

Recognition of Mental State Terms Clinical Findings in Children with Autism and a Functional Neuroimaging Study of Normal Adults

SIMON BARON-COHEN, HOWARD RING, JOHN MORIARTY, BETTINA SCHMITZ, DURVAL COSTA
and PETER ELL

Background. The mind's ability to think about the mind has attracted substantial research interest in cognitive science in recent decades, as 'theory of mind'. No research has attempted to identify the brain basis of this ability, probably because it involves several separate processes. As a first step, we investigated one component process - the ability to recognise mental state terms. **Method.** In Experiment 1, we tested a group of children with autism (known to have theory of mind deficits) and a control group of children with mental handicap, for their ability to recognise mental state terms in a word list. This was to test if the mental state recognition task was related to traditional theory of mind tests. In Experiment 2, we investigated if in the normal brain, recognition of mental state terms might be localised. The procedure employed single photon emission computerised tomography (SPECT) in normal adult volunteers. We tested the prediction (based on available neurological and animal lesion studies) that there would be increased activation in the orbito-frontal cortex during this task, relative to a control condition, and relative to an adjacent frontal area (frontal-polar cortex). **Results.** In Experiment 1, the group with autism performed significantly worse than the group without autism. In Experiment 2, there was increased cerebral blood flow during the mental state recognition task in the right orbito-frontal cortex relative to the left frontal-polar region. **Conclusions.** This simple mental state recognition task appears to relate to theory of mind, in that both are impaired in autism. The SPECT results implicate the orbito-frontal cortex as the basis of this ability.

Thinking about thinking seems to be one of the abilities that makes human cognition unique. This capacity for reflective thought has been variously dubbed theory of mind (Premack & Woodruff, 1978; Astington *et al.*, 1988), metarepresentation (Pylyshyn, 1978; Leslie, 1987), recursive mental model building (Johnson-Laird, 1983), self-consciousness (Johnson-Laird, 1988), mindreading (Whiten, 1991), or the intentional stance (Dennett, 1978). Here we refer to it as 'theory of mind', a generic term within cognitive science to denote the mind's ability to think about the range of mental states (intentions, desires, thoughts, beliefs, dreams, pretence, etc.).

Researching theory of mind is important in its own right, and also likely to throw light on the severe childhood psychiatric syndrome of autism. Children with autism have been shown to be impaired in a range of theory of mind tasks (Baron-Cohen *et al.*, 1985; Leslie & Frith, 1988; Baron-Cohen, 1989*a,b*, 1992*a*; Perner *et al.*, 1989; Reed & Peterson, 1990; Leekam & Perner, 1991; Charman & Baron-Cohen, 1992; Leslie & Thaiss, 1992; Swettenham, 1992; Phillips, 1993; see Baron-Cohen *et al.*, 1993, for an overview of this area). This cognitive deficit is

thought to underlie the social and communicative difficulties such children show (Baron-Cohen, 1988, 1990; Frith, 1989; Happé, 1993; Tager-Flusberg, 1993).

We report studies which are part of our longer term aim of identifying where in the brain theory of mind might be localised; and where in autistic children a specific dysfunction in this process might be occurring.

Functional neuroimaging techniques such as positron emission tomography (PET) or single photon emission computerised tomography (SPECT) require cognitive activation tasks comprising simple processes for repetition during the period of radioactive ligand uptake in the brain, in order to maximise the production of a clear signal. This period varies from about 40 seconds (for bolus $H_2^{15}O$ injections in PET) to about 5 minutes in the SPECT techniques used below. Traditional theory of mind tasks (Wimmer & Perner, 1983) are not simple enough, and take too long (at least 60 seconds each) to be repeated many times in the limited time period available. We therefore selected to investigate one aspect of theory of mind, namely the ability to

identify mental state terms, using a novel paradigm designed to allow repeated performance of this process.

While such a task does not require inferences about the content of mental states, or about the relation between mental states and action, we reasoned that mental state terms might be processed separately in the normal brain. The idea that different semantic categories might be localised neurologically has some plausibility, given studies of specific naming and comprehension deficits in some neurological patients (Warrington & Shallice, 1984; Hillis & Caramazza, 1991; Goodglass *et al*, 1966). If mental state terms are processed separately, this would be consistent with the notion that theory of mind is modular, both in cognitive and neurobiological terms (Leslie, 1991, 1994; Baron-Cohen, 1990, 1992*b*, 1994; Baron-Cohen & Ring, 1994).

In Experiment 2, we employed an experimental task (see Method) that fulfilled the essential criteria for cognitive activation in neuroimaging studies (George *et al*, 1991). We wanted first to check that this task was indeed related to theory of mind. Children with autism mostly fail theory of mind tests (see references cited above). We had every reason to predict that children with autism would fail on this task, since their spontaneous speech in most cases specifically lacks mental state terms (Tager-Flusberg, 1993). This was the basis for the first experiment.

Experiment 1

Subjects

We tested a group of 15 children with autism, diagnosed according to established criteria (Rutter, 1978; American Psychiatric Association, 1987), who were attending special schools for autism in the London area. Their age range was 8–16 years; verbal

mental age (Test of Reception of Grammar (TROG); Bishop, 1983) was 4–6 years age-equivalent (with one subject scoring at an age-equivalent of 10 years). We also tested a second group of 15 children who had moderate mental handicap but not autism. They had a similar age and mental age range, the latter assessed on the TROG. They were attending schools for children with learning difficulties in the London area. The details of the subjects are shown in Table 1.

