Psychological Medicine, 2000, **30**, 1169–1187. Printed in the United Kingdom

Linguistic processing in high-functioning adults with autism or Asperger's syndrome. Is global coherence impaired?

THERESE JOLLIFFE[1] AND SIMON BARON-COHEN

From the Departments of Experimental Psychology and Psychiatry, University of Cambridge

ABSTRACT

Background. Linguistic processing was explored in normally intelligent adults with either autism or Asperger's syndrome, to test if global coherence was impaired. Global coherence is the ability to establish causal connections and interrelate local chunks into higher-order chunks so that most linguistic elements are linked together thematically. Since individuals with autism are hypothesized to have weak central coherence then one would predict that the clinical groups would have difficulty integrating information globally so as to derive full meaning.

Methods. Two experiments were designed to test global coherence. Experiment 1 investigated whether individuals on the autism spectrum condition could arrange sentences coherently. Experiment 2 investigated whether they were less able to use context to make a global inference.

Results. The clinical groups were less able to arrange sentences coherently and use context to make a global inference.

Conclusions. The results suggest that individuals on the autism spectrum have impaired global coherence. Arranging sentences and making global inferences correlated highly, suggesting that central coherence may be a unitary force in these different tasks. Of the two clinical groups, the autism group had the greater deficit. The effect that such a deficit would have on one's daily life is discussed, along with possible explanations for the clinical groups' greater difficulty, and suggestions for future research.

INTRODUCTION

In the last three decades attempts have been made to elucidate the nature of the linguistic impairment in autism. Much of this early work centred on individuals with learning difficulties. Although learning difficulties are characteristic of the majority of individuals with classical autism, in the present study we concentrate on individuals without general learning difficulties. These higher functioning individuals enable us to test for deficits that cannot be explained by impaired general intelligence. We recruited two such groups on the autism spectrum, one meeting the criteria for autism, and the other meeting the criteria for Asperger's syndrome. Although language delay is not necessary for a diagnosis of autism, our autism participants were selected on the basis of early language development and were required to have had a clinically significant language delay. Our other group were individuals diagnosed with Asperger's syndrome who did not have a clinically significant language delay. We separated the groups on the basis of their early language development in order to assess the impact of early language development on the ability to process language in context.

Early experimental work demonstrated that the basic computational aspects of language (phonology and syntax) and semantic and

[1] Address for correspondence: Dr Therese Jolliffe, Departments of Experimental Psychology and Psychiatry, University of Cambridge, Downing Street, Cambridge CB2 3EB.

conceptual development of language are mental age or IQ appropriate in autism (Bartolucci *et al.* 1976; Tager-Flusberg, 1981*a*, *b*, 1988; Minshew *et al.* 1995). This led researchers to conclude that it must be higher-level linguistic abilities that are impaired (Frith & Snowling, 1983; Loveland *et al.* 1990; Tager-Flusberg, 1995). Not surprisingly, analyses of autistic language have moved from the assessment of 'basic' language abilities to the assessment of higher-level linguistic abilities in more natural and varied contexts. Higher functioning individuals with autism or Asperger's syndrome provide an appropriate testing ground for higher-level linguistic abilities.

The experiments reported in this paper explore the understanding of language in context and at the global level. The global level is generally information that is not in short-term or working memory at one time. This is often very loosely defined as being five or more sentences (see McKoon & Ratcliff, 1992; O'Brien, 1995). This paper investigates global coherence, which requires establishing connections between widely separated pieces of information. This is usually achieved by establishing causal connections (Long *et al.* 1992) and by interrelating local chunks into higher-order chunks (Graesser *et al.* 1994) so that most linguistic elements are linked together contextually. Investigating global coherence provides a test of Frith's (1989) weak central coherence, which predicts an impairment in integrating information globally so as to derive or make use of global meaning.

