

EDITORIAL

Schizophrenia and theory of mind¹

A BRIEF HISTORY OF THEORY OF MIND STUDIES

We suspect that people have an everyday theory of mind because they explain and frequently talk about the behaviour of others and themselves in terms of beliefs and desires. Having a theory of mind means that we believe that other people have minds like ours and that we understand the behaviour of these others in terms of the contents of their minds: their knowledge, beliefs and desires. But how can we demonstrate experimentally that people are using their theory of mind to predict the behaviour of others. This problem is particularly acute in the case of animals or young human children when they do not have language. Dennett (1978) discussing Premack & Woodruff's (1978) seminal paper 'Does the chimpanzee have a theory of mind?', suggested that the use of false beliefs to explain behaviour would provide convincing evidence. When their belief is true (i.e. corresponds to the actual state of the world) we can explain peoples' behaviour on the basis of the state of the world without needing to know about their beliefs. This ambiguity does not arise when the belief is false. The first experiment to use this approach was published by Wimmer & Perner (1983). They showed that at around 4 years of age a child knows that Maxi will look for his chocolates where Maxi believes them to be, even though the child knows that this belief is false because he has seen Maxi's mother moving the chocolates. In the English-speaking world the task involving Maxi and the chocolates has become the Sally-Anne task (see this issue, Lee *et al.* 2004). Since Wimmer & Perner's publication the development of theory of mind abilities in children has been studied intensively (e.g. Baron-Cohen *et al.* 2000; Wellman *et al.* 2001).

Theory of mind abilities have also been studied in various non-human species including apes, monkeys, dolphins and birds, but it remains a matter of controversy as to whether any creatures other than humans have such abilities (see Povinelli & Vonk, 2003; Tomasello *et al.* 2003). It is important not to equate theory of mind with social cognition. Social cognition includes many abilities for which a theory of mind is not necessary. This is clearly the case for monkeys. It is generally agreed that monkeys do not have any knowledge of the mental state of others (Cheney & Seyfarth, 1990), but at the same time they have very sophisticated social abilities, being able to recognize facial expressions, use a variety of calls and keep track of alliances among their peers (Seyfarth & Cheney, 2000).

Interest in the study of theory of mind received a considerable boost from the finding that children with autism have specific difficulties with theory of mind tasks (Baron-Cohen *et al.* 1985). The observation that performance of theory of mind tasks could be impaired while other abilities remained intact suggested that a dedicated neural system might underlie theory of mind abilities. A number of brain imaging studies have confirmed that a rather restricted group of brain regions are reliably activated when volunteers perform a variety of theory of mind tasks. In contrast there has been a surprising lack of reports that theory of mind problems can be acquired as a result of brain damage, but in last few years a few such studies have appeared. A useful review of the neural basis of theory of mind may be found in Lee *et al.* (2004). Much work still needs to be done before we can specify precisely what the brain regions associated with theory of mind tasks actually do (for some speculation see Frith & Frith, 2003). Does having a theory of mind depend upon being able to perform special kinds of computation, or does it depend on being able to represent a special kind of material (i.e. mental states)? There is some evidence that the material is critical. Autistic children can understand the concept of a false photograph, but not a false belief (Leslie & Thaiss, 1992) and

¹ Address for correspondence: Professor C. D. Frith, Wellcome Department of Imaging Neuroscience, Institute of Neurology, 12 Queen Square, London WC1N 3BG.

the same distinction is found in the brain (Saxe & Kanwisher, 2003). Our understanding of the neural underpinnings of theory of mind awaits the development of a computational account of the cognitive mechanism.