Method

Each child was tested individually in a quiet room in the school. They were first asked a control question: "Do you know what is inside your head?" If the subject replied "No", or gave an answer that did not include the word 'mind', then they were told that their mind was also in their head. All of the group with mental handicap, but only a third of the group with autism, spontaneously mentioned their mind in answer. (The others gave responses which included reference to the brain, hair, blood, and bones.) After the clarification that their mind was in their head, they were then asked "Do you know what the mind can do?". If the subject did not spontaneously mention 'thinking' or any other related mental activity, then this was also supplied by the experimenter. All but two subjects in the group with mental handicap, but only four children with autism, spontaneously mentioned 'thinking' or a related mental activity (remembering, dreaming, etc.) in response to this question. Following this clarification of the mind's activity (which served as a warm-up and brief training phase), the experiment began.

The subject was asked to read aloud the words on the first word list (see Appendix A), in order to check that they could read the words. If the subject had any difficulty with this (and very few of them did), the experimenter helped to read the word. The experimenter then said: "OK. Now let's look at each word. I want you to say if each word is something the mind can do". The experimenter then pointed to the first word in the list, and said, "Let's begin here". If the child spontaneously worked through the list, then no further instruction was given. If however the child did not begin the task, the experimenter pointed to the first word in the list and asked "Can the mind do this?". The child then worked down the list, in fixed order, the experimenter pointing to each word in turn, asking "And can the mind do this?". Word list A contained 8 mental state words and 8 non-mental state words. The child received a final score out of 16. This guarded against a Yes- or No-bias.

Table 1
Subjects' ages and mental ages, in Experiment 1

	Chronological age (years: months)	Verbal mental age (TROG) (years: months)
Autism		
mean	12:3	4:8
s.d.	2:1	1:5
range	8:3-16:0	4:0-10:0
Mental handicap		
mean	11:9	4:6
s.d.	2:0	1:2
range	8:9-17:6	4:0-6:0

Table 2
Number of children in each group, passing each Word List,
in Experiment 2

	List A		List B	
	Pass	Fail	Pass	Fail
Autism ($n = 15$)	4*	11	14	1
Mental handicap ($n = 15$)	15	0	15	0

* $P < 0.0002$.

Following this, the child was given a second word list (see Appendix B). In place of the 8 mental state words were 8 words that described what the body could do, alongside 8 foil words neither body- nor mind-related. The subject was again instructed to read through the word list and the same procedure was followed, the question being "Can the body do this?". The score was again out of 16.

Finally, the order of presentation of word lists A and B was counterbalanced in each group, in order to guard against order-effects.

Results

Table 2 shows the number of children in each group passing the test, where a pass is defined as scoring equal to or more than 6 out of 8 on the target words, and equal to or more than 6 out of 8 on the foil words, on each list. The criterion of 6 or more out of 8 was based on establishing above chance performance, given that for each word, a correct score could be obtained with a probability of 0.5. The probability of passing 6 target words, or 6 foil words, by chance would therefore be extremely small ($P = 0.0156$).

Only 26.7% of the group with autism passed on list A, while 93.3% of them passed on list B. In contrast, all subjects in the group with mental handicap passed on lists A and B. This group difference on list A was highly significant ($\chi^2 = 14.35$, d.f. = 1, $P = 0.0002$). Finally, the four children with autism who passed on list A did not differ from the 11 who failed in terms of either chronological or verbal mental age (both $P > 0.05$).

Discussion

As predicted, children with autism were significantly impaired in recognising which words in a word-list were mind-related, compared with performance by a group of non-autistic children with mental handicap (and of an equivalent verbal mental age). This deficit is further evidence for an autism-specific impairment in the child's concept of mind (Baron-Cohen *et al*, 1993), and provides us with preliminary evidence

that the recognition task we used taps a component process in the larger domain of 'theory of mind'. Given its brevity and potential for repeatability, this task is well-suited to the requirements of neuro-imaging. In Experiment 2, we used an adult version of the task with a group of normal adult volunteers, in order to test if a specific brain region is involved in theory of mind tasks. We did not include individuals with autism in this investigation for ethical reasons.

Experiment 2

The use of SPECT employing ^{99m}Tc -hexamethyl propylene amine oxime (^{99m}Tc -HMPAO) and a triple-headed, brain-dedicated gamma-camera, allowed the production of high quality images with a shorter scan-acquisition time, and using lower doses of labelled ligand, than earlier SPECT paradigms. This technique is described elsewhere (Costa *et al*, 1986, 1989; George *et al*, 1991, review; Kouris *et al*, 1992). ^{99m}Tc -HMPAO is a lipophilic radiotracer which rapidly crosses the blood-brain barrier after intravenous administration, entering active neuronal cells in proportion to blood flow and then becoming trapped intracellularly in a hydrophilic state. Entering active neurons within 1-2 minutes after administration, it remains stable in those neurons for several hours, giving an accurate representation of the activity of the brain at the time of administration.

Experiment 2 required a comparison of cerebral blood flow during an experimental and a control task. As in Experiment 1, our experimental task involved judging if each word in a word list was a mind-related term or not. In our control task, the subject had to judge if each word in a new word list was a body-related term or not. This ensured that the target words in each condition related to people, and that the task demands were identical. We reasoned that if any difference in cerebral activation pattern was found, this might be due to one key difference between the tasks: that of processing mental state terms.