There is evidence for, and against, the central coherence hypothesis. Thus, it has been shown that some individuals with autism succumb to visual illusions (Happé, 1996) while others do not (Ropar & Mitchell, 1999). Some individuals with autism or Asperger's syndrome show superior disembedding on the Embedded Figures Test (Shah & Frith, 1983; Jolliffe & Baron-Cohen, 1997), while others do not (Brian & Bryson, 1996). Similarly, Mottron & Belleville (1993) found an autistic adult who showed a lack of global precedence on hierarchical figures, although Ozonoff *et al.* (1994) did not replicate this (though see the discussion in Jolliffe & Baron-Cohen, 1997, for an explanation of why this finding may have been an artefact of experimental design). This paper attempts to test the central coherence hypothesis with two experiments, both of which require participants to achieve global coherence.

EXPERIMENT ONE: THE GLOBAL INTEGRATION TEST

The Global Integration Test required participants to rearrange sentences in accordance with a theme in order to tell the most coherent story. This made use of behavioural material. Half the stories contained temporal cues, which referred to times of the day or year (and as such these stories would assist in sentence rearrangement), while the other half contained no temporal cues.

The stories without temporal cues required participants to rely on and interpret the theme provided and rearrange sentences coherently on the basis of integrating information within the context of the theme. It was hoped that these stories would challenge those individuals with weak central coherence who might be forced to rely on a 'chaining' strategy, i.e. linking pairs of sentences by identifying concepts in each sentence that relate to each other, or that refer to previously mentioned information. For instance, some stories, such as the one about a young child starting school for the first time, might 'trip up' those who rely on a chaining strategy, since they would be more likely to place the sentence 'When Amy was *older* she had to start at a *new* school', first, rather than last. Such errors would demonstrate a lack of appreciation of the global nature (theme) of the story. These stories place considerable demands on one's ability to establish coherence and if individuals with autism or Asperger's syndrome have weak central coherence, then they should be less proficient at rearranging these sentences, since they did not have temporal cues to help them.

Participants

All our clinical participants had been diagnosed at reputable hospitals and clinical centres. For every participant clinicians were contacted (with the patients' and parents' consent) and medical files were inspected. Individuals with autistic conditions were distinguished on the basis of their early language development. In cases where either the clinician or the parents could not be sure about early language development, these individuals were excluded. Individuals whose

early language development was considered to be borderline were also excluded.

Distinguishing individuals on the basis of their early language development gave rise to two groups, one of which clinicians regarded as meeting the criteria for autism with language delay (DSM-IV, American Psychiatric Association, 1994) and the other which clinicans regarded as meeting the criteria for Asperger's syndrome (ICD-10, World Health Organization, 1992). For the Asperger's syndrome group a number of individuals had not received a diagnosis until adulthood, and for two individuals, clinicians revised an early diagnosis of autism to that of Asperger's syndrome on the basis that these individuals did not have a clinically significant delay in early language development. Whereas it is possible that the DSM-IV might classify these participants as having autism without language delay, we have not questioned the clinicians' diagnosis of Asperger's syndrome, especially as this leaves unaffected the key differentiator between the two groups: early language development.

Fifty-one adults participated in the experiment. These comprised 17 with high-functioning autism, 17 with Asperger's syndrome and 17 normal adult control participants. The normal adults acted as a comparison group for the two clinical groups. The majority of clinical participants were tested in their place of residence, except where some preferred to be tested at the university. All control participants were tested in a quiet room at the university. The 17 control participants were taken from the general population of Cambridge and were chosen to match the clinical groups as closely as possible with respect to the characteristics of age, IQ, sex and handedness.

