Journal of Autism and Developmental Disorders, Vol. 29, No. 5, 1999

Recognition of Faux Pas by Normally Developing Children and Children with Asperger Syndrome or High-Functioning Autism

Simon Baron-Cohen,[1] Michelle O'Riordan,[1] Valerie Stone,[1] Rosie Jones,[1] and Kate Plaisted[1]

Most theory of mind (ToM) tests are designed for subjects with a mental age of 4–6 years. There are very few ToM tests for subjects who are older or more able than this. We report a new test of ToM, designed for children 7–11 years old. The task involves recognizing faux pas. Study 1 tested 7–9, and 11-year-old normal children. Results showed that the ability to detect faux pas developed with age and that there was a differential developmental profile between the two sexes (female superiority). Study 2 tested children with Asperger syndrome (AS) or high-functioning autism (HFA), selected for being able to pass traditional 4- to 6-year level (first- and second-order) false belief tests. Results showed that whereas normal 9- to 11-year-old children were skilled at detecting faux pas, children with AS or HFA were impaired on this task. Study 3 reports a refinement in the test, employing control stimuli. This replicated the results from Study 2. Some patients with AS or HFA were able to recognize faux pas but still produced them. Future research should assess faux pas production.

KEY WORDS: Theory of mind; faux pas recognition; Asperger syndrome; high-functioning autism.

INTRODUCTION

First-order false belief tasks (e.g., "Sally thinks it's x, when really it's y") or second-order false belief tests (e.g., "Sally thinks Mary thinks x, but both Sally and Mary are wrong") are the main ways in which researchers assess a person's theory of mind (ToM). The main findings are that 4-year-olds pass first-order false belief tasks, whereas 6-year-olds pass second-order false belief tasks (Perner & Wimmer, 1985; Sullivan, Zaitchik & Tager-Flusberg, 1994; Wimmer & Perner, 1983).

Adherence to these sorts of tasks leaves the impression that theory of mind does not develop beyond this level. Yet passing these tests should really be considered as relatively early points in the acquisition of a ToM, rather than the endpoint. Just as in the development of literacy, where learning to read normally happens around 4–6 years of age but a person's reading skills continue to develop until at least adolescence or beyond, so there is a lot of development of "mindreading" beyond 4–6 years old.

One clue that this is true is that there are now three more "advanced" tests of theory of mind (Baron-Cohen, Jolliffe, Mortimore, & Robertson, 1997; Baron-Cohen, Wheelwright, & Jolliffe, 1997; Happé, 1994). Happé's task, for example, involves the detection of sarcasm, bluff, irony, and double-bluff. These are not well understood until 8 years old. Baron-Cohen *et al.*'s tasks involve detecting both basic and complex mental states from the information around people's eyes. This has only been tested in adults. A second clue that ToM develops beyond the 4- to 6-year level is that whereas the majority of children with autism fail first- and second-order false belief tests (Baron-Cohen, 1989a; Baron-Cohen, Leslie, & Frith, 1985; see Baron-Cohen, 1995, for a review), a subgroup of people with autism (often

[1] Departments of Experimental Psychology and Psychiatry, University of Cambridge, Downing Street, Cambridge, CB2 3EB, United Kingdom.

0162-3257/99/1000-0407$16.00/0

more able adolescents and adults) can pass these despite reporting that they have difficulty understanding people's thoughts, intentions, and actions (Bowler, 1992; Dahlgren & Trillingsgaard, 1996; Happé, 1994). There is therefore a clear need for tests of theory of mind which go beyond the 4- to 6-year-old level. The present work describes a new ToM test pitched at the 9- to 11-year-old level. The test comprises assessing the ability of subjects to recognize "faux pas."

Defining a faux pas turns out to be far from straightforward. Socially normal individuals can usually recognize when someone has committed a faux pas, though specifying the necessary and sufficient criteria for this is difficult. A working definition of faux pas might be when a speaker says something without considering if it is something that the listener might not want to hear or know, and which typically has negative consequences that the speaker never intended. Consider an example:

> Steve, a scientist, is traveling on a plane with his wife. Suddenly, he is tapped on the shoulder by another scientist. Steve looks up, sees that he knows this man, and says "Oh hi! How nice to run into you! Let me introduce you to my wife, Betsy. Betsy, this is Jeffrey, a good friend of mine from Harvard days." Betsy says "Oh, hi Jeffrey, pleased to meet you." The other man replies "Er, my name isn't Jeffrey, it's Mike."

What makes Steve's utterance a faux pas is the "uh-oh!" emotion most people would feel if they were in Steve's shoes. Having claimed this is a good friend, it is then a bit insulting to get his name wrong, since the implication is that that this is not such a good or important friend after all. It could be that part of the definition of a faux pas should include this "uh-oh" reaction, but this itself is hard to define. It is a mix of regret, embarrassment, and feeling bad for the listener, perhaps best summed up by the phrase "I wish I hadn't said that!," or by the cartoon gesture of putting one's fist into one's own mouth.

Utterances of this type are suitable for inclusion in a more advanced theory of mind test because detecting a faux pas requires both an appreciation that there may be a difference between a speaker's knowledge state and that of their listener, and an appreciation of the emotional impact of a statement on the listener. A faux pas detection test has not (as far as we know) been previously reported. Note that we report a test of *detecting* faux pas, rather than *producing* faux pas. (The latter would require a completely different design and procedure.)

In the test reported here, a series of stories was played to the subject. For example, in one story, Mary says "I don't think I've met this little boy," to a child's mother. In fact, the child is a little girl. The speaker did not say this out of any malicious intent, but out of a mistaken belief. The faux pas in this case lies in the fact that it may unintentionally upset parents for their little girl to be thought of as a boy. In another story in the test, Tim is in a restaurant and spills his coffee. He turns to the waiter and says, "I've spilt my coffee. Would you be able to mop it up?." In fact, the other person is not a waiter but simply another customer. Once again there is no malice involved. The speaker was simply mistaken. However, the faux pas in this case arises because it is a bit rude to ask a bystander to clean up your mess.

