Psychological Medicine, 2000, **30**, 735–738. Printed in the United Kingdom
© 2000 Cambridge University Press

CORRESPONDENCE

To the Editor:

In this letter we would like to draw the attention of clinicians to the possibility of different forms of theory of mind impairment in psychiatric and developmental disorders. We argue that not all psychiatric and/or developmental disorders associated with theory of mind impairment can be appropriately described in terms of diminished capacity to represent the mental states of self and other. There is evidence from studies of schizophrenia that suggests impairment to theory of mind is not an absence of understanding that others have mental states, as has been argued in the past for schizophrenia (Corcoran *et al*. 1995; 1997; Frith & Corcoran, 1996; Sarfati *et al*. 1999) and for autism (Leslie & Frith, 1988; Baron-Cohen, 1989; 1991; Frith, 1989; Leekam & Perner, 1991; Baron-Cohen & Ring, 1994). Individuals, such as those with positive symptom schizophrenia, may have an impaired theory of mind due to an overattribution of knowledge to others that will result in inaccurate inferences about others' mental states. This impairment has been described as 'hyper-theory of mind' (Abu-Akel, 1999).

We propose that hyper-theory of mind may be due to unconstrained generation of hypotheses that impacts either the content or the number of representations generated. On the one hand, the overattribution of knowledge to one's self and other people may result from generating a particular hypothesis about the content of the mind. For example, believing that someone can know what you have in your pocket even if he/she were absent at the time you pocketed the contents. On the other hand, this impairment may be due to difficulties in deciding among several competing hypotheses about the mind. This could result in an increase in the likelihood of choosing an erroneous hypothesis about the content of the mind.

To date, research on theory of mind in psychiatric and developmental disorders has primarily focused on the presence or absence of conceptual abilities or the processes required for the successful application of these conceptual abilities. The research provides several accounts of where the deficit is thought to be. Some researchers have argued the deficit is in representational abilities (i.e. a conceptual deficit) whereby an individual is unaware that others can hold a false belief about a state of affairs (e.g. Frith, 1989; Sodian & Frith, 1992; Bishop, 1993; Baron-Cohen & Ring, 1994; Philips *et al*. 1995; Frith & Corcoran, 1996). Others have argued that the deficit is in application (i.e. deficit in performance) whereby an individual has awareness of other's mental life but fails to demonstrate this knowledge due to processing constraints (e.g. Bachevalier, 1991; Ozonoff *et al*. 1991; Russell *et al*. 1991; Bowler, 1992).

Looking first at the literature on autism, where much of the research on theory of mind deficit has been centered, the representational deficit account argues that individuals with autism do not go through a radical conceptual shift that typically developing children go through at about age four (e.g. Gopnik & Astington, 1988). Perner *et al*. (1989) show that children with autism fail to understand that their own and others' minds can hold false beliefs. Baron-Cohen (1991) later extended this line of research to show that children with autism have a conceptual deficit pertaining to an understanding of appearance and reality. Bishop (1993) entertains the possibility that there is a deficit underlying both theory of mind skills and executive functioning in autism, which is failure to form high-level second order representations of the world.

Where researchers have invoked an application deficit rather than a conceptual deficit it has been argued that individuals with autism have intact theory of mind but they fail to exhibit this ability because they cannot inhibit reality in their answers to false belief questions (Russell *et al*. 1991). Bowler (1992), based on evidence from Asperger's syndrome (a mild form of autism), attributed the application deficit to impairment in specific cognitive processes, such as deficits in accessing stored information

https://doi.org/10.1017/S0033291799002123 Published online by Cambridge University Press

(Ozonoff *et al*. 1991) and in memory (Bachevalier, 1991). Cognitive processes such as these are prerequisite for holding information in short-term memory in order to theorize about the contents of the mind.