THEORY OF MIND AND SCHIZOPHRENIA

The term 'autism' was originally coined by Bleuler to describe a feature of schizophrenia. The term 'autistic aloneness' is also used and describes a withdrawal from other people into the self. This 'qualitative impairment in social interaction' (DSM-IV) is a key diagnostic feature of autism and is plausibly accounted for by a deficit in theory of mind (Frith *et al.* 1991). It therefore seems reasonable to suggest that a similar cognitive deficit might underlie autism in schizophrenia (Frith & Frith, 1991). The major difference between autism and schizophrenia is that autism is present from early childhood, while many patients with schizophrenia show no signs of abnormal social interactions until they reach adulthood. Clearly, these people had a theory of mind before the onset of their illness. We might therefore expect the effects of the subsequent deficit to be different. It is misleading to refer to the problem as a theory of mind deficit in the case of schizophrenia. A person who does not have a theory of mind takes no account of the beliefs and desires of other people when trying to understand their behaviour. Indeed, he may not have any concept of beliefs and desires. This may be the case in schizophrenic patients with predominantly negative features, but not in those with positive symptoms. The patient with paranoid delusions has no problem ascribing intentions to other people. His problem is that he ascribes the wrong intentions. He has a theory of mind since he explains the behaviour of others in terms of their intentions. The fault lies in the mechanism that is used to discover what these intentions are. This mechanism is better referred to as 'mind reading' or 'mentalizing'.

A number of studies have shown that patients with schizophrenia perform theory of mind tasks badly (see Lee *et al.* 2004 for a review). Their problem with these tasks does not seem to be a secondary consequence of general cognitive impairments (Doody *et al.* 1998) or problems with executive function (Pickup & Frith, 2001). As in autism the problem is domain specific since patients can deal with false maps, but not false beliefs (Pickup & Frith, 2001). Whether there is a special relationship between mentalizing problems and symptoms remains unclear. Patients with negative features and patients with thought disorder tend to perform theory of mind tasks badly, but the results for patients with paranoid symptoms are equivocal. One problem is that, for most of the tasks used, poor performance does not distinguish between lack of theory of mind (or under-mentalizing) as might be associated with autism and over-mentalizing as might be associated with paranoia. Either problem would lead to a wrong answer. It might be possible to address this problem with a detailed error analysis. Someone with a lack of theory of mind would predict behaviour on the basis of the actual state of the world rather than beliefs. Someone who over-mentalized would predict behaviour on the basis of the wrong beliefs. In the right situation these strategies would lead to different kinds of error. The purest evidence for over-mentalizing would be to show that paranoid patients ascribe attentions to behaviour that the rest of us see as mechanical or random (Blakemore *et al.* 2003).

MENTALIZING AND COMMUNICATION

Most human social interactions involve language and the successful use of language leads to communication, i.e. the exchange of knowledge and beliefs. This aspect of language is called pragmatics and clearly involves mentalizing. Grice (1957) was among the first to consider how such communications work. He proposed a series of maxims for conversation that would help to ensure a meaningful exchange of information (Grice, 1975). For example, there is the maxim of quantity 'Be as informative as is required but do not make your contribution more informative than is required'. In other words do not tell people things they already know. Just tell them what they need to know. Following this maxim depends upon mentalizing since we have to know something about the current state of knowledge of the person to whom we are talking. People who have difficulty mentalizing should have problems with this communicative aspect of language. There are many

anecdotal accounts of people with autism either failing to tell other people important things that they need to know or telling them things that they know very well already. An example of the latter is the young man who often telephones his favourite aunt, but always announces himself, 'This is M. C. Smith, your nephew speaking'.

If patients with schizophrenia have problems with mentalizing this should lead to very specific problems in communicating through language. Many of the deviant aspects of language associated with schizophrenia can be plausibly explained in terms of mentalizing problems. For example, a patient with 'poverty of content of speech' converses at adequate length, but supplies little information clearly violating Grice's maxim of quantity. 'Tangentiality' and 'derailment' arise when the patient says things that are apparently irrelevant to what has just been said. This might arise from over-mentalizing whereby the patient ascribes knowledge to the listener that the listener does not have.

Corcoran & Frith (1996) investigated implicit knowledge of Gricean maxims by presenting patients with vignettes in which different responses were made. For example, Jo is in the tobacconist's. The assistant asks him what he would like. Jo replies: (a) 'A packet of cigarettes please'; or (b) 'Twenty Benson and Hedges please'. The patient was asked to choose which Jo was more likely to say. In this example, (a) is not sufficiently informative (flouting the maxim of quantity) and therefore (b) is the better answer. In general, the schizophrenic patients performed this task worse than normal and psychiatric control subjects. Within the schizophrenic group, however, it was only the patients with negative features (including poverty of content of speech) who performed badly. The patients with predominantly paranoid symptoms performed in the normal range. However, a more recent study (Tenyi *et al.* 2002) found that patients with 'paranoid schizophrenia' were impaired in detecting violations of relevance (another Gricean maxim) in conversational vignettes.