The aim of Study 1 was to test when normal children, selected for being able to pass second-order false belief tasks, recognize faux pas. As well as testing for changes with age on this test, we were also interested in possible sex differences. Specifically, we predicted that females may be superior to males, given other evidence of females being socially more developed than age-matched males (Baron-Cohen & Hammer, 1997; Eibl, Eibesfeldt, 1989; Goodenough, 1957; Halpern, 1992; Happé, 1995; McGuiness & Pribram, 1979). Having carried out the normative study (Study 1), subsequent studies (2 and 3) were planned to include children with Asperger syndrome (AS) or high-functioning autism (HFA).

The stories included in the test were gathered by asking people to give us examples of faux pas incidents from their own experience. Having collected a list of these, they were then presented to a panel of four judges, to establish which stories were categorized as good examples of faux pas by all four judges. On this basis, 10 good stories were included in the test.

STUDY 1: THE NORMAL DEVELOPMENT OF FAUX PAS DETECTION

Method

Participants

Fifty-nine normal subjects were tested in three age groups. To be included in the study the child had to pass two criteria: (a) passing first- and second-order false belief tasks, and (b) answering correctly at least 7 of 10 comprehension questions (listed in Appendix A) to ensure the children understood the stories. Three children were eliminated from the study at this stage due to their failure to reach these criteria. Of the 56 remaining children, 20 were age 7, 20 were age 9, and 16 were 11 years old. The proportions of boys and girls were equal in each age group.

All children were assessed for verbal mental age (VMA) using the British Picture Vocabulary Scale (BPVS) Long Form (Dunn, Dunn, Whetton, & Pintilie, 1982) and for an aspect of non-verbal mental age (non-VMA) using the WISC-R Block Design Task. The BPVS correlates well with tests of more complex language skills, such as syntax, despite only being a measure of single-word comprehension (Jarrold, Boucher, & Russell, 1997). An unpaired t test of chronological age revealed that male and female subjects within each age group did not differ significantly ($p < .05$). The same was true for VMA and Block Design ability (both $p < .05$). These data are presented in Table I.

Materials

Ten faux pas stories were used for the study. Each story involved two or three characters and at least two separate statements. The stories are listed in Appendix A. The language used was simple so that it could be understood by young children. In addition the stories were designed so that the faux pas occurred either in the last phrase, one phrase before last, or two phrases before the end on an equal number of occasions. This was to ensure that a child could not pass by simply quoting the last phrase heard (parroting) or using some similar strategy.

The stories were recorded onto TDK60 tape. An audio cassette was used in preference to a video, to avoid the possibility of facial expressions giving direct feedback about the emotional response of the recipient of the faux pas (Keasey, 1977). If a video had been used the subjects may have detected the faux pas as a result of recognizing the key, relevant emotional expression (such as horror, shame, or embarrassment). Furthermore, there was no explicit reference to the characters' reactions, so that subjects had to detect faux pas without explicit cues, just as in real life. The stories were recorded in a soundproof room at the speed of normal speech. One narrator was used throughout, and the remaining parts were read by various people. The parts were all clearly spoken without distinctive accents or speech impediments which may have affected their salience. The narratives were played to the subjects on a Dictaphone held close to their ear in a quiet room in their school.

In between the stories the following questions were used to assess the child's understanding.

Faux Pas Detection Question. In the story did someone say something that they should not have said?

Identification Question. What did they say that they should not have said?

Comprehensive Question. This question was different for each story. (See Appendix A.)

False Belief Question. Did they know/remember that ? (See Appendix A.)

The first question assessed whether the child had detected whether a faux pas was present. The second question ensured that the subject had identified the correct utterance as the faux pas. The third question ensured that the child had understood the story and was paying attention, so that failure of a faux pas question could not be due to verbal comprehension problems or distraction. The final question checked that the child understood the faux pas was a consequence of the speaker's false belief rather than being an action with malicious intent.

Procedure

The test session began by giving the subjects first- and second-order false belief tasks. The first-order task was the traditional "Sally-Anne" task (Baron-Cohen *et al.*, 1985). The second-order task was a modification of the "Sally-Anne" task which allowed testing of a second-order false belief (suggested by Angel Rivière, personal communication). This modification was that when Sally departs, she looks back through the keyhole while Anne is moving the marble. When Sally

Table I. Mean Chronological Age, Verbal Mental Age, and Block Design Verbal Mental age for All Subjects

			CA (months)		VMA (months)		Block Design (months)	
Group	*n*	Age (years)	*M*	*SD*	*M*	*SD*	*M*	*SD*
Girls	(10)	7	88.5	5.0	92.7	12.2	100.6	30.0
Boys	(10)	7	87.8	7.0	89.3	13.2	107.6	30.1
Girls	(10)	9	109.8	4.0	108.3	12.6	108.6	22.0
Boys	(10)	9	110.4	5.2	105.5	21.5	130.0	35.6
Girls	(8)	11	136.3	5.0	141.0	28.3	138.3	32.8
Boys	(8)	11	135.4	3.4	143.0	32.4	141.0	30.2

returns the test question is no longer "Where will Sally look for the marble?", but rather "Where does Anne think Sally will look for the marble?" These ToM tests were administered so as to check that any failure on the faux pas test was not due to a ToM deficit at an equivalent of a 4- to 6-year-old level. Then the Block Design test and BPVS were administered. These latter tests were to have a brief measure of cognitive and language function.

