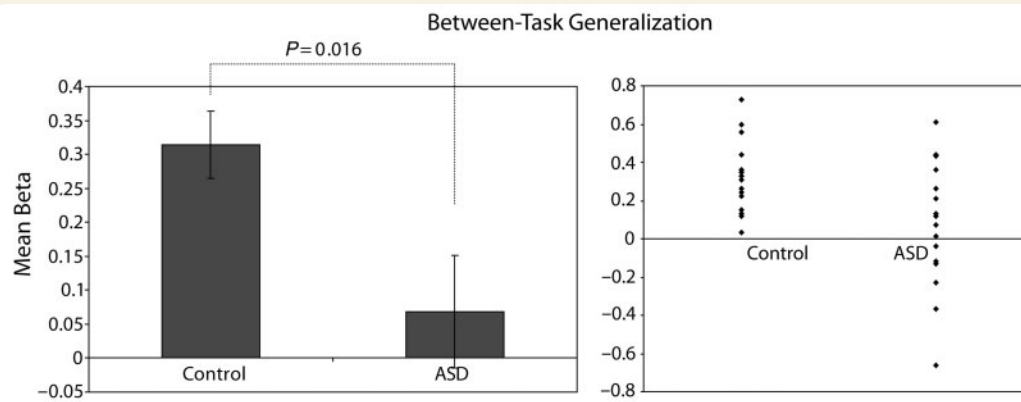
Figure 4 Generalization of functional specialization between tasks in the two groups. In these analyses, a single map of functional specialization was calculated within each task by subtracting normalized activity related to the Attention-related contrast from normalized activity related to the Mentalizing-related contrast, at each voxel. The voxel maps, representing functional specialization in each task, were then compared with each other. Left graph indicates the mean beta value, which is significantly greater than zero in the Control group, but not the ASD group. Error bars represent standard error of the mean. Right graph indicates beta values for each participant in the two groups.

control group, the Same-Function relationships were significantly positive, but the Cross-Function relationships were not significantly different from zero. However, in the ASD group, only one Same-Function relationship was significantly positive (Alphabet Attention versus Spatial Attention; $t(15) = 2.5$, $P = 0.012$, one-tailed), and in addition there was a positive Cross-Function relationship that just missed significance (Alphabet Attention versus Spatial Mentalizing; $t(15) = 2.0$, $P = 0.06$, two-tailed; Fig. 3A). No other pairwise comparison was significant [$t(15) < 1.62$, $P > 0.12$]. It therefore seems that, unlike the control group, the ASD group did not produce results that distinguished between Same-Function and Cross-Function comparisons.

Same-Function versus Cross-Function comparisons

In order to test formally for differences between Same-Function and Cross-Function relationships, two figures were calculated for each participant, representing the mean Same-Function and Cross-Function beta values. These figures were analysed in an ANOVA with factors Contrast (Same-Function or Cross-Function) and Group (control or ASD), revealing a significant main effect of Contrast [$F(1,30) = 20$, $P = 0.0001$] and, importantly, a Contrast × Group interaction [$F(1,30) = 10$, $P = 0.003$]. In the control group there was a highly significant difference between the Same-Function and Cross-Function relationships [$t(15) = 5.4$, $P = 0.00008$] whereas this comparison did not approach significance in the ASD group [$t(15) = 0.9$, $P = 0.4$]. Direct comparison of the two groups showed that Same-Function relationships tended to be higher in the control than the ASD group [$t(30) = 1.8$, $P = 0.08$] whereas Cross-Function relationships tended to be lower in the control than the ASD group [$t(30) = 1.8$, $P = 0.08$]. These results are shown in Fig. 3B.

Consistency of functional specialization

In a further analysis, a single figure was calculated for each participant, representing similarity of functional specialization (for Attention versus Mentalizing contrasts) in the two task contexts (Alphabet and Spatial). In the control group, there were a significant positive relationship between the maps representing functional specialization in the two task contexts [mean beta = 0.31; $t(15) = 6.4$, $P = 0.00001$]. Thus, in the control group the map of specialization for Attention versus Mentalizing functions generalized significantly from one task context to the other. However, there was no such relationship between the two maps in the ASD group [mean beta = 0.07; $t(15) = 0.8$, $P = 0.4$]. Direct comparison of the beta value reflecting consistency of functional specialization revealed a significant group difference [$t(30) = 2.6$, $P = 0.016$; Fig. 4].