All participants were required to be of at least normal intelligence (i.e. scoring $\geqslant 85$) on the WAIS-R (Wechsler, 1981, full scale, performance and verbal IQ). Participants were matched on IQ and age (see Table 1). Table 1 gives the participant details of verbal IQ (VIQ), performance IQ (PIQ), full-scale IQ (FSIQ) and chronological age (CA). Four one-way ANOVAs revealed no significant differences between groups on any of these variables: VIQ $F(2, 48) = 0{\cdot}51$, $P = 0{\cdot}60$; PIQ $F(2, 48) = 0{\cdot}58$, $P = 0{\cdot}57$; FSIQ $F(2, 48) = 0{\cdot}10$, $P = 0{\cdot}91$ and CA $F(2, 48) = 0{\cdot}59$, $P = 0{\cdot}56$. The sex ratio in all three groups was 15:2 (m:f), reflecting the sex ratio found in these clinical groups in other studies (Wing, 1981; Klin *et al.* 1995). The groups were closely matched on handedness, there were 15 right-handed and two left-handed individuals in the normal and high-functioning autism group, and 14 right-handed and three left-handed in the Asperger's syndrome group. All participants were born in England and English was their first language. Several individuals within each group were either studying for, or held, formal qualifications such as a university degree or diploma. All three groups contained participants from various socio-economic backgrounds. Thus, in each group: two individuals were in socio-economic class 1; three or four were in socio-economic class 2; three or four were in socio-economic class 3; two or three were in socio-economic class 4; two or three were in socio-economic class 5; three or four were in socio-economic class 6.

Since the second experiment required an appreciation of the mental concept of desire, an inclusion criteria was that the clinical groups had to be able to pass both first- and second-order belief tasks. All the clinical participants were found to possess second-order mentalizing ability. The fact that we did not need to exclude any clinical participants on the basis of their mentalizing ability is consistent with other research with high-functioning individuals, which similarly documents high success rates on second-order belief tasks (Bowler, 1992). However, when more sophisticated tests of mentalizing ability are employed, such individuals do have difficulties (Baron-Cohen *et al.* 1997). The first-order task was a version of Perner *et al.*'s (1989) Smarties task. The second-order task was Baron-Cohen's (1989) ice cream van test. Whereas all participants passed the first-order task, five out of 51 participants failed the second-order task. These included one participant with Asperger's syndrome, two with high-functioning autism and two normal control participants. These participants were re-tested on a new variation of the second-order belief task and all passed.

Materials

The stimuli presented consisted of two sets of eight stories and a trial story to accompany each set. These stories consisted of just five sentences

Table 1. *Participant characteristics*

Participant group	CA	VIQ	PIQ	FSIQ
Normal ($N = 17$)				
Mean	30·00	106·47	105·24	106·35
S.D.	9·12	10·94	14·00	12·72
Range	(18–49)	(87–127)	(85–134)	(88–133)
Autism ($N = 17$)				
Mean	30·71	107·59	101·77	105·12
S.D.	7·84	14·37	13·06	13·47
Range	(19–46)	(88–135)	(85–132)	(90–133)
Asperger's syndrome ($N = 17$)				
Mean	27·77	110·82	100·29	107·12
S.D.	7·81	13·51	14·23	14·34
Range	(18–49)	(89–130)	(85–133)	(86–132)

in each. Each of the sentences appeared on separate cards. There were two types of set; one set had temporal cues (the Temporal condition), and the other set was without temporal cues (the Coherence condition). Each of the five sentences within a story contained a number from 1–5, which was placed in the top left hand corner on the reverse side and which was to assist with their laying out in a specific order.[1] The exposure to each story's title card was timed and response time was gathered on how long it took participants to rearrange the sentences.

All the stories used in this experiment were behavioural and none of them were mentalistic or were particularly social in nature. The stories used were presented, in a pilot study, to 12 individuals who found them simple to read and comprehend. Examples of the stimuli are shown in Appendix 1.

Procedure

Each participant was tested alone in a quiet room which was free from distractions.

The trial item

The two types of story were placed in two separate piles face down on the table. In order to prevent order effects, half of the participants in each group received the randomized Temporal condition first, the other half received the randomized Coherence condition first. Participants were randomly assigned to one order or the other and the trial item for this condition was administered first. The participants were told that they were going to see some sentences that form a story. The experimenter mentioned that this was just a trial story, so they could get use to the procedure before they started the test items. The experimenter used a screen to hide the sentences, while she laid them in the order circumscribed by the number 1–5, which appeared on the reverse of the sentences. The screen was removed and the participants were told that they would be seeing sentences similar to those presented.