The final stage of the experiment involved playing the tape of faux pas stories to the children. The Experimenter introduced the set of stories by saying "Now I'm going to play you some stories. I want you to listen very carefully because afterwards I am going to ask you some questions to see what you think of them. Are you ready?" Initially, a practice story containing no faux pas, was played to the subject after which a simple comprehension question was asked. This allowed the child to get used to the Dictaphone, for the volume to be adjusted if necessary, and to act as control stories. The 10 faux pas stories were presented in two experimental orders, either in the order they appear in Appendix A, or in the reverse order, to avoid possible order effects. This was counterbalanced in each sex and age group.

After each faux pas story the subject was asked the four questions. If the child answered the first question incorrectly then the second question was not asked (i.e., if the child said no one had said something they should not have said, then they were not asked what should not have been said). Positive comments were made throughout the testing session to encourage the subject, but no feedback was given about the correctness of their answers. No additional prompts were given and repetition of a story was allowed only if interruptions occurred. In practice this was not needed.

Scoring

The children were given 1 point for each faux pas they identified correctly. To detect a faux pas the child had to answer all the questions correctly, that is, identify that someone had said something that they should not have, identify what it was that they should not have said, or some approximation of it, answer a comprehension question, and recognize that the faux pas was a consequence of a false belief. Failure of any of these questions led to a score of zero for that story.

It was not possible to detect a faux pas by using some simple strategy. Although the correct answer to Question 1 was always "yes," the correct answers to Questions 2 and 3 were parts of the narrative which varied in position between stories, and the correct answer to Question 4 was always "no." Furthermore, the child was given no feedback about the nature of their response and so had no basis on which to form a strategy. Detecting a faux pas by chance therefore had a very small probability. The child would have to answer "yes" to the first question ($p = .5$), then pick the correct utterance of which there were a minimum of two in each story ($p = .5$), recall a part of the story to answer the control question (this cannot be assigned a chance probability) and finally answer "no" to the belief question ($p = .5$). Therefore the probability of detecting a faux pas by chance alone was $p < .125$. Hence, the number of faux pas that could be detected by chance would be < 1.25 stories.

Results

For all analyses, a criterion p value of .05 was used to establish statistical significance. The faux pas scores were analyzed by a 3 × 2 ANOVA with Age (7, 9, or 11) and Sex (boy or girl) as factors. This analysis showed significant main effects of age, $F(2, 50) = 22.03$; and sex, $F(2, 50) = 5.33$. Neuman-Keuls pairwise comparisons revealed that faux pas scores were significantly higher in the 11-year-olds than the two younger age groups and that the scores for the 9-year-olds were higher than those of the 7-year-olds ($p < .05$). They also revealed that as predicted, faux pas scores were higher for female than male subjects ($p < .05$).

Although the overall interaction failed to reach significance at the .05 level, $F(1, 50) = 1.08$, $p = .349$, improvement in task performance between 7 and 9 years was greater for the girls than the boys and that boys showed a more dramatic improvement in performance than girls between 9 and 11 years.

The fact that, overall, girls' performance was statistically better than boys' merited further analysis. Therefore one-sample t tests were conducted for each age and sex group to see whether their mean scores differed from chance. It was found that all groups scored significantly higher than chance at the 98% confidence level except for the 7-year-old boys. The mean scores and standard deviations of each age and sex group are shown in Table II.

The data were analyzed to identify which part of the task the children were answering incorrectly. Paired t tests for each age and sex group compared the total number of failed faux pas with the number of faux pas stories associated with a failed false belief question. This revealed that significantly more faux pas were failed than were associated with incorrectly answered false belief questions. The t-test results are listed here:

Table II. Mean Score on the Faux Pas Test, Study 1

Group	Age (years)	Faux pas score		
		M	*SD*	Range
Girls	7	3.8	2.20	1–7
Girls	9	7.2	1.99	3–9
Girls	11	8.5	0.93	7–10
Boys	7	2.9	2.60	0–7
Boys	9	4.6	2.80	0–9
Boys	11	7.9	2.03	5–10

Girls: 11 years, $t(7) = 3.74$; 9 years, $t(9) = 3.37$; 7 years, $t(9) = 8.33$; Boys: 11 years, $t(7) = 3.0$; 9 years, $t(9) = 6.33$; 7 years, $t(9) = 4.67$. This suggests that a failure to detect faux pas was not due to a first-order deficit in mental state attribution.

Paired *t* tests were also used to compare the number of failed faux pas with the number of failed comprehension questions for each group of age and sex. This revealed that were significantly more failed faux pas than there were stories associated with incorrectly answered comprehension questions. The *t*-test results are listed here: Girls: 11 years, $t(7) = 4.58$; 9 years, $t(9) = 4.26$; 7 years, $t(9) = 8.11$; Boys: 11 years, $t(7) = 2.70$; 9 years, $t(9) = 5.92$, 7 years, $t(9) = 8.30$. This shows that failure to detect faux pas was equally not due to lack of understanding or memory.

Pearson's correlational analysis of VMA, Block Design scores, and faux pas scores for each subject revealed a positive correlation ($r = .5$) between BPVS and Block Design. Furthermore, the VMA and faux pas scores were positively correlated ($r = .52$). This is perhaps not surprising as faux pas detection is a verbal task. However, there was no correlation between performances on the spatial task and faux pas tasks ($r = .26$). This suggests that the faux pas task is testing more than the general intelligence of the subject.

Discussion

Study 1 assessed the performance of normal children on a more advanced theory of mind test, involving faux pas detection. The results show a differential developmental profile for male and female subjects. As predicted, girls were more advanced than boys on this task, with girls performing above chance at age 7 and boys at age 9. Given that all children were selected on the basis of being able to pass second-order false belief tests, these results confirm that faux pas detection is more complex than standard false belief tests.