Test–retest reliability

Test–retest reliability was significantly greater than zero in both groups [control: mean beta = 0.19, $t(15) = 4.8$, $P = 0.0002$; ASD: mean beta = .19, $t(14) = 4.7$, $P = 0.0004$], and did not differ significantly between them [$t(29) = 0.02$, $P = 0.99$]. Note that the absolute beta values here are not directly comparable to the earlier analyses, because they are based on analyses of individual sessions whereas the earlier analyses collapsed over the two sessions representing each task context. Finally, when the signal was averaged over voxels in mrPFC, the overall level of activity did not differ significantly between the groups for any of the contrasts [$t(30) < 1.52$, $P > 0.14$].

Discussion

In this study, we compared participants with ASD to age- and IQ-matched controls on tasks involving (i) attentional selection

between stimulus-oriented and stimulus-independent information and (ii) reflecting on the mental states of another person (i.e. mentalizing), both of which functions are thought to recruit mrPFC (Gilbert et al., 2007). Conventional analyses of fMRI data, looking for regions of mrPFC showing differences in overall levels of activity after averaging over participants, did not strongly distinguish the two groups. However, when the data were analysed on a participant-by-participant basis to investigate fine-grained functional specialization, highly significant group differences emerged. As in an earlier study using the same protocol (Gilbert et al., 2007), in the control group the spatial distribution of activation over mrPFC showed a significant positive relationship between one task context and the other, when the same function (attention-related or mentalizing-related) was examined (i.e. same-function relationships). But when patterns of activity associated with the two functions (attention and mentalizing) were compared with each other (i.e. cross-function relationships), the only result to approach significance was a *negative* relationship.

This distinctive pattern was not observed in the ASD group. In the six pairwise comparisons between the various contrasts, the attention-related contrasts, but not the mentalizing-related contrasts, generalized significantly from one task context to the other (Fig. 3A). However, the cross-function relationships also tended to be positive, so that data from the ASD group (unlike the control group) did not distinguish between same-function and cross-function relationships (Fig. 3B). In addition, the map of functional specialization between attention and mentalizing functions generalized significantly from one task context to the other in the control group, but not the ASD group (Fig. 4). These results indicate that the functional organization of mrPFC differs between participants with ASD and control participants. However, behavior was not significantly different between the two groups in the present study. This is consistent with the suggestion that even though high-functioning participants with ASD can perform similarly to control participants on many tests, they may accomplish this through the use of atypical mechanisms (Pierce et al., 2001).

Multi-voxel pattern analysis techniques have been shown to offer enhanced sensitivity in previous studies (Norman et al., 2006; O'Toole et al., 2007) but have most commonly been applied to investigations of the visual system in typical populations. The present results extend these findings in two regards. First, they demonstrate the applicability of such analysis techniques to the study of prefrontal cortex (Gilbert et al., 2007; Haynes et al., 2007). For instance, although conventional analysis of the mentalizing versus non-mentalizing contrast in the control group did not produce activation within mrPFC that met standard statistical thresholds (see Supplementary material), multi-voxel pattern analysis revealed a positive relationship between mentalizing-related activation in the two task contexts, indicating reproducible mentalizing-related activity in this region. Second, the present results demonstrate that multi-voxel pattern analysis can offer enhanced sensitivity for detecting differences between typical and atypical populations, compared with conventional analyses. This suggests that multi-voxel pattern analysis may prove to be an important technique for future studies investigating functional brain differences between various populations.