Planned analysis of scans, and predictions

We decided to perform a planned comparison of specific brain regions, based on four lines of evidence. Firstly, given that children with autism have specific impairments in the development of a theory of mind, and are specifically impaired on this mental state term recognition task (see Experiment 1, above), one possibility is that in autism the brain system responsible for theory of mind may be specifically

damaged (Baron-Cohen, 1990; Morton *et al*, 1991; Baron-Cohen & Ring, 1994). Although no consistent brain damage has been identified across cases, children with autism show classic signs of frontal lobe damage, such as repetitive behaviours, an inability to inhibit attention to salient objects (Russell *et al*, 1991), and deficits in tests of 'executive function', such as the Tower of Hanoi, the Wisconsin Card Sorting Task, and Milner's Maze Test (Prior & Hoffman, 1990; Ozonoff *et al*, 1991; Hughes & Russell, 1993). Since executive function tasks appear to be localised in the frontal cortex of the brain (Shallice, 1988), we surmised that theory of mind might also be localised in the frontal lobes, at least in part (an idea also discussed by Bishop, 1992), and we decided to restrict our planned analysis of brain regions to the frontal lobes.

Secondly, clinical descriptions of adult patients with known frontal lobe lesions not only show perseveration and failure to control attention, but also show social abnormalities (Shallice, 1988), suggesting that acquired frontal lobe damage may indeed disrupt social understanding. More specifically, patients with orbito-frontal lesions show loss of interest in social contact, and impaired social judgement (Eslinger & Damasio, 1985; Price *et al*, 1990). Thirdly, animal studies show that monkeys with orbito-frontal lesions (but not lateral-frontal lesions) show increased social avoidance (Butter *et al*, 1968). Fourthly, patients with orbito-frontal lesions have difficulties in the pragmatics of discourse (Kaczmarek, 1984), itself thought to require a theory of mind (Baron-Cohen, 1988; Happe, 1993; Tager-Flusberg, 1993). The evidence for the orbito-frontal hypothesis of theory of mind and autism is reviewed and discussed more extensively elsewhere (Baron-Cohen & Ring, 1994).

On the strength of these strands of evidence, we predicted that the orbito-frontal region of the brain would be significantly more activated during the mental state term recognition task than an adjacent frontal lobe region – the frontal polar region. To test for laterality effects, we examined four areas: left and right orbito-frontal cortex, and left and right frontal-polar cortex.

Subjects

Our subjects were 12 healthy male volunteers, all students aged 20–24 years. They were thus of broadly homogeneous intelligence and educational level. As in Experiment 1, each subject participated in both the experimental and control conditions, and the order of these was counterbalanced.

Procedure

The experimenter first explained to the subject that they should relax and lie as still as possible in the scanner, and not say anything during the procedure, in order to minimise unnecessary brain activation. The experimenter read the following instructions prior to the experimental condition: "You are going to hear a list of words. Some of these may repeat. Your job is to listen to each word. Most of the words describe what goes on in the mind, or what the mind can do, for example, words like think, believe, or imagine. Every time you hear a word which describes what goes on in the mind, raise the index finger of your left hand. OK?" Before the control condition, the same set of instructions was read to the subject, differing only in: "Most of the words describe things in the body, or what the body can do. For example, words like head, teeth, or walk." The experimenter then provided a warm-up of 5 words, to ensure the subject had understood and could perform the task, after which the experiment began.

The word lists contained 120 words, of which two-thirds were the target words, in random order. The proportion of two-thirds was chosen so as to maximise the time during which the relevant brain system(s) would be activated. Words were presented at a rate of 1 every 3 seconds. Mind-related words, which only appeared in the experimental condition, did not include any affect terms; body-related words only appeared in the control condition; foil words (which appeared in both conditions and were identical) were neither mind- nor body-related. The full sets of words (without the repetitions) are listed in Appendices C and D. The two lists were matched for both familiarity (Nelson & Kucera, 1982), and for the approximate number of nouns and verbs each list contained. Two independent judges confirmed that target words in each condition could be unambiguously identified.

Stereo headphones were placed over the subject's ears, and two experimenters remained present throughout the procedure to check that the subject was indeed making the correct judgements, as indicated by the subject's index finger movements after each target word. Each word list lasted for 6 minutes, during which half the total dose of ^{99m}Tc -HMPAO was injected intravenously into an indwelling cannular in the antecubital fossa. The dose administered on each occasion was 250 (± 15) MBq. Following each word list, a 22-minute scan was taken. Subjects had their eyes open during the word list presentation. Both conditions were administered in the same session, without interruption, so as to avoid the subject changing position.

SPECT scan acquisition and identification of cortical regions

Subjects were scanned on a triple-headed brain-dedicated SPECT system (GE/CGR Neurocam), using general purpose collimators. In air, the transaxial spatial resolution in the centre of the field of view, using such collimators, is 9.5 mm full width at half maximum (FWHM) (Moore *et al*, 1992). Images were acquired over 22 minutes, as 128 projections in a 64×64 matrix. Images were reconstructed using software provided by the scanner manufacturers. A Hanning prefilter was used with 1.0 cycle/cm cut off. A ramp filter was used for back-projection, and attenuation correction assumed a uniform linear attenuation coefficient (Chang type = 0.12 cm^{-1}). Transaxial images were reconstructed parallel to the orbito-meatal plane. The reconstructed horizontal and sagittal slices were 0.8 cm (2 pixels) thick.

For each subject, a region of interest (ROI) analysis was performed after normalising the two scans (control and experimental task) to the same total number of brain counts. Raters were blind to condition. Regions were identified from the functional images, with reference to a standard stereotaxic atlas of brain structure (Talairach & Tournoux, 1988). Counts were measured by placing 4×4 pixel ROIs in the target areas.