The same arguments used to explain theory of mind deficits in autism have been used to explain the failure of schizophrenic patients on theory of mind tasks (e.g. Doody *et al*. 1998; Corcoran *et al*. 1995; Frith & Corcoran, 1996). With few exceptions, these studies argue that the failure of patients with various schizophrenias on theory of mind tasks can be explained by their inability to make inferences about the mental states of others. Memory deficit has also been implicated in the failure of patients with schizophrenia on theory of mind tasks, but found relevant only for patients with negative symptoms (e.g. flat affect, poverty of speech) (Corcoran *et al*. 1995). The memory deficit account, however, cannot explain the impairment to theory of mind in patients with positive symptom schizophrenia (e.g. paranoia, delusions) for whom no memory deficit has been found (Corcoran *et al*. 1995; Frith & Corcoran, 1996). According to Bowler (1992), patients with chronic positive symptom schizophrenia possess knowledge about other people's minds as measured by their success on theory of mind tasks, but they are claimed to be deficient in applying this knowledge during interactions with others in everyday contexts.

If the impairment to theory of mind is one of conceptual deficit or deficit in application, and theory of mind is essential for social interaction and language use, then one would expect language and social dysfunctions in individuals with positive symptom schizophrenia and autism (including Asperger's syndrome) not to differ. However, we believe that there are several arguments for the inadequacy of these accounts. The first, and most salient aspect, is that these disorders exhibit different social and language dysfunctions (American Psychiatric Association, 1994). Secondly, patients with positive symptom schizophrenia contrast with patients with autism or Asperger's syndrome in that they initially develop theory of mind normally without any delay (e.g. Frith, 1992, 1994; Corcoran *et al*. 1995), i.e. these patients do not have impairment to theory of mind until the onset of their illness (typically during late adolescence and early adulthood (American Psychiatric Association, 1994)). Unlike patients with autism, patients with positive symptom schizophrenia would have to lose their representational understanding of the mind rather than never develop it, and, unlike patients with Asperger's syndrome, patients with positive symptom schizophrenia would have to lose their understanding of theory of mind application rules rather than never develop them.

However, there is evidence to suggest that patients with some types of schizophrenia do apply their knowledge of the content of the mind, but in an incorrect or a biased way (Bentall, 1995; Abu-Akel, 1999). For example, a patient with positive symptom schizophrenia of the disorganized type studied by Abu-Akel (1999) demonstrated recognition of others' minds in that the patient would uphold a conversation governed by the rules of discourse such as turn-taking and responding to requests for further information, but the content of the discourse was biased in that the events alluded to were not recoverable by the interlocutor.

We argue that theory of mind impairment, in terms of lack of representational abilities, does not apply to positive symptom schizophrenia and misses the vital distinction between this and other types of schizophrenia, as well as other prominent psychiatric and developmental disorders (e.g. autism). While previous research has focused on deficits either in conceptualization (lack of representational abilities) or in the application of these skills (a non-social cognitive deficit), a hyper-theory of mind explanation allows us to focus on what patients do know of their own and other's minds. It also explains anomalies in the theory of mind functioning mechanism in schizophrenia as a cognitive bias that reflects a predilection to use information about the mind in a particular way (i.e. over-attributing knowledge to self and others). This cognitive bias may underlie the difficulties that paranoid schizophrenics have with social attributions (Bentall *et al*. 1991), social judgement (LaRusso, 1978), and the problems that schizophrenic patients with delusions have in drawing appropriate conclusions from the evidence presented (Garety *et al*. 1991). In sum, we concur with previous findings suggesting that patients with positive symptom schizophrenia do have a conceptual understanding of others' minds (Bowler, 1992), however, rather than

https://doi.org/10.1017/S0033291799002123 Published online by Cambridge University Press

attribute impairment to the inability to apply theory of mind, our hypothesized notion of hyper-theory of mind allows for the possibility that these patients are able to apply theory of mind, albeit atypically.