An important aspect of the design in the present study was the use of two task contexts (Alphabet and Spatial), each of which was subject to experimental manipulation of two orthogonal factors (Attention and Mentalizing). This approach is important in studies of executive function that seek to examine 'central processes' (Fodor, 1983; Burgess et al., 2006), i.e. processes that are insensitive to the precise stimuli or responses involved in a task, but which may potentially be involved in a wide range of different situations. In order to examine such processes, we used two task contexts which were dissimilar in terms of peripheral features such as the use of particular stimulus classes (e.g. letters versus non-alphabetical stimuli), but subject to analogous experimental manipulations with respect to putative central processes such as attentional selection between stimulus-oriented and stimulus-independent cognition (for other examples of this approach, see Burgess et al., 2001, 2003; Gilbert et al., 2005; Simons et al., 2006). The finding of significant relationships between the spatial distribution of activity related to these putative central processes, despite peripheral differences between the task contexts, suggests the involvement of mrPFC in such processes regardless of the precise nature of the stimuli or responses involved (Burgess et al., 2006).

Although the pattern of activity related to the Attention and Mentalizing contrasts generalized significantly from one task context to the other in the control group, cross-function relationships were non-significant (or even negative). These results suggest relatively little overlap between neural populations involved in Attention and Mentalizing functions, despite the close proximity of peak co-ordinates within mrPFC associated with both functions (Gilbert et al., 2007). The pattern of results in the ASD group was rather different. There was no clear distinction between same-function and cross-function relationships, unlike the control group. In addition, the single map representing functional specialization within a particular task context, calculated by subtracting the normalized response to one contrast from the response to the other within that task context, correlated highly significantly between the two task contexts in the control group, but not significantly in the ASD group (Fig. 4).

The finding that the map of functional specialization within one task context did not generalize to the other task context in the ASD group could be interpreted in several ways. One possibility might be that the spatial distribution of activity within mrPFC was completely random or undifferentiated in the ASD group (i.e. that there was no systematic functional specialization within mrPFC at all). However, this interpretation is not consistent with the analysis of test–retest reliability, which was highly significant in the ASD group, and not significantly different between the two groups. This indicates that the spatial distribution of activation related to a particular contrast was reproducible from one session to the next in both groups, rather than being random or uniform. If activity across mrPFC had been uniform for a particular contrast, the only variation between voxels would arise from noise, which would not be expected to be consistent from one session to the next. This finding additionally shows that global factors such as motion or stimulus-correlated motion cannot account for the group difference, seeing as such factors would influence test–retest reliability as much as other measures.

Below, we consider two other possibilities, both of which would be consistent with the present results, and are not mutually exclusive. The first of these possibilities is that the ASD group tended to activate a particular subset of voxels within mrPFC for the two task contexts, but that this subset of voxels was similar for the Attention and Mentalizing contrasts (unlike the control group, in which there was little overlap between the spatial distribution of activity associated with the two functions). In this case, individual voxels would have no reliable preference for the Attention versus Mentalizing contrasts, and hence there would be nothing to compare across the two task contexts. This hypothesis is consistent with the relatively higher cross-function relationships in the ASD than the control group (Fig. 3B). An alternative possibility is that the ASD group did exhibit functional distinctions between the Attention and Mentalizing contrasts within mrPFC, but that this group also recruited separable subregions of mrPFC for the Alphabet and Spatial task contexts, meaning that functional specialization in one task context did not generalize to the other. This hypothesis is consistent with the relatively lower same-function relationships in the ASD than the control group (Fig. 3B). At a physiological level, this hypothesis would be consistent with the suggestion of reduced synaptic pruning in ASD (Courchesne *et al*., 2003; C.D. Frith, 2003), which might lead to an overabundance of neural pathways and encourage the formation of relatively separate populations specialized for different tasks, rather than the need to reuse the same circuits for multiple tasks. At a behavioral level, this would be consistent with an increase in rote-learning and instance-learning, and a decrease in generalization and prototype extraction (Cohen, 1994; Gustafsson, 1997; Beversdorf *et al*., 2000; McClelland, 2000; U. Frith, 2003). Further studies will be required to distinguish these possibilities. In addition, further studies will be required to clarify the relationship between the present results, focusing on mrPFC, and differences in other brain regions associated with ASD. Of course, any particular behavior supported by mrPFC will depend on its interactions with other brain regions, and changes in brain development associated with ASD are unlikely to affect mrPFC alone. There are therefore likely to be complex interactions between multiple brain regions undergoing atypical development in ASD, and this may yield idiosyncratic changes in the abilities of different individuals. However, the present study establishes clear evidence for abnormal functional specialization within mrPFC in high-functioning participants with ASD, and the utility of multi-voxel pattern analysis techniques for investigating atypical populations.