Regular ROIs ($1.6 \times 1.6 \times 0.8 \text{ cm}^3$) were used. Frontal-polar activity was measured by placing an ROI (2.05 cm^3) in both hemispheres on sagittal slices 0.8 to 2.4 cm lateral to the midline of the brain, extending from 0.8 to 2.4 cm above the plane of the basal orbito-frontal cortex. Visual comparison with the atlas indicated that this area was in the region of Brodmann's area 10. Orbito-frontal activity was measured by placing an ROI (2.05 cm^3) in both hemispheres, again on sagittal slices, 1.6 to 3.2 cm lateral to the midline, in the plane of the basal orbito-frontal cortex, approximately corresponding to Brodmann's area 11, but possibly also the most inferior parts of areas 10 and 32. (The orbito-frontal region is shown in Fig. 1). Finally, activation in the occipital cortex was measured by placing an ROI (2.05 cm^3) in the midline occipital region in the slice in which the corpus striatum and anterior cingulate cortex are best seen. This was so that frontal blood flow could be computed relative to extra-frontal blood flow, for each subject.

Results

For each of the frontal lobe ROIs in each condition, the mean ratio to occipital ROI activity in the

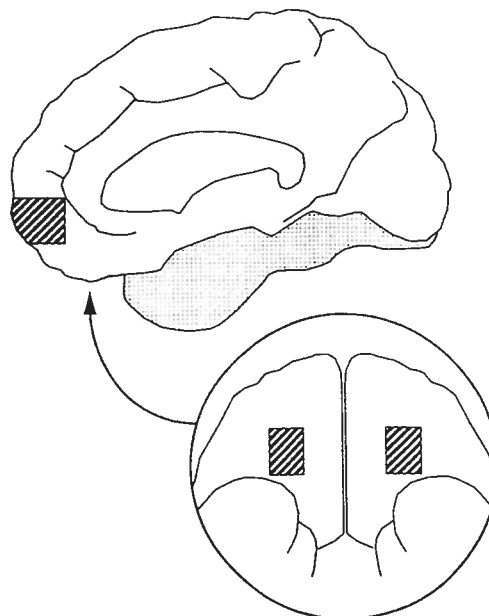


Fig. 1 The orbito-frontal cortex. The main figure represents the medial aspect of the mid-line sagittal face of the right hemisphere. The diagonally striped area is a diagrammatic representation of the location of the region of interest (ROI) described in the paper as frontal-polar. The circular inset represents a view of the anterior portion of the inferior aspect of the brain. The diagonally striped areas are representations of the locations of the left and right orbital frontal ROIs.

corresponding condition was calculated. These frontal/occipital ratios are shown in Table 3. Subsequently, the frontal to occipital ratio for each ROI in the experimental (E) condition was divided by the corresponding value for the control (C) condition. This generated what will be referred to as the corrected E/C ratio. The mean corrected E/C ratios for each ROI are shown in Table 4.

Paired *t*-test comparisons between these mean corrected E/C ratios, for each of the four frontal ROIs against each other, are shown in Table 5. Only

Table 3
Mean frontal/occipital ratios for the four frontal regions of interest (ROIs)

	RFP	LFP	ROF	LOF
Experimental task				
mean	0.798	0.768	0.737	0.716
s.d.	0.08	0.05	0.03	0.07
Control task				
mean	0.783	0.802	0.718	0.703
s.d.	0.07	0.05	0.07	0.05

Table 4
Mean experimental over control task ratio (E/C ratio), corrected, for each region of interest, in Experiment 2

	RFP	LFP	ROF	LOF
Mean	1.02	0.96	1.03	1.02
s.d.	0.06	0.07	0.06	0.06

RFP, right frontal polar; LFP, left frontal polar; ROF, right orbito-frontal; LOF, left orbito-frontal.

Table 5
t-test results comparing the mean experimental over control task ratios (E/C ratios), corrected, for each region of interest, against each other. (Experiment 2)

	RFP	LFP	ROF	LOF
RFP		3.78	-0.59	0.06
LFP			-5.04*	-2.73
ROF				0.58
LOF				

RFP, right frontal polar; LFP, left frontal polar; ROF, right orbito-frontal; LOF, left orbito-frontal.
* = $P < 0.001$.

six tests were performed, in order to test the predictions outlined earlier. Using a significance level of $P < 0.008$ (or $0.05/6$), in order to guard against chance significant results, this analysis revealed that activity in the right orbito-frontal cortex was significantly different from the activity in the left frontal-polar cortex during the mental state term recognition task ($t = -5.04$, $P < 0.001$). None of the other differences between regions were significant at this conservative probability level. Hence the comparison of experimental to control tasks indicates that performance of the mental state recognition task is associated with increased activity in the right orbito-frontal ROI relative to decreased activity in the left frontal polar ROI.

Discussion

In Experiment 2 we tested the prediction that the orbito-frontal region of the brain would be significantly more active during a mental state term recognition task than during a control task, relative to an adjacent frontal region (frontal-polar cortex). This prediction was confirmed, although only for the right orbito-frontal region relative to the left frontal-polar region. This unilateral and reciprocal effect was not predicted. The result is compatible with the idea that the orbito-frontal region plays a role in the processing of mental state concepts, as implicated

by evidence from acquired lesions to that region (Kaczmarek, 1984; Eslinger & Damasio, 1985; Price *et al.*, 1990).