It could be argued that the faux pas task merely assesses verbal ability. It is true that there is a correlation between VMA and performance on the faux pas detection task, but there are many reasons for thinking that the faux pas task measures more than verbal ability. First, it is not surprising to find a correlation between faux pas scores and VMA, since VMA increases with chronological age. Indeed, we would expect a correlation with any other ability that also increases with age. The correlation between the faux pas and VMA merely indicates that one increases as the other does—it obviously does not demonstrate a causal direction. Second, children had to pass a minimum of 7 out of 10 comprehension questions to be included in the study. This suggests that all the children could cope with the syntactic demands of the stories. Failure on the faux pas task cannot thus be attributed to insufficient linguistic ability per se.

Having established normal age and gender patterns of performance on this new test, Study 2 tested children with Asperger syndrome or high-functioning autism, some of whom pass first- and second-order false belief tests. If they have a delayed development of ToM, then this might only show up by testing them with a more advanced, age-appropriate test of this.

STUDY 2: THE DETECTION OF FAUX PAS BY CHILDREN WITH ASPERGER SYNDROME OR HIGH-FUNCTIONING AUTISM

Some authors consider that autism lies on a continuum, with AS lying closer to the normal end (Wing, 1997). Individuals with AS show the triad of impairments typical in autism, but have no associated mental handicap and no history of cognitive and language delay (ICD-10; World Health Organization [WHO], 1994). However, in both DSM-IV (American Psychiatric Association [APA], 1994) and ICD-10 the hierarchical decision rule states that a person cannot be diagnosed with AS if they have at any time met criteria for autism. This means that most cases of apparent AS may in fact have autism, since one can meet the crieria for autism if one shows onset of behavioral difficulties in social reciprocity, communication in any form (e.g., nonverbal), or difficulties in play, or restricted and repetitive interests prior to the age of 3 years old. For this reason, it has proven difficult to distinguish AS from HFA both diagnostically and experimentally, since both of these may have no associated mental handicap, but would meet these criteria for autism. In the studies to be reported next, we therefore use the terms AS and HFA without any attempt to argue for a distinction of these.

A subgroup of children with autism and AS clearly appreciate first-order false beliefs (Baron-Cohen, 1989b) and furthermore, some of these individuals also appreciate second-order false beliefs (Bowler, 1992; Dahlgren & Trillingsgaard, 1996; Happé, 1994). Ozonoff, Pennington, and Rogers interpreted this as showing that ToM problems are not universal in autism and AS (Ozonoff, Rogers & Pennington, 1991). However, the ToM deficit in HFA or AS may be masked by only using the existing first- and second-order false belief tests. The results from Baron-Cohen, Jolliffe, *et al.*, (1997), Baron-Cohen, Wheelwright, *et al.*, (1997), and Happé, (1994) suggest that even individuals with HFA or AS have deficits when tested on *advanced* ToM tasks.

In Study 2 reported here, we extended the advanced theory or mind studies by employing the faux pas test reported in Study 1. We predicted that children with AS or HFA who pass false-belief tasks would nevertheless fail the faux pas detection task, despite having an adequate mental age.

Method

Participants

Sixteen normal children and 12 subjects with AS or HFA were tested. To be included in the study the child had to pass two criteria: (a) passing first- and second-order false belief tasks and (b) answering correctly at least 7 of the 10 comprehension questions (listed in Appendix A) to ensure a general level of understanding of the stories. One child from the AS/HFA group was eliminated from the study at this stage due to failure to reach these criteria. Of the remaining subjects the male to female ratio was 1:1 in the control group, and 9:2 in the AS/HFA group.

Diagnosis of the children with AS or HFA was carried out by an independent clinician with extensive experience in this area (Patrick Bolton) using ICD-10 (WHO, 1994) criteria. They were recruited via the clinic in Cambridge for Communication Disorders. Age of onset of language data showed that in all cases, these children either had single words by 2 years old or phrase speech by 3 years old, so that technically there was no language delay. However, in most cases there was evidence of other communication anomalies prior to age 3 years (e.g., delays in protodeclarative pointing). There was no history of general cognitive delay in these cases.

As in Study 1, the children were assessed for VMA using the British Picture Vocabulary Scale (BPVS) Long Form (Dunn *et al.*, 1982) and for an aspect of non-VMA using the WISC-R Block Design Task. An unpaired *t* test of chronological age revealed that the two groups did not differ significantly ($p > .05$). The same was true for VMA ($p < .05$). Two individuals in the group of subjects with AS/HFA performed at ceiling on the Block Design Task and were therefore given a non-VMA of greater than 202 months. This is not uncommon in autism spectrum conditions (Happé, 1996). Since this value cannot be incorporated in a statistical analysis, a score of 202 months was assigned to these subjects. The analysis of the Block Design data may thus be a conservative one. The mean block design score of the subjects with AS/HFA was higher than the control subjects although this difference was not significant, $t(25) = 1.31$ $p = .2$. Matching was thus based on group mean data. These data are presented in Table III.

Materials and Procedure

This was identical to Study 1.

Results

Performance is summarized in Table IV. For all analyses, a criterion p value of .05 was used to establish statistical significance. The scores were analyzed by a one-way ANOVA with Group as the factor. This revealed that the subjects with AS/HFA were significantly impaired relative to the normal subjects, $F(1, 25) = 17.77$.

The scores from the subjects with AS/HFA were compared to that of the boys from the control sample. This was done to ensure that our finding was not attributable to the fact that there was a higher male to female ratio in the AS/HFA group than in the control sample. The scores were analyzed by a one-way ANOVA

Table III. Mean Chronological Age, Verbal Mental Age, and Block Design Mental Age for All Children in Study 2

		CA (months)		VMA (months)		BDMA (months)	
Group	*n*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Control	16	135.9	15.1	142.0	46.6	140.0	44.7
AS/HFA	11	144.0	32.9	159.0	53.8	160.5	51.7

Table IV. Mean Score on Faux Pas Test, Study 2

	Faux pas score		
Group	*M*	*SD*	Range
Control	8.2	1.56	5–10
AS/HFA	4.9	2.73	1–9

with Group as the factor. This revealed that the subjects with AS/HFA were still significantly impaired relative to the normal boys, $F(1, 17) = 7.62$.