Non-literal interpretations of speech that are required to understand figures like metaphor and irony are also believed to depend upon mentalizing (e.g. Happé, 1993). Langdon *et al.* (2002) found that understanding of irony was related to understanding of false beliefs (tested with non-verbal picture sequences) and that both were impaired in patients with high ratings of positive formal thought disorder.

MENTALIZING IN REAL LIFE

I have outlined how problems with mentalizing could explain many of the symptoms associated with schizophrenia including their problems with communication. While the relation between task specific symptoms and task performance is not clear-cut there is now much evidence that many patients have problems with mentalizing tasks. What then are we to make of the study, in this issue, by McCabe *et al.* (2004)? This study reports an intensive analysis of clinical interviews with patients with schizophrenia. The results are very clear. The patients did guide their utterances on the basis of what the listener needed to know. The patients knew that the listener had different beliefs from their own and took account of these differences in generating their replies. In these real life conversations the patients used accurate mentalizing to achieve successful communication. They did not show theory of mind problems. The only identifiable problem was that they could not generate understandable justifications for their beliefs. They recognized that their justifications were not understandable, but were not able to produce better justifications. Furthermore, they did not modify their beliefs even though they recognized that these beliefs were not shared by anyone else.

At first sight this results seems incompatible with the many previous studies using more traditional mentalizing tasks. How can they be reconciled? One possibility is that the patients studied by McCabe *et al.* did not have mentalizing problems and hence would not have performed badly on traditional mentalizing tasks. This would make them unusual since the evidence is that patients with schizophrenia tend to perform such tasks badly unless they are currently symptom free. Nevertheless, it would be very interesting to conduct a study in which both discourse performance and performance on mentalizing tasks was measured in the same patients. I predict such a study will show that many patients will perform the mentalizing tasks badly while at the same time showing intact mentalizing in their discourse.

There is a fundamental difference between the use of mentalizing in discourse and the use of mentalizing in theory of mind tasks. In other domains this difference has been characterized as 'on-line' versus 'off-line' processing (e.g. Tyler, 1992). During discourse mentalizing is used implicitly and automatically in the service of communicating. In this sense it is used on-line. In most theory of mind tasks mentalizing is carried out off-line. The patient is not taking part in the interaction, but must make explicit use of mentalizing to answer questions about an interaction that has been described. This requirement puts more weight on working memory and on meta-cognitive processes (i.e. reflecting on mentalizing). As might be expected off-line tasks are more difficult than on-line tasks. For example, Kempler *et al.* (1998) showed that patients with Alzheimer's disease were impaired in detecting violations of grammar in sentences like 'Susan went meetings'. However, when asked to read the sentences aloud they showed the normal pattern of naming the ungrammatical words more slowly. This latency difference indicates that, at an implicit level, they detected the grammatical violations.

There is evidence for a similar distinction in mentalising tasks. Clements & Perner (1994) used a version of the Sally-Anne task in which Sally enters the scene by a different door depending on where she is going to look for her marble. When told that Sally was about to return 90% of 3-year olds looked at the correct door (i.e. the one consistent with Sally's false belief). However, only 45% of these same children gave the right answer when asked explicitly where Sally would look for her marble. I predict that many patients with schizophrenia would show the same pattern of performance.

A DISORDER OF CONSCIOUSNESS?

Do patients with schizophrenia have a specific problem with explicit mentalizing? Before considering this we need to be sure that the problem is not with explicit mentalizing, but rather with the increased working memory load and meta-cognitive processes that are inevitably associated with explicit tasks. I think this is an unlikely explanation since Pickup & Frith (2001) have shown that patients who fail false belief tasks can still do false map tasks. These explicit tasks differ only in the mentalizing component.

Implicit mentalizing is not sufficient for all aspects of discourse. For example explicit mentalizing is required if we are confronted with the problem of how to change someone's belief. This is the one problem identified by McCabe *et al.* (2004) in their analysis of schizophrenic discourse.