Supplementary material

Supplementary material is available at *Brain* online.

Acknowledgements

References

Baird G, Simonoff E, Pickles A, Chandler S, Loucas T, Meldrum D, et al. Prevalence of disorders of the autism spectrum in a population cohort of children in South Thames: the Special Needs and Autism Project (SNAP). Lancet 2006; 368: 210–5.

Beversdorf DQ, Smith BW, Crucian GP, Anderson JM, Keillor JM, Barrett AM, et al. Increased discrimination of ''false memories'' in autism spectrum disorder. Proc Natl Acad Sci USA 2000; 97: 8734–7.

Bird G, Catmur C, Silani G, Frith C, Frith U. Attention does not modulate neural responses to social stimuli in autism spectrum disorders. Neuroimage 2006; 31: 1613–24.

Burgess PW, Alderman N, Forbes C, Costello A, Coates LM-A, Dawson DR, et al. The case for the development and use of ''ecologically valid'' measures of executive function in experimental and clinical neuropsychology. J Int Neuropsychol Soc 2006; 12: 194–209.

Burgess PW, Dumontheil I, Gilbert SJ. The gateway hypothesis of rostral prefrontal cortex (area 10) function. Trends Cogn Sci 2007; 11: 290–8.

Burgess PW, Quayle A, Frith CD. Brain regions involved in prospective memory as determined by positron emission tomography. Neuropsychologia 2001; 39: 545–55.

Burgess PW, Scott SK, Frith CD. The role of the rostral frontal cortex (area 10) in prospective memory: a lateral versus medial dissociation. Neuropsychologia 2003; 41: 906–18.

Castelli F, Frith C, Happé F, Frith U. Autism, Asperger syndrome and brain mechanisms for the attribution of mental states to animated shapes. Brain 2002; 125: 1839–49.

Chiu PH, Kayali MA, Kishida KT, Tomlin D, Klinger LG, Klinger MR, et al. Self responses along cingulate cortex reveal quantitative neural phenotype for high-functioning autism. Neuron 2008; 57: 463–73.

Cohen IL. An artificial neural network analogue of learning in autism. Biol Psychiatry 1994; 36: 5–20.

Courchesne E, Carper R, Akshoomoff N. Evidence of brain overgrowth in the first year of life in autism. J Am Med Assoc 2003; 290: 337–44.

Fodor J. The Modularity of Mind. Cambridge, MA: MIT Press; 1983.

Frith CD. What do imaging studies tell us about the neural basis of autism? In: Rutter M, editor Autism: Neural basis and treatment possibilities. Chichester, Wiley: Novartis Foundation; 2003.

Frith CD, Frith U. The neural basis of mentalizing. Neuron 2006; 50: 531–4.

Frith U. Autism: Explaining the enigma. 2nd edn. Oxford: Blackwell; 2003.

Frith U, Frith CD. Development and neurophysiology of mentalizing. Phil Trans R Soc B 2003; 358: 459–73.

Gallagher HL, Jack AI, Roepstorff A, Frith CD. Imaging the intentional stance in a competitive game. Neuroimage 2002; 16: 814–21.

Gilbert SJ, Bird G, Brindley R, Frith CD, Burgess PW. Atypical recruitment of medial prefrontal cortex in autism spectrum disorders: an fMRI study of two executive function tasks. Neuropsychologia 2008; 46: 2281–91.

Gilbert SJ, Frith CD, Burgess PW. Involvement of rostral prefrontal cortex in selection between stimulus-oriented and stimulus-independent thought. Eur J Neurosci 2005; 21: 1423–31.