Further analysis investigated whether failure of faux pas detection in either the AS/HFA group or the control group could be attributed to difficulties in belief attribution. *T* tests showed that failure to detect faux pas was not due to a basic lack of appreciation of mental states, as the false belief questions were answered correctly for a significant proportion of the failed faux pas, AS/HFA group, $t(10) = 6.53$; Control group, $t(15) = 4.44$.

Pearson's correlation analysis of the faux pas scores with the VMA scores and the Block Design scores were conducted for each subject. There was no correlation between VMA and the faux pas scores in either group (AS/HFA group, $r = .27$; Control group, $r = -.18$). Furthermore, there was no correlation between performance on the Block Design task and the faux pas task (AS/HFA group, $r = .22$; Control group, $r = .12$). As in Study 1, the lack of correlation between faux pas scores and performance on the IQ subtest is an important finding since it demonstrates that the faux pas task is testing more than general intelligence. Further, the absence of a correlation between faux pas performance and VMA shows that in Study 2, results are not an artifact of verbal ability.

There were no items that proved to be significantly more difficult than others in the test, for either group. Also, for the subjects with AS/HFA, there was no trend towards a subgroup of high and low scorers. For all subjects, performance was simply depressed relative to controls. If we define a Pass on the test as a whole as scoring equal to or above 8 out of 10, then 2 out of 11 children with AS/HFA passed (18%), versus 12 out of 16 normal children (75%). This difference is highly significant ($\chi^2 = 64.84$, $p < .01$).

Discussion

Study 2 assessed the performance of children with AS/HFA and a matched control group of normal children on a faux pas detection task. The results show that the subjects with AS/HFA performed at a significantly lower level on the task in comparison to the control group. This confirms our prediction that despite passing the false belief tasks, individuals with AS/HFA do have a deficit in theory of mind. The sample of 11 children with AS/HFA inevitably means that we must be cautious about generalizing these findings until replications have been attempted. However, finding a significant difference even with such a small sample indicates that this difference is likely to be robust.

Although the faux pas task is a verbal task there are several reasons why we can conclude that the differential profile of results in the two groups is not an artifact of verbal ability. First, the group of individuals with AS/HFA and the control group were not significantly different in terms of their VMA. Second, there was no correlation between VMA and faux pas performance in either group. Furthermore, the AS/HFA group reached the same level of performance as the control subjects on the Comprehension and Belief Questions. Note that many of the Comprehension and Belief Questions are just as long and as embedded (syntactically) as the Faux Pas Questions. Although we think the Faux Pas Questions may differ in *psycholinguistic* terms, we do not think they are harder than these Control Questions in purely *linguistic* terms.

One criticism of the design of the test as used above is that all 10 stories involve a faux pas, so that a subject could in principle appear to pass the first of the four questions on each trial just because of a "yes" bias. An even stronger version of the test would therefore include an equal number of control stories where no faux pas occurred. In Study 3 we report data from a modified test, from a new sample of children with AS/HFA, in order to check test if the results from Study 2 replicate under these more stringent conditions.

A second potential confound in Study 2 is that the group with AS/HFA was predominantly male, while in the control group the sex ratio was 1 : 1. Given that Study 1 showed that male subjects were worse than female subjects on this task, it could be argued that the result from Study 2 merely reflects the differential proportions of male and female subjects in each group. There are two reasons we consider this is not the case. First, our earlier study showed that while girls and boys ages 7 and 9 performed differently on the faux pas task, there was no significant difference in the performance levels of 11-year-old girls and boys. Our control sample were 11 years of age and therefore the girls and boys in this group would be expected to have reached the same level of performance. Second, when the data for the subjects with AS/HFA were compared to that of the normal 11-year-old boys, the AS/HFA group were still relatively impaired. However, to be cautious, in Study 3, the groups were matched exactly for sex (all male).

STUDY 3: THE MODIFIED FAUX PAS TEST, WITH CHILDREN WITH AS/HFA

Study 3 employed a modified test of Faux Pas detection, using the same 10 stories which contained a Faux Pas (from Studies 1 and 2), but with these randomly interleaved with 10 stories which did not contain a Faux Pas, to test if the earlier deficit identified in children with AS/HFA could be replicated while addressing a weakness of the original version.

Participants

A new set of 15 children with AS/HFA, diagnosed using the same criteria (ICD-10) as used in Study 2 but attending a school for AS, were tested. These children had been referred to this school from different regions of the country. Their CA and IQ (as reported in school notes) is shown in Table V. IQ in all cases was based on the WISC-R, and had been assessed at the point of school entry, and within the last 3 years. This shows that all the subjects with AS/HFA had an IQ in the normal range, defined as over 85. IQ was assumed to be normal in the control group. School notes also confirmed that in all cases there was no history of language or general cognitive delay, but for the reasons discussed earlier, this was insufficient as a criterion for distinguishing AS from HFA. As before, therefore, these two terms are used without distinction. A new sample of 15 normal controls was also tested, selected to be younger than the group with AS/HFA, so that if anything, the group with AS/HFA would have an advantage. Independent t tests showed that the two groups differed significantly in terms of age, $t(28) = 9.51$, $p = .001$, two-tailed. All subjects were male.

Method

The same procedure as was used in Study 2 was used here, but the 10 control stories shown in Appendix B were interleaved randomly onto the audiotape, resulting in a mix of 10 faux pas and 10 control stories being presented. Two orders were presented (one the reverse of the other) in a counterbalanced design, to avoid order effects. Children were told that some stories might sound similar so they should listen very carefully before answering the question after each story. Comprehension questions were asked after each story, as in Study 2.