The idea that schizophrenia is associated with intact implicit processing, but impaired explicit processing is closely related to the much earlier idea contrasting intact automatic processing and impaired controlled processing (Callaway & Naghdi, 1982). The difference is that the new version of this idea puts much more emphasis on the conscious aspect of controlled processing, in particular reflexive consciousness when we think about our own mental states or those of others. Evidence for a problem of this type in memory tasks has recently been reported in a series of studies by Danion and his colleagues (e.g. Danion *et al.* 1999; Bacon *et al.* 2001). These authors conclude that the memory problems shown by patients with schizophrenia lie in the domain of meta-memory and reflect an impaired conscious awareness of one's own knowledge. More automatic memory processing remains intact.

A similar dissociation has recently been observed by Dehaene *et al.* (2004) in a motor conflict task. Subjects are presented with numbers and have to indicate with button presses whether these are larger or smaller than five. Performance on this task is affected by the prior subliminal presentation of another, visually masked number. Response times are faster when the prime and target number are congruent and slower when they are incongruent. In this version of the task patients with schizophrenia showed normal effects of congruent and incongruent primes. However, if the primes were visible the behaviour of the patients was markedly different. Their responses were slowed down to a much greater extent than those of the controls, and furthermore the patients no longer showed congruency effects. Dehaene and his colleagues conclude that 'the cognitive impairment in schizophrenia is rather selectively linked to the anomalous monitoring and regulation of conscious conflicts'.

These observations provide an interesting link between mentalizing problems and the more general cognitive problems associated with schizophrenia. A possible common factor is that all involve conscious reflection upon states of mind whether in one's self or in others. I suspect that in normal life we do not spend very much time consciously reflecting on states of mind. Much of ordinary day-to-day life can go on perfectly well without such reflection. Indeed, one of the key problems for research on consciousness is to identify tasks for which consciousness is actually necessary. It is striking that patients with hallucinations and delusions, but without pronounced negative features, perform within normal limits on traditional cognitive tests (e.g. Johnstone & Frith, 1996). Thus, the cognitive anomaly associated with schizophrenia should not have much impact on many of their daily activities. A disorder of conscious reflection would be such an anomaly.