Gilbert SJ, Spengler S, Simons JS, Steele JD, Lawrie SM, Frith CD, et al. Functional specialization within rostral prefrontal cortex (area 10): A meta-analysis. J Cogn Neurosci 2006; 18: 932–48.

Gilbert SJ, Williamson IDM, Dumontheil I, Simons JS, Frith CD, Burgess PW. Distinct regions of medial rostral prefrontal cortex supporting social and non-social functions. Soc Cogn Affect Neurosci 2007; 2: 217–26.

Gustafsson L. Inadequate cortical feature maps: a neural circuit theory of autism. Biol Psychiatry 1997; 42: 1138–47.

Hadjikhani N, Chabris CF, Joseph RM, Clark J, McGrath K, Aharon I, et al. Early visual cortex organization in autism: an fMRI study. Neuroreport 2004; 15: 267–70.

Haxby JV, Gobbini MI, Furey ML, Ishai A, Schouten JL, Pietrini P. Distributed and overlapping representations of faces and objects in ventral temporal cortex. Science 2001; 293: 2425–30.

Haynes JD, Rees G. Decoding mental states from brain activity in humans. Nat Rev Neurosci 2006; 7: 523–34.

Haynes JD, Sakai K, Rees G, Gilbert S, Frith C, Passingham RW. Reading hidden intentions in the human brain. Curr Biol 2007; 17: 323–8.

Just MA, Cherkassky VL, Keller TA, Kana RK, Minshew NJ. Functional and anatomical cortical underconnectivity in autism: evidence from an fMRI study of an executive function task and corpus callosum morphometry. Cereb Cortex 2007; 17: 951–61.

Kamitani Y, Tong F. Decoding the visual and subjective contents of the human brain. Nat Neurosci 2005; 8: 679–85.

Kana RK, Keller TA, Cherkassky VL, Minshew NJ, Just MA. Sentence comprehension in autism: Thinking in pictures with decreased functional connectivity. Brain 2006; 129: 2484–93.

Kleinhans NM, Müller R-A, Cohen DN, Courchesne E. Atypical functional lateralization of language in autism spectrum disorders. Brain Res 2008; 1221: 115–25.

Lord C, Risi S, Lambrecht L, Cook EJ, Levanthal B, DiLavore PC, et al. The autism diagnostic observation schedule-generic: a standard measure of social and communication deficits associated with the spectrum of autism. J Autism Dev Disord 2000; 30: 205–23.

McClelland JL. The basis of hyperspecificity in autism: A preliminary suggestion based on properties of neural nets. J Autism Dev Disord 2000; 30: 497–502.

Minshew NJ, Goldstein G, Siegel DJ. Neuropsychologic functioning in autism: Profile of a complex information processing disorder. J Int Neuropsychol Soc 1997; 3: 303–16.

Müller RA, Cauich C, Rubio MA, Mizuno A, Courchesne E. Abnormal activity patterns in premotor cortex during sequence leaning in autistic patients. Biol Psychiatry 2004; 56: 323–32.

Norman KA, Polyn SM, Detre GJ, Haxby JV. Beyond mind-reading: multi-voxel pattern analysis of fMRI data. Trends Cogn Sci 2006; 10: 424–30.

O'Toole AJ, Jinag F, Abdi H, Pénard N, Dunlop JP, Parent MA. Theoretical, statistical, and practical perspectives on pattern-based classification approaches to the analysis of functional neuroimaging data. J Cogn Neurosci 2007; 19: 1735–52.

Pierce K, Müller RA, Ambrose J, Allen G, Courchesne E. Face processing occurs outside the fusiform 'face area' in autism: evidence from functional MRI. Brain 2001; 124: 2059–73.

Schmitz N, Rubia K, Daly E, Smith A, Williams S, Murphy DG. Neural correlates of executive function in autistic spectrum disorders. Biol Psychiatry 2006; 59: 7–16.

Simons JS, Schölvinck M, Gilbert SJ, Frith CD, Burgess PW. Differential components of prospective memory? Evidence from fMRI. Neuropsychologia 2006; 44: 1388–97.

Wechsler D. Wechsler adult intelligence scale. 3rd edn. London: Harcourt Assessment; 1996.