Table V. Mean Chronological age and IQ for Children in Study 3

	CA		IQ	
Group[a]	*M*	*SD*	*M*	*SD*
Control	10.47	0.22	—	—
AS/HFA	13.35	1.18	96.93	0.82

[a] 15 in each group.

Table VI. Mean Score on Faux Pas Stories in Study 3

	Faux pas score		
Group	*M*	*SD*	Range
Control	9.2	0.94	7–10
AS/HFA	6.6	2.03	4–10

Results

The two groups were at ceiling on all Belief and Comprehension questions (maximum marks for both). Table VI shows mean number of correct scores for the 10 faux pas stories. Table VII shows the mean number of correct scores for the 10 control stories. Repeated-measures ANOVA showed an effect of Group, $F(1, 28) = 13.56$, $p = .001$; and an effect of Type, $F(1, 28) = 23.75$, $p = .0001$; and a Group × Type interaction, $E(1, 28) = 19.54$, $p = .0001$. Post-hoc Neuman-Keuls tests revealed this was due to the group with AS/HFA performing significantly lower on the Faux Pas than on the Control Stories, relative to the normal group. As can be seen, the two groups did not differ on the Control Stories, both being at ceiling.

DISCUSSION OF STUDY 3 AND GENERAL DISCUSSION

Study 3 replicates the results of Study 2, using an even more stringent method. While the children with AS/HFA could identify that no one had said anything wrong in the control stories, at a level equivalent to

Table VII. Mean Score on Control Stories in Study 3

	Faux pas score		
Group	*M*	*SD*	Range
Control	9.3	0.82	8–10
AS/HFA	9.3	0.82	8–10

matched normal controls, they again showed a statistically significant impairment in identifying when someone had committed a faux pas. This modified method rules out any explanation of the results in terms of a general response bias. Therefore there appears to be a genuine deficit in the performance of the group with AS/HFA on the faux pas task. We can conclude that individuals with AS/HFA have a deficit in the utilization of mental state knowledge.

One possible explanation for this deficit is derived from weak central coherence theory (Frith, 1989): that although children with AS/HFA can identify the mental states of the characters individually, they find it hard to bring all of this information together to make a coherent picture of the faux pas and its possible impact. This is plausible, and needs to be tested against a specific theory of mind theory in future studies. The executive dysfunction theory of autism (Ozonoff, 1995; Ozonoff, Pennington, & Rogers, 1991; Pennington & Ozonoff, 1996) has more difficulty in providing a plausible explanation for the present results. Executive function incorporates the abilities of planning, flexibility, inhibition, and working memory. While individuals with autism may not have problems in working memory (Russell, Jarrold, & Henry, 1996) there is evidence to suggest that there may be problems in the other areas of executive functioning such as planning (Hughes, Russell, & Robbins, 1994; Ozonoff, Pennington, & Rogers, 1991). However, while a deficit in planning might be expected to lead to a higher rate of *producing* faux pas, it does not explain why individuals with autism or AS should be poor at faux pas *detection.* There is evidence that individuals with autism have a deficit in inhibition (Hughes & Russell, 1993; Hughes *et al.*, 1994; McEvoy, Rogers, & Pennington, 1993) and again this could certainly account for the production of faux pas. However, a child suffering from disinhibition would not be likely to miss detecting faux pas, since they would be likely to answer the question, "Did someone say something that they shouldn't have said?" (the faux pas identification question) by saying "Yes." In fact, the children with AS/HFA tended to underdetect faux pas, answering this by saying "No." However, some children with AS/HFA could detect faux pas within normal levels, and yet were observed to commit faux pas in their everyday behavior. Future research should also consider methods for assessing this productive deficit.

In conclusion, the performance of individuals with AS/HFA on the faux pas detection task is an experimental demonstration of their ToM deficit at a higher level than either first- or second-order false belief tasks. Whether this occurs for purely ToM-related reasons, or reasons related to central coherence, requires further research. Equally, the extent to which this deficit is specific to AS/HFA requires further testing of other clinical groups. We have recently applied an adult version of this test to patients with orbitofrontal cortex and amygdala lesions, and demonstrated similar impairments to those reported in the patients with Asperger syndrome, (Stone, Baron-Cohen, & Knight, 1998; Stone, Baron-Cohen, Young, Calder, & Green, 1999). We consider the Faux Pas test to be a first step towards creating methods for assessing the use of a theory of mind in increasingly naturalistic social settings.

ACKNOWLEDGMENTS

Experiments 1 and 2 reported here were carried out by M. A. O'Riordan and R. Jones as part fulfillment of a Part II Project in Experimental Psychology, University of Cambridge. M. A. O'Riordan and S. Baron-Cohen were supported by the MRC during the period of this work. We are grateful to Southlands School, Hampshire, for their assistance in Study 3, and to Patrick Bolton and St. Joseph's RC Primary School, Upminster, Essex; Edwin Lambert Primary School, Hornchurch, Essex; Queen Edith County Primary, Cherry Hinton, Cambridge; The Marlborough Unit, Rainham, Gillingham Kent; Hope Lodge School; and Southampton schools for their assistance in Studies 1 and 2.

APPENDIX A: FAUX PAS STORIES (STUDIES 1–3)

Following each story, the Comprehension and False Belief Questions are shown.

1. All of the class took part in a story competition. Emma really wanted to win. Whilst she was away from school, the results of the competition were announced: Alice was the winner. The next day, Alice saw Emma and said "I'm sorry about your story." "What do you mean?" said Emma. "Oh nothing," said Alice.

   Who won the story competition?
   Did Alice realize that Emma hadn't heard the results of the competition?

2. Robert had just started at a new school. He said to his new friend, Andrew, "My Mum is a dinner lady at this school." Then Claire came over and said, "I hate dinner ladies. They're horrible." "Do you want

to come and play rounders?" Andrew asked Claire. "No" she replied "I'm not feeling very well."

What job does Robert's Mum do?
Did Clare know that Robert's Mum was a dinner lady?