C. D. FRITH

REFERENCES

- Bacon, E., Danion, J. M., Kauffmann-Muller, F. & Bruant, A. (2001). Consciousness in schizophrenia: a metacognitive approach to semantic memory. *Consciousness and Cognition* **10**, 473–484.
- Baron-Cohen, S., Leslie, A. M. & Frith, U. (1985). Does the autistic-child have a theory of mind. *Cognition* **21**, 37–46.
- Baron-Cohen, S., Tager-Flusberg, H. & Cohen, D. J. (2000). *Understanding Other Minds: Perspectives from Developmental Cognitive Neuroscience*. Oxford University Press: Oxford.
- Blakemore, S.-J., Sarfati, Y., Bazin, N. & Decety, J. (2003). The detection of intentional contingencies in simple animations in patients with delusions of persecution. *Psychological Medicine* **33**, 1433–1441.
- Callaway, E. & Naghdi, S. (1982). An information processing model for schizophrenia. *Archives of General Psychiatry* **39**, 339–347.
- Cheney, D. L. & Seyfarth, R. M. (1990). *How Monkeys See the World: Inside the Mind of Another Species*. University of Chicago Press: Chicago, IL.
- Clements, W. A. & Perner, J. (1994). Implicit Understanding of Belief. *Cognitive Development* **9**, 377–395.
- Corcoran, R. & Frith, C. D. (1996). Conversational conduct and the symptoms of schizophrenia. *Cognitive Neuropsychiatry* **1**, 305–318.
- Danion, J. M., Rizzo, L. & Bruant, A. (1999). Functional mechanisms underlying impaired recognition memory and conscious awareness in patients with schizophrenia. *Archives of General Psychiatry* **56**, 639–644.
- Dehaene, S., Artiges, E., Naccache, L., Martelli, C., Viard, A., Schürhoff, F., Recasens, C., Martinot, M. L. P., Leboyer, M. & Jean-Luc Martinot, J.-L. (2004). Conscious and subliminal conflicts in normal subjects and patients with schizophrenia: the role of the anterior cingulate. *Proceedings of the National Academy of Sciences of the United States of America* (in the press).
- Dennett, D. C. (1978). Beliefs about beliefs. *Behavioral and Brain Sciences* **1**, 568–570.
- Doody, G. A., Gotz, M., Johnstone, E. C., Frith, C. D. & Owens, D. G. (1998). Theory of mind and psychoses. *Psychological Medicine* **28**, 397–405.
- Frith, C. D. & Frith, U. (1991). Elective affinities in schizophrenia and childhood autism. In *Social Psychiatry: Theory, Methodology, and Practice* (ed. P. E. Bebbington), pp. 65–88. Transaction Publishers: New Brunswick, NJ.
- Frith, U. & Frith, C. D. (2003). Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society of London, Series B* **358**, 459–473.
- Frith, U., Morton, J. & Leslie, A. M. (1991). The cognitive basis of a biological disorder: autism. *Trends in Neurosciences* **14**, 433–438.
- Grice, H. P. (1957). Meaning. *Philosophical Review* **66**, 377–388.
- Grice, H. P. (1975). Logic and conversation. In *Syntax and Semantics. Vol. 3. Speech Acts* (ed. P. Cole and J. L. Morgan), pp. 41–58. Academic Press: New York.
- Happé, F. G. (1993). Communicative competence and theory of mind in autism: a test of relevance theory. *Cognition* **48**, 101–119.
- Johnstone, E. C. & Frith, C. D. (1996). Validation of three dimensions of schizophrenic symptoms in a large unselected sample of patients. *Psychological Medicine* **26**, 669–679.
- Kempler, D., Almor, A., Tyler, L. K., Andersen, E. S. & MacDonald, M. C. (1998). Sentence comprehension deficits in Alzheimer's disease: a comparison of off-line vs. on-line sentence processing. *Brain and Language* **64**, 297–316.
- Langdon, R., Coltheart, M., Ward, P. B. & Catts, S. V. (2002). Disturbed communication in schizophrenia: the role of poor pragmatics and poor mind-reading. *Psychological Medicine* **32**, 1273–1284.
- Lee, K.-H., Farrow, T. F. D., Spence, S. A. & Woodruff, P. W. R. (2004). Social cognition, brain networks and schizophrenia. *Psychological Medicine* **34**, 391–400.
- Leslie, A. M. & Thaiss, L. (1992). Domain specificity in conceptual development: neuropsychological evidence from autism. *Cognition* **43**, 225–251.
- McCabe, R., Leudar, I. & Antaki, C. (2004). Do people with schizophrenia display theory of mind deficits in clinical interactions? *Psychological Medicine* **34**, 401–412.
- Pickup, G. J. & Frith, C. D. (2001). Theory of mind impairments in schizophrenia: symptomatology, severity and specificity. *Psychological Medicine* **31**, 207–220.
- Povinelli, D. J. & Vonk, J. (2003). Chimpanzee minds: suspiciously human? *Trends in Cognitive Sciences* **7**, 157–160.
- Premack, D. & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences* **1**, 515–526.
- Saxe, R. & Kanwisher, N. (2003). People thinking about thinking people. The role of the temporo-parietal junction in 'theory of mind'. *Neuroimage* **19**, 1835–1842.
- Seyfarth, R. M. & Cheney, D. L. (2000). Social awareness in monkeys. *American Zoologist* **40**, 902–909.
- Tenyi, T., Herold, R., Szili, I. M. & Trixler, M. (2002). Schizophrenics show a failure in the decoding of violations of conversational implicatures. *Psychopathology* **35**, 25–27.
- Tomasello, M., Call, J. & Hare, B. (2003). Chimpanzees understand psychological states – the question is which ones and to what extent. *Trends in Cognitive Sciences* **7**, 153–156.
- Tyler, L. K. (1992). The distinction between implicit and explicit language function: Evidence from aphasia. In *The Neuropsychology of Consciousness* (ed. A. D. Milner and M. D. Rugg), pp. 159–178. Academic Press: San Diego, CA.
- Wellman, H. M., Cross, D. & Watson, J. (2001). Meta-analysis of theory-of-mind development: the truth about false belief. *Child Development* **72**, 655–684.
- 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.