Other possible interpretations of these data must also be considered. Firstly, while one key difference between the experimental and control conditions was the presence of mind-related words, another difference was that in the experimental condition the target words were also more abstract. Future studies to compare mind-related with other abstract words will be important and will also be relevant to questions concerning the modularity of theory of mind (Baron-Cohen, 1990, 1994; Leekam & Perner, 1991; Leslie, 1991; Charman & Baron-Cohen, 1992; Leslie & Thaiss, 1992; Leslie & Roth, 1993). We suspect that to the extent that the SPECT study is revealing a neural system that is used in the mental state recognition task, this is likely to be a system dedicated to the processing of mental state terms, rather than one dedicated to all abstract terms. Other studies have shown that children with autism are not globally impaired in the comprehension of abstract words (Eskes *et al.*, 1990; Hobson & Lee, 1989); as demonstrated in Experiment 1 here, they are specifically impaired in recognition of mind-related words.

Secondly, might the orbito-frontal cortex be just one part of a wider brain system involved in theory of mind? This possibility is very likely. In the present experiment, we carried out a focused comparison of four regions within the frontal lobes. Future studies will need to look either at the whole brain, or at the very least at more target regions, in order to test this question. One possibility is that subcortical regions such as the limbic system also play a role in theory of mind, given the evidence for the role of the limbic system (especially the amygdala) in emotion-processing (Brothers, 1990).

Thirdly, might subjects with autism have abnormalities in the orbito-frontal cortex? The common symptoms produced by orbito-frontal lesions include impaired social judgement (Eslinger & Damasio, 1985), utilisation behaviour (L'Hermitte, 1984), pragmatic/discourse breakdown (Kaczmarek, 1984), diminished aggression (Mateer & Williams, 1991), increased indifference (in monkeys: Meyer, 1972), decreased appreciation for dangerous situations (Mateer & Williams, 1991), hyper-olfactory exploration (in monkeys: Thorpe *et al.*, 1983), diminished response to pain (Goldman-Rakic, 1987), and excessive activity (in monkeys: Ferrier, 1886). These are all behaviours which are documented in clinical descriptions of autism (Wing, 1976; Frith, 1989; Baron-Cohen & Bolton, 1993). Since adults with autism can participate in SPECT scan studies (George *et al.*, 1992), and given the specific pattern

of cerebral activity shown in Experiment 2, we suggest that it will be important to use tasks like the one used here with subjects with autism, during functional neuroimaging. At the very least, we would predict a different pattern of brain activity within these particular regions to that seen in the normal adults in Experiment 2.

Fourthly, could our results reflect a possible confounding effect of the greater imageability of the body-related words? This question can be considered by examining the findings of other studies that have specifically investigated imageability. In a series of studies, Goldenberg and colleagues investigated the cerebral basis of visual imagery in normal volunteers (Goldenberg *et al*, 1987, 1989a,b, 1991). Combining the results from these studies, the authors concluded that while there was no evidence for an 'imagery centre', there was support for the hypothesis that the cerebral correlates of visual imagery, involving inferior temporal and occipital regions, are different from that of non-imaginal thinking. Note that no pre-frontal cortical regions were implicated in the process of visual imagery, making it unlikely that the prefrontal findings from Experiment 2 (above) simply reflect the use of imagery.

Fifthly, could our results simply be an effect of semantic monitoring? This explanation is important to consider, since Petersen *et al* (1988), using PET, demonstrated that activation in the prefrontal cortex was associated with semantic monitoring. However, Petersen *et al*'s results come from a comparison of semantic monitoring with a control condition consisting of the passive presentation of words. In our study both conditions involved semantic monitoring, thereby controlling for this process. Not surprisingly then, their results were different from ours, involving activation in the left prefrontal cortex 0.6 cm below the anterior cingulate/posterior cingulate (AC-PC) line. This more lateral region is distinct from both the orbital frontal and frontal polar regions investigated in our own study.

Related to the last point, could the difference between conditions be due to neural pathways responsible for verb recognition versus noun-recognition? This possibility is raised since, although both conditions contained a mix of nouns and verbs, and although the relative frequency of these did not differ significantly between the two conditions, the nouns in the mental state condition were often derived from verbs (e.g. knowledge derives from know). However, PET studies have not demonstrated that noun-noun v. noun-verb comparisons activate distinct cortical areas. Rather, in both types of semantic task, an increase in superior temporal gyri is seen (Wise *et al*, 1991). Given these findings, and

given that our analysis was of frontal regions, this suggests that our results are unlikely to reflect noun v. verb processing differences.

Finally, one puzzling question is the significance of the reciprocal responses between the left frontal polar and right orbito-frontal cortices. It may be that the former is involved in the body-term recognition task, just as the latter is activated by the mental state term recognition task. However, there appears to be no other evidence supporting this possibility. Alternatively, it may be that both of these two cerebral regions are involved, in a reciprocal manner, in the performance of the experimental task here. There are certainly other reports in the imaging literature which propose a reciprocal relationship between distant brain sites, associated with specific tasks or states. Such a relationship has been proposed, for example, between anterior and posterior cingulate regions in patients with depression (Bench *et al*, 1993). However, unlike that situation, there are no known strong and specific functional connections between contralateral orbito-frontal and frontal polar regions. This reciprocal effect merits further investigation.

Conclusions

From our evidence, we suggest that this cortical region is related to the processes involved in theory of mind. This is consistent with Stuss's (1991, p. 258) idea that a key function of the prefrontal cortex is for "knowing about knowing". It is also consistent with the clinical evidence suggesting frontal lobe damage can lead to a lack of self-reflective capacity (Ackerly & Benton, 1948; Luria, 1969). Our contribution has been to demonstrate indications of a role for the orbito-frontal cortex in this function. This result is compatible with studies showing that damage to the orbito-frontal cortex can lead to a failure to introspect (de Noble, 1835; cited in Blumer & Benson, 1975). Previous studies have all been based on cases of acquired damage to the orbito-frontal cortex. In contrast, given the clinical results from Experiment 1 reported here, our theory of autism posits early developmental damage to the same area. From other evidence from autism, we might expect this area to be part of a wider brain system (Baron-Cohen & Ring, 1994).