3. Mike was in one of the cubicles in the toilets at school. Joe and Peter were at the sinks nearby. Joe said "You know that new boy in the class, his name is Mike. Doesn't he look really weird!" Mike then came out of the cubicles. Peter said "Oh hello Mike, are you going to play football now?".

Where were Joe and Peter when they were talking?
Did Joe Know that Mike was in the cubicles?

4. Kim helped her Mum make an apple pie for her uncle when he came to visit. She carried it out of the kitchen. "I made it just for you," said Kim. "Mmm", replied Uncle Tom, "That looks lovely. I love pies, except for apple, of course!"

What kind of pie had Kim made?
Did Uncle Tom know that the pie was an apple pie?

5. James bought Richard a toy airplane for his birthday. A few months later, they were playing with it, and James accidentally dropped it. "Don't worry" said Richard, "I never liked it anyway. Someone gave it to me for my birthday."

What did James give Richard for his birthday?
Did Richard remember James had given him the toy airplane for his birthday?

6. Sally has short blonde hair. She was at her Aunt Carol's house. The doorbell rang. It was Mary, a neighbor. Mary said "Hello," then looked at Sally and said "Oh, I don't think I've met this little boy. What's your name?" Aunt Carol said "Who'd like a cup of tea?."

Whose house was Sally at?
Did Mary know that Sally was a little girl?

7. Mrs. West, the teacher, had something to tell her class, "One of the boys in our class, Simon, is very seriously ill" she said. The class were all very sad and were sitting quietly when a little girl, Becky, arrived late. "Have you heard my new joke about sick people?" she asked. The teacher said to her "Sit down and get on with your work."

What did the teacher tell the class at the beginning of the story?
Did Becky know Simon was sick?

8. Tim was in a restaurant. He spilt his coffee on the floor by accident. Jack was another person in the restaurant, standing by the cash desk waiting to pay. Tim went up to Jack and said "I'm terribly sorry, but I've spilt my coffee. Would you be able to mop it up?".

Where did the story take place?
Did Tim know Jack was a customer?

9. Jill had just moved into a new house. She went shopping with her Mum and bought some new curtains. When Jill had just put them up, her best friend Lisa came round and said, "Oh, those curtains are horrible, I hope you're going to get some new ones." Jill asked, "Do you like the rest of my bedroom?."

What had Jill just bought?
Did Lisa know the curtains were new?

10. Helen's mum was having a surprise party for Helen's birthday. She invited Nicky and said, "Don't tell anyone, especially Helen!" The day before the party Nicky and Helen were playing together and Nicky ripped her new dress. "Oh!" said Nicky, "I was going to wear this to your party." "What party?" said Helen. "Come on," said Nicky "Let's go and see if my mum can mend the tear."

Who was the surprise party for?
Did Nicky remember the party was a surprise?

APPENDIX B: TEN CONTROL STORIES (STUDY 3 ONLY)

Comprehension and False Belief Questions are shown following each story.

1. All of the class took part in a poetry competition. Jane really wanted to win. While she was away, the results of the competition were announced: Mary was the winner. The next day, Jane bumped into Mary. Mary said "How are you feeling?". "Fine thanks?" said Jane, "Oh good" said Mary.

Who won the story competition?
Did Mary know that Jane hadn't heard the results of the competition?

2. David had just started at a new school. He said to his new friend, Mike, "My Mum is a teacher in this school." Then Jeff came over. "I hate school" he told them, "It's so small." "Do you want to come and play rounders?" Mike asked Jeff. "No" he replied "I'm not feeling very well."

What job does David's Mum do?
Did Jeff know that David's Mum was a teacher?

3. John was in one of the cubicles in the toilets at school. Sam and Eddy were at the sinks nearby. Sam said "You know that new boy in the class—you know, his name is John. Doesn't he look cool!" John then came out of the cubicles. Peter said "Oh, hi John. Are you going to play football now?"

Where were Sam and Eddy when they were talking?
Did Sam know that John was in the cubicles?

4. Kate helped her Mum make a fruit pie for her neighbor when he came to visit. She carried it out of the kitchen. "I made it just for you," said Kate. "Mmm", replied her neighbor, "That looks lovely—I love pies, especially fruit ones!"

What kind of pie had Kate made?
Did the neighbor know that the pie was a fruit pie?

5. Simon bought Robert a toy car for his birthday. A few months later, they were playing with it, and Simon dropped it. "Don't worry," said Robert, "it was only an accident".

What did Simon give Robert for his birthday?
Did Simon know Robert had given him the toy car for his birthday?

6. Jill has short brown hair. She was at her Uncle Ted's house. The doorbell rang. It was Mrs. Smith, a neighbor. Mrs. Smith said "Hello," then looked at Jill and said "Oh, I don't think I've met this little girl. What's your name?" Uncle Ted said "Who'd like some tea?".

Whose house was Jill at?
Did Mrs. Smith know that Jill was a little girl?

7. Mrs. Jones, the teacher, had something to tell her class, "You all know George in our class, well, he is very seriously ill" she said. The class were all very sad and were sitting quietly when a little girl, Amy, arrived late. "I've just been to the dentist" she said. The teacher said to her "Sit down and get on with your work."

What did the teacher tell the class at the beginning of the story?
Did Amy know George was sick?

8. Alan and Ed were having a drink in a restaurant. Alan spilt his coke on the floor by accident. He said to Edward "Oh dear, haven't I been clumsy—I've spilt my coke!" Edward said "I'll order another one."

Where did the story take place?
Did Alan know Ed was a customer?

9. Michelle had just moved into a new house. Michelle went shopping with her Mum and bought a new rug for her bedroom. When Michelle had just put it down, her best friend, Samantha, came round and said, "Oh, your new rug is just like my new one." Michelle asked, "Do you like the house?"

What had Michelle just bought?
Did Samantha know the rug was new?