Additionally, we suspect that there may be overattribution of mental states to self even in the presence of no impairment to the ability to accurately represent mental states for others as is the case of schizophrenics with passivity symptoms (e.g. delusions of control). Work by Frith & Done (1989) and Frith & Corcoran (1996) shows that patients with passivity have a specific impairment in the ability to reason about their own intentions to act but no such impairment in understanding others' mental states.

By invoking the notion of hyper-theory of mind we are able to reframe impairment to mental representation on a continuum, with one variety of theory of mind impairment being characterized as having no representational understanding of mental states, another variety as having representational understanding of mental states, but a deficit in application of this understanding (i.e. due to a reality bias), and a third variety as having representational understanding of mind but overattributing mental states or overgenerating hypotheses about mental life. This renders the dichotomous classic tests of theory of mind (presence or absence of mentalization) inadequate in distinguishing between the reasons why people fail to answer belief questions accurately. It is for future research to define methods to capture this proposed continuum of theory of mind impairments.

To conclude, the notion of hyper-theory of mind may have important implications for distinguishing between disorders in which a theory of mind impairment has been implicated. Specifically, hyper-theory of mind can be a useful conceptual construct in distinguishing between individuals with autism, individuals with positive symptom schizophrenia, and individuals with other schizophrenias (e.g. negative symptom schizophrenia). Furthermore, under the proposed hyper-theory of mind account, autism can be described as the retention of the reality bias beyond early childhood, whereas childhood schizophrenia can best be described as the retention of hyper-theory of mind beyond early childhood. Finally, it is our understanding that any adequate theory about a cognitive faculty not only has to be falsifiable, but has to account for the full range of observed phenomena (i.e lack, presence and *excess* of mentalization). We argue that the notion of hyper-theory of mind provides the additional piece to complete this range of phenomena in the domain of representational understanding of mind.

REFERENCES

Abu-Akel, A. (1999). Impaired theory of mind in schizophrenia. *Pragmatics and Cognition* **7**, 247–282.

American Psychiatric Association (1994). *Diagnostic and Statistical Manual of Mental Disorders, 4th edn*. American Psychiatric Association: Washington, DC.

Bachevalier, J. (1991). An animal model for childhood autism: memory loss and socioemotional disturbances following neonatal damage to the limbic system in monkeys. In Tamminga, C. A., Schulz, S. C. (Eds.), *Advances in Neuropsychiatry and Psychopharmacology* (*Vol 1*). Raven Press, NY, pp. 129–140.

Baron-Cohen, S. (1989). Are autistic children 'behaviorists'? An examination of their mental-physical and appearance-reality distinctions. *Journal of Autism and Developmental Disorders* **19**, 579–600.

Baron-Cohen, S. (1991). The development of theory of mind in autism: deviance or delay? *Psychiatric Clinics of North America* **14**, 33–51.

Baron-Cohen, S. & Ring, H. (1994). A model of the mind reading: neuropsychological and neurobiological perspectives. In *Origins of Understanding Mind* (ed. P. Mitchell and C. Lewis), pp. 183–207. Erlbaum: New Jersey.

Bentall, R. P. (1995). Brains, biases, deficits and disorders. *British Journal of Psychiatry* **167**, 153–155.

Bentall, R. P., Kaney, S. & Dewey, M. E. (1991). Paranoia and social reasoning: an attribution theory analysis. *British Journal of Clinical Psychology* **30**, 13–23.

Bishop, D. M. V. (1993). Annotation: autism, executive functions and theory of mind: a neuropsychological perspective. *Journal of Child Psychology and Psychiatry* **34**, 279–293.

Bowler, D. M. (1992). 'Theory of mind' in Asperger's syndrome. *Journal of Child Psychology and Psychiatry* **33**, 877–893.

Corcoran, R., Cahill, C. & Frith, C. D. (1997). The appreciation of visual jokes in people with schizophrenia: a study of 'mentalizing' ability. *Schizophrenia Research* **24**, 319–327.