Acknowledgements

Howard Ring and John Moriarty were supported by the Raymond Way Neuropsychiatry Fund. Simon Baron-Cohen was at the Institute of Psychiatry, London, and in receipt of grants from the MRC and the Mental Health Foundation

during the period of this work. This project was supported by the Sir Jules Thorn Charitable Trust. We are grateful to Chris Frith for his comments on an earlier version of this paper. This work was presented at the Society for Philosophy and Psychology Conference, McGill University, Montreal, June 1992, and at the British Neuropsychiatry Association Annual Conference, Robinson College, Cambridge, July, 1993.

Appendix A

Mind-related word list (score out of 16).
 Letter Car Dream Think Tape Horse Want Computer Know
 Flower School Remember Pretend Cover Idea Understand

Appendix B

Body-related word list (score out of 16).
 Letter Car Hand Eye Tape Horse Face Computer Walk
 Flower School Blood Mouth Cover Finger Nose

Appendix C

Mind-related word list used in Experiment 2.
 Know Think Want Mind Idea Believe Expect Plan
 Remember Understand Decide Attend Recognise Hope
 Wish Assume Thought Realise Knowledge Forget Doubt
 Imagine Memory Desire Dream Belief Guess Realisation
 Pretence Intend Imagination Intention Ignore Pretend
 Reason Decision Learn Fantasy Consider Disbelieve

Appendix D

Body-related word list used in Experiment 2.
 Hand Eyes Move Run Face Foot Body Sit Walk Stand Head
 Arm Heart Digest Hair Fight Face Hit Tooth Blood Build
 Eat Mouth Finger Head Speak Muscle Artery Ear Smell
 Throat Jump Chest Smile Swim Bleed Tongue Elbow Sleep
 Neck

Foil words (these appeared in both lists in Experiment 2, randomly inserted among the target words).
 Business Present Nation Plant District Buy Note Table
 Clock Leaf Pen Computer Letter Car Tape Horse Flower
 School Cover Country

References

- ACKERLEY, S. & BENTON, A. (1948) Report of a case of bilateral frontal lobe defect. *Association for Research in Nervous and Mental Disease*, 27, 479-504.
- AMERICAN PSYCHIATRIC ASSOCIATION (1987) *Diagnostic and Statistical Manual of Mental Disorders* (3rd edn, revised) (DSM-III-R). Washington, DC: APA.
- ASTINGTON, J., HARRIS, P. & OLSON, D. (1988) *Developing Theories of Mind*. New York: Cambridge University Press.
- BARON-COHEN, S. (1988) Social and pragmatic deficits in autism: cognitive or affective? *Journal of Autism and Developmental Disorders*, 18, 379-402.
- (1989a) The autistic child's theory of mind: a case of specific developmental delay. *Journal of Child Psychology and Psychiatry*, 30, 285-298.
- (1989b) Are autistic children behaviourists? An examination of their mental-physical and appearance-reality distinctions. *Journal of Autism and Developmental Disorders*, 19, 579-600.
- (1990) Autism: a specific cognitive disorder of "mind-blindness". *International Review of Psychiatry*, 2, 79-88.
- (1991a) Do people with autism understand what causes emotion? *Child Development*, 62, 385-395.
- (1991b) The development of a theory of mind in autism: deviance and delay? *Psychiatric Clinics of North America*, 14, 33-51.
- (1992a) Out of sight or out of mind: another look at deception in autism. *Journal of Child Psychology and Psychiatry*, 33, 1141-1155.
- (1992b) On modularity and development in autism: a reply to Burack. *Journal of Child Psychology and Psychiatry*, 33, 623-629.
- (1994) How to build a baby that can read minds: cognitive mechanisms in mindreading. *Cahiers de Psychologie Cognitive (Current Psychology of Cognition)*, 13, 1-40.
- , LESLIE, A. M. & FRITH, U. (1985) Does the autistic child have a 'theory of mind'? *Cognition*, 21, 37-46.
- & BOLTON, P. (1993) *Autism: the Facts*. Oxford: Oxford University Press.
- & RING, H. (1994) A model of the Mindreading System: Neuropsychological and neurobiological perspectives. In *Origins of an Understanding of Mind* (eds P. Mitchell & C. Lewis). Hove: Lawrence Erlbaum Associates.
- , TAGER-FLUSBERG, H. & COHEN, D. J. (eds) (1993) *Understanding Other Minds: Perspectives From Autism*. Oxford: Oxford University Press.
- BENCH, C., FRISTON, K., BROWN, R., *et al* (1993) Regional cerebral blood flow in depression measured by PET: the relationship with clinical dimensions. *Psychological Medicine*, 23, 579-590.
- BISHOP, D. (1983) *Test of Reception of Grammar*. University of Manchester.
- (1992) Autism and frontal-limbic functions. *Journal of Child Psychology and Psychiatry*, 34, 279-294.
- BLUMER, D. & BENSON, D. (1975) Personality changes with frontal and temporal lobe lesions. In *Psychiatric Aspects of Neurological Disease Vol 1* (eds D. Benson & D. Blumer), pp. 151-170. New York: Grune & Stratton.
- BROTHERS, L. (1990) The social brain: a project for integrating primate behaviour and neurophysiology in a new domain. *Concepts in Neuroscience*, 1, 27-51.
- BUTTER, C., MISHKIN, M. & MIRSKY, A. (1968) Emotional responses towards humans in monkeys with selective frontal lesions. *Physiology and Behaviour*, 3, 213-215.
- CHARMAN, T. & BARON-COHEN, S. (1992) Understanding beliefs and drawings: a further test of the metarepresentation theory of autism. *Journal of Child Psychology and Psychiatry*, 33, 1105-1112.
- COSTA, D., ELL, P., CULLUM, I., *et al* (1986) The *in vivo* distribution of 99mTc-HMPAO in normal men. *Nuclear Medicine Communication*, 7, 646-658.
- , LUI, D., SINHA, A., *et al* (1989) Intracellular localization of 99mTc-HMPAO and Thallian-DDC in rat brain. *Nuclear Medicine Communication*, 10, 459-466.
- DENNETT, D. (1978) *Brainstorms: Philosophical Essays on Mind and Psychology*. New York: Harvester Press.
- ESKES, G., BRYSON, S. & MCCORMICK, T. (1990) Comprehension of concrete and abstract words in autistic children. *Journal of Autism and Developmental Disorders*, 20, 61-73.
- ESLINGER, P. & DAMASIO, A. (1985) Severe disturbance of higher cognition after bilateral frontal lobe ablation: Patient EVR. *Neurology*, 35, 1731-1741.
- FERRIER, D. (1886) *Functions of the Brain* (2nd edn). London: Smith & Elder.
- FRITH, U. (1989) *Autism: Explaining the Enigma*. Oxford: Basil Blackwell.
- GEORGE, M., COSTA, D., KOURIS, K., *et al* (1992) Cerebral blood flow abnormalities in adults with infantile autism. *Journal of Nervous and Mental Diseases*, 180, 413-417.
- , RING, H., COSTA, D., *et al* (1991) *Neuroactivation and Neuroimaging with SPET*. London: Springer-Verlag.
- GOLDMAN-PAKIC, P. (1987) Circuitry of primate prefrontal cortex and regulation of behaviour by representational memory.