10. Annette's mum was having a surprise party for Annette's birthday. She invited Bridget and said to her. "Don't tell anyone, especially not Annette!" The day before the party Annette and Bridget were playing together and Annette ripped her new shirt. "Oh!" said Annette, "My mother will be cross." "Don't worry" said Bridget. "Let's go and see if my mum can mend the tear."

Who was the surprise party for?
Did Annette know there was a surprise party planned?

REFERENCES

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed). Washington, DC: Author.

Baron-Cohen, S. (1989a). The autistic child's theory of mind: A case of specific developmental delay. *Journal of Child Psychology and Psychiatry, 30,* 285–298.

Baron-Cohen, S. (1989b). Thinking about thinking: how does it develop? Critical Notice. *Journal of Child Psychology and Psychiatry, 30,* 931–933.

Baron-Cohen, S. (1995). *Mindblindness.* Boston: MIT Press/Bradford Books.

Baron-Cohen, S., & Hammer, J, (1997). Is autism an extreme form of the male brain? *Advances in Infancy Research, 11,* 193–217.

Baron-Cohen, S., Jolliffe, T., Mortimore, C., & Robertson, M. (1997). Another advanced test of theory of mind: Evidence from very high functioning adults with autism or Asperger Syndrome. *Journal of Child Psychology and Psychiatry, 38,* 813–822.

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., Wheelwright, S., & Jolliffe, T. (1997). Is there a "language of the eyes"? Evidence from normal adults and adults with autism or Asperger syndrome. *Visual Cognition, 4,* 311–331.

Bowler, D. M. (1992). Theory of mind in Asperger syndrome. *Journal of Child Psychology and Psychiatry, 33,* 877–895.

Dahlgren, S., & Trillingsgaard, A. (1996). Theory of mind in non-retarded children with autism and AS. A research note. *Journal of Child Psychology and Psychiatry, 37,* 759–763.

Dunn, L. M., Dunn, L. M., Whetton, C., & Pintilie, D. (1982). *British Picture Vocabulary Test.* London: NFER-Nelson.

Eibl-Eibesfeldt, J. (1989). *Human ethology.* Athene de Gruyter.

Frith, U. (1989). *Autism: explaining the enigma.* Oxford: Basil Blackwell.

Goodenough, E. (1957). Interest in persons as an aspect of sex difference in the early years. *Genetic Psychology Monographs. 55,* 287–323.

Halpern, D. (1992). *Sex differences in cognitive ability.* Hillsdale, NJ: Erlbaum.

Happé, F. (1994). An advanced test of theory of mind: Understanding of story characters' thoughts and feelings by able autistic, mentally handicapped, and normal children and adults. *Journal of Autism and Development Disorders, 24,* 129–154.

Happé, F. (1995). The role of age and verbal ability in the theory of mind task performance of subjects with autism. *Child Development, 66,* 843–855.

Happé, F. (1996). *Autism.* London: UCL Press.

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.

Hughes, C., Russell, J., & Robbins, T. (1994). Specific planning deficit in autism: Evidence of a central executive dysfunction. *Neuropsychologia, 32,* 477–492.

Jarrold, C., Boucher, J., & Russell, J. (1997). Language profiles in children with autism: theoretical and methodological implications. *International Journal of Research and Practice; Autism. 1,* 57–76.

Keasey, C. B. (1977). Children's developing awareness and usage of intentionality and motives. *Nebraska Symposium on Motivation* (pp. 219–260). Lincoln: University of Nebraska Press.

McEvoy, R. E., Rogers, S. J., & Pennington, B. F. (1993). Executive function and social communication deficits in young autistic children. *Journal of Child Psychology and Psychiatry, 34,* 563–578.

McGuiness, D., & Pribram, K. H. (1979). The origins of sensory bias in the development of gender differences in perception and cognition. In M. Borner (Ed.), *Cognitive growth and development: Essays in memory of Herbert G Birch.* New York: Brunner/Mazel.

Ozonoff, S. (1995). Executive functions in autism. In E. Schopler & G. Mesibov (Eds.), *Learning and cognition in autism,* New York: Plenum Press.

Ozonoff, S., Pennington, B., & Rogers, S. (1991). Executive function deficits in high-functioning autistic children: Relationship to theory of mind. *Journal of Child Psychology and Psychiatry, 32,* 1081–1106.

Ozonoff, S., Rogers, S., & Pennington, B. (1991). Asperger's syndrome: Evidence of an empirical distinction from high-functioning autism. *Journal of Child Psychiatry and Psychology, 32,* 1107–1122.

Pennington, B., & Ozonoff, S. (1996). Executive functions and developmental psychopathology. *Journal of Child Psychology and Psychiatry, 37,* 51–87.

Perner, J., & Wimmer, H. (1985). "John thinks that Mary thinks that . . ." Attribution of second-order beliefs by 5–10 year old children. *Journal of Experimental Child Psychology, 39,* 437–471.

Russell, J., Jarrold, C., & Henry, L. (1996). Working memory in children with autism and with moderate learning difficulties. *Journal of Child Psychology and Psychiatry, 37,* 673–686.

Stone , V., Baron-Cohen, S., & Knight, R. (1998). Frontal lobe contributions to theory of mind. *Journal of Cognitive Neuroscience, 10,* 640–656.

Stone, V., Baron-Cohen, S., Young, A., Calder, A., & Green, J. (1999). *Patients with amygdalectomy show impairments on theory of mind.* Unpublished manuscript. University of Denver.

Sullivan, K., Zaitchik, D., & Tager-Flusberg, H. (1994). Preschoolers can attribute second-order beliefs. *Developmental Psychology, 30,* 395–402.

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. (1997). *The Autistic Spectrum.* Oxford: Pergamon Press.

World Health Organization. (1994). *International classification of disease* (10th ed.). Geneva, Switzerland: Author.