Corcoran, R., Mercer, G. & Frith, C. D. (1995). Schizophrenia, symptomatology and social inference: Investigating 'theory of mind' in people with schizophrenia. *Schizophrenia Research* **17**, 5–13.

Doody, G. A., Gotz, M., Johnstone, E. C., Frith, C. D. & Cunningham Owens, D. G. (1998). Theory of mind and psychoses. *Psychological Medicine* **28**, 397–405.

Frith, U. (1989). *Autism: Explaining the Enigma*. Blackwell: Oxford.

Frith, C. D. (1992). *The Cognitive Neuropsychology of Schizophrenia*. Erlbaum: Hove.

Frith, C. D. (1994). Theory of mind in schizophrenia. In *The Neuropsychology of Schizophrenia* (ed. A. S. David and J. C. Cutting), pp. 147–161. Erlbaum: Hove.

Frith, C. D. & Corcoran, R. (1996). Exploring 'theory of mind' in people with schizophrenia. *Psychological Medicine* **26**, 521–530.

Frith, C. D. & Done, D. J. (1989). Experiences of alien control in schizophrenia reflect a disorder in the central monitoring of action. *Psychological Medicine* **19**, 359–363.

Garety, P., Hemsley, D. & Wesseley, S. (1991). Reasoning in deluded schizophrenic and paranoid patients. *Journal of Nervous and Mental Disease* **179**, 194–201.

https://doi.org/10.1017/S0033291799002123 Published online by Cambridge University Press

Gopnik, A. & Astington, J. (1988). Children's understanding of representational change and its relation to the understanding of appearance-reality distinction. *Child Development* **59**, 98–110.
LaRusso, L. (1978). Sensitivity of paranoid patients to nonverbal cues. *Journal of Abnormal Psychology* **87**, 463–471.
Leekam, S. R. & Perner, J. (1991). Does the autistic child have a metarepresentational deficit? *Cognition* **40**, 203–218.
Leslie, A. M. & Frith, U. (1988). Autistic children's understanding of seeing, knowing and believing. *British Journal of Developmental Psychology* **6**, 315–324.
Ozonoff, S., Rogers, S. J. & Pennington, B. F. (1991). Asperger's syndrome: evidence of an empirical distinction from high functioning autism. *Journal of Child Psychology and Psychiatry* **32**, 1107–1122.
Perner, J., Frith, U., Leslie, A. M. & Leekam, S. R. (1989). Extrapolation of the autistic child's theory of mind: knowledge, belief and communication. *Child Development* **60**, 689–700.
Philips, W., Gomez, J. C., Baron-Cohen, S., Laa, V. & Riviere, A. (1995). Treating people as objects, agents or 'subjects': how young children with and without autism make requests. *Journal of Child Psychology and Psychiatry* **36**, 1383–1398.
Russell, J., Mauthner, N., Sharpe, S. & Tidswell, T. (1991). The 'windows task' as a measure of strategic deception in preschoolers and autistic subjects. *British Journal of Developmental Psychology* **9**, 331–349.
Sarfati, Y, Hardy-Baylé Brunet, E. & Widlöcher, D. (1999). Investigation of theory of mind in schizophrenia: Influence of verbalization in disorganized and non-disorganized patients. *Schizophrenia Research* **37**, 183–190.
Sodian, B. & Frith, U. (1992). Deception and sabotage in autistic, retarded and normal children. *Journal of Child Psychology and Psychiatry* **33**, 591–605.

AHMAD ABU-AKEL AND ALISON L. BAILEY

Address correspondence to:
Dr A. Abu-Akel,
Department of Applied Linguistics,
UCLA,
Box 951531, 405 Hilgard Avenue,
Los Angeles, *CA 90095-1531*, *USA*.

Dr. A.L. Bailey,
Department of Education,
UCLA,
Box 951521, 405 Hilgard Avenue,
Los Angeles, *CA 90095-1521*,
USA.

https://doi.org/10.1017/S0033291799002123 Published online by Cambridge University Press