- In *Handbook of Physiology: Section 1: Nervous System: Part 1*. Bethesda, MD: American Physiological Society.
- GOLDENBERG, G., PODREKA, I., STEINER, M., *et al* (1987) Patterns of regional blood flow related to memorizing of high and low imagery words: An emission computer tomography study. *Neuropsychologia*, **25**, 473-485.
- , ———, ———, *et al* (1989a) Regional cerebral blood flow patterns in visual imagery. *Neuropsychologia*, **27**, 641-664.
- , ———, UHL, F., *et al* (1989b) Cerebral correlates of imagining colours, faces, and a map: 1. SPECT of regional cerebral blood flow. *Neuropsychologia*, **27**, 1315-1328.
- , ———, STEINER, M., *et al* (1991) Contributions of occipital and temporal brain regions to visual and acoustic imagery: a SPECT study. *Neuropsychologia*, **29**, 695-702.
- GOODGLASS, H., KLEIN, B., CAREY, P., *et al* (1966) Specific semantic word categories in aphasia. *Cortex*, **2**, 74-89.
- HAPPÉ, F. (1993) Understanding communication in high functioning autistic subjects. *Cognition*, **48**, 101-119.
- HILLIS, A. & CARAMAZZA, A. (1991) Category-specific naming and comprehension impairment: a double dissociation. *Brain*, **114**, 2081-2094.
- HOBSON, R. P. & LEE, A. (1989) Emotion related and abstract concepts in autistic people: evidence from the British Picture Vocabulary Scale. *Journal of Autism and Developmental Disorders*, **19**, 601-623.
- HUGHES, C. & RUSSELL, J. (1993) Autistic children's difficulty with mental disengagement from an object: its implications for theories of autism. *Developmental Psychology*, **29**, 498-510.
- JOHNSON-LAIRD, P. (1983) *Mental Models*. Cambridge: Cambridge University Press.
- (1988) A computational analysis of consciousness. In *Consciousness in Contemporary Science* (eds A. Marcel & E. Bisiach). Oxford: Oxford University Press.
- KACZMAREK, B. (1984) Neurolinguistic analysis of verbal utterances in patients with focal lesions of frontal lobes. *Brain and Language*, **21**, 52-58.
- KOURIS, K., JARRITT, P., COSTA, D., *et al* (1992) Physical assessment of the GE/CGR Neurocam and comparison with a single rotating gamma camera. *European Journal of Nuclear Medicine*, **19**, 236-242.
- LEEKAM, S. & PERNER, J. (1991) Does the autistic child have a metarepresentational deficit? *Cognition*, **40**, 203-218.
- LESLIE, A. M. (1987) Pretence and representation: the origins of "theory of mind". *Psychological Review*, **94**, 412-426.
- (1991) The theory of mind impairment in autism: evidence for a modular mechanism of development? In *Natural Theories of Mind* (ed. A. Whiten). Oxford: Basil Blackwell.
- (in press) ToMM ToBy, and Agency: core architecture and domain specificity. In *Domain Specificity in Cognition and Culture* (eds L. Hirschfeld & S. Gelman). New York: Cambridge University Press.
- & FRITH, U. (1988) Autistic children's understanding of seeing, knowing, and believing. *British Journal of Developmental Psychology*, **6**, 315-324.
- & THAISS, L. (1992) Domain specificity in conceptual development: evidence from autism. *Cognition*, **43**, 225-251.
- & ROTH, D. (1993) What can autism teach us about metarepresentation? In *Understanding Other Minds: Perspectives From Autism* (eds S. Baron-Cohen, H. Tager-Flusberg & D. Cohen). Oxford: Oxford Medical Publications.
- LURIA, A. (1969) Frontal lobe syndromes. In *Handbook of Clinical Neurology*, Vol 2 (eds P. Vinken & G. Bruyn), pp. 725-757. Amsterdam: North Holland.
- L'HERMITTE, F. (1984) 'Utilization behaviour' and its relation to lesions of the frontal lobes. *Brain*, **106**, 237-255.
- MATEER, C. & WILLIAMS, D. (1991) Effects of frontal lobe injury in childhood. *Developmental Neuropsychology*, **7**, 359-376.
- MEYER, D. (1972) Some features of the dorsolateral frontal and inferotemporal syndromes in monkeys. *Acta Neurobiologica Experimenta*, **32**, 235-260.
- MOORE, S., KOURIS, K. & COLLUM, I. (1992) Collimator design for Single Photon Emission Tomography. *European Journal of Nuclear Medicine*, **19**, 138-150.
- MORTON, J., FRITH, U. & LESLIE, A. (1991) The cognitive basis of a biological disorder: autism. *Trends in Neurosciences*, **14**, 434-438.
- NELSON, W. & KUCERA, H. (1982) *Frequency Analysis of English Usage*. Boston: Houghton Mifflin.
- OZONOFF, S., PENNINGTON, B. & ROGERS, S. (1991) Executive function deficits in high-functioning autistic individuals: relationship to theory of mind. *Journal of Child Psychology and Psychiatry*, **32**, 1081-1105.
- PERNER, J., FRITH, U., LESLIE, A. M., *et al* (1989) Exploration of the autistic child's theory of mind: knowledge, belief, and communication. *Child Development*, **60**, 689-700.
- PETERSON, S., FOX, P., POSNER, M., *et al* (1988) Positron emission tomographic studies of the cortical anatomy of single-word processing. *Nature*, **331**, 585-589.
- PHILLIPS, W. (1993) Understanding intention and desire by children with autism. Unpublished PhD thesis, Institute of Psychiatry, University of London.
- PREMACK, D. & WOODRUFF, G. (1978) Does the chimpanzee have a 'theory of mind'? *Behaviour and Brain Sciences*, **4**, 515-526.
- PRICE, B., DAFFNER, K., STOWE, R., *et al* (1990) The computational learning disabilities of early frontal lobe damage. *Brain*, **113**, 1383-1393.
- PRIOR, M. & HOFFMAN, W. (1990) Brief report: neuropsychological testing of autistic children through an exploration with frontal lobe tests. *Journal of Autism and Developmental Disorders*, **20**, 581-590.
- PYLYSHYN, Z. (1978) When is attribution of beliefs justified? *Behavioural and Brain Sciences*, **1**, 592-593.
- REED, T. & PETERSON, C. (1990) A comparative study of autistic subjects' performance at two levels of visual and cognitive perspective taking. *Journal of Autism and Developmental Disorders*, **20**, 555-568.
- RUSSELL, J., MAUTHNER, N., SHARPE, S., *et al* (1991) The Windows Task as a measure of strategic deception in preschoolers and autistic children. *British Journal of Developmental Psychology*, **9**, 331-349.
- RUTTER, M. (1978) Diagnosis and definition. In *Autism: a Reappraisal of Concepts and Treatment* (eds M. Rutter & E. Schopler), pp. 1-26. New York: Plenum Press.
- SHALLICE, T. (1988) *From Neuropsychology to Mental Structure*. Cambridge: Cambridge University Press.
- STUSS, D. (1991) Self, awareness, and the frontal lobes: a neuropsychological perspective. In *The Self: Interdisciplinary Approaches* (eds J. Straus & G. Goethals). New York: Springer-Verlag.
- SWETTENHAM, J. (1992) The autistic child's theory of mind: a computer-based investigation. Unpublished PhD Thesis, University of York.
- TAGER-FLUSBERG, H. (1993) What language reveals about the understanding of minds in children with autism. In *Understanding Other Minds: Perspectives From Autism* (eds S. Baron-Cohen, H. Tager-Flusberg & D. J. Cohen). Oxford: Oxford University Press.
- TALAIRACH, J. & TOURNOUX, P. (1988) *Coplanar Stereotaxic Atlas of the Human Brain*. New York: Thieme Medical.
- THORPE, S., ROLLS, E. & MADDISON, S. (1983) The orbito-frontal cortex: neuronal activity in the behaving monkey. *Experimental Brain Research*, **49**, 93-115.
- WARRINGTON, E. & SHALLICE, T. (1984) Category specific semantic impairments. *Brain*, **107**, 829-854.

- WIMMER, H. & PERNER, J. (1983) Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, **13**, 103-128.
- WING, L. (1976) *Early Childhood Autism*. Pergamon Press.
- WISE, R., CHOLLET, F., HADER, U., *et al* (1991) Distribution of cortical neural networks involved in word comprehension and word retrieval. *Brain*, **114**, 1803-1817.

Simon Baron-Cohen, PhD, University of Cambridge; **Howard Ring**, MRCPsych, Academic Department of Psychiatry, London Hospital Medical College, Turner Street, London E1 2AD; **John Moriarty**, MRCPsych; **Bettina Schmitz**, MD, Raymond Way Neuropsychiatry Research Group, Institute of Neurology, University of London, Queen Square, London WC1N 3BG; **Durval Costa**, MD; **Peter Ell**, MD, Department of Nuclear Medicine, Middlesex Hospital, Mortimer St, London WC1

Correspondence: Dr Simon Baron-Cohen, Departments of Experimental Psychology and Psychiatry, University of Cambridge, Downing Street, Cambridge CB2 3EB

(First received 17 January 1994, final version 22 March 1994, accepted 28 April 1994)