The experimenter drew the participant's attention to the fact that there were five sentences and they were asked to read them aloud. Then the story's theme was placed down on the table at the head of the story. Participants were told that this was the theme of the story and they were asked to read this aloud. They were given a further 10 s to reflect on what they had read. Participants were then asked to rearrange the sentences into the most sensible (coherent) story for the title presented. They were told they would be given as much time as they wanted. In the rare instance where a participant failed to arrange the practice items correctly, the experimenter demonstrated the correct order and then re-administered the item. This procedure was done to provide some reassurance that individuals understood the task.

[1] This numbering facilitated three things: first, that the correct sentence for the beginning of each story was never placed at the beginning; secondly, that the correct sentence for the end of each story was never placed at the end; thirdly, that the sentences in the story that preceded or followed each other were never placed next to each other when they were laid out in the test. For example, when the starting order (1–2–3–4–5) was rearranged, the correct order of sentences might now appear, according to their labels on the reverse, as the following: 4–1–3–5–2. For each set type the following orders were used: 2–4–1–5–3; 2–5–3–1–4; 3–1–5–2–4; 3–5–2–4–1; 4–2–5–3–1; 4–1–3–5–2; 5–2–4–1–3; 5–3–1–4–2.

The test items

It was clear from the participants' approach to the trial story that they fully understood the requirements of the task, so the experimenter proceeded with the test items. Participants were told that with these they would be timed to see how long it took them to put the sentences into the most coherent order for the title. They were instructed to tell the experimenter as soon as they had finished.

On completion of their first condition, participants were administered their second condition. When they were given their second condition they were always given its respective trial item before starting the test items. All participants completed both conditions in one session, although participants were permitted to take a break between them.

Scoring

Participants were granted a score of 1 or 0 for each story, depending on whether or not the story had been coherently arranged up to a maximum score of 8 for each set. A stopwatch was used to time the rearrangement of sentences and was started as soon as the experimenter had given the instruction to rearrange the sentences.

Results

Table 2 shows the mean accuracy scores and response times for rearranging sentences by Condition for the three groups.

Accuracy on rearranging sentences

A two-factor repeated measures ANOVA was performed, which had a between-participant variable of Group, and a within participant variable of Condition (Temporal and Coherence). It revealed a significant main effect of Condition ($F(2, 48) = 543$, $P < 0{\cdot}001$) and Group ($F(2, 48) = 543$, $P < 0{\cdot}01$) and there was a Group by Condition interaction ($F(1, 48) = 72$, $P < 0{\cdot}01$).

The Group by Condition interaction was investigated further to examine whether the group effect applied to one or both of the conditions. As predicted, simple effects showed the effect of Group to be significant only for the Coherence condition (F Coherence $(2, 48) = 8{\cdot}70$, $P = 0{\cdot}001$; F Temporal $(2, 48) = 0{\cdot}35$, $P = 0{\cdot}71$). The source of the interaction was investigated further using t tests. Planned contrasts indicated that the Coherence mean for the autism (aut.) and Asperger's syndrome (Asp.) groups, was as predicted, significantly lower than that of the normal control group (t aut. $(48) = 4{\cdot}16$, $P < 0{\cdot}001$; t Asp. $(48) = 2{\cdot}30$, $P < 0{\cdot}05$), while the Coherence mean for the autism group was not significantly different to that of the Asperger's syndrome group, it did however approach significance ($t(48) = 1{\cdot}87$, $P = 0{\cdot}07$).

Given that the Condition effect was significant, it was useful to see whether there were different Condition effects for the three groups. Simple effects showed the effect of Condition to be significant for the clinical groups (F aut. $(1, 48) = 43{\cdot}56$, $P < 0{\cdot}001$; F Asp. $(1, 48) = 24{\cdot}04$, $P < 0{\cdot}001$), but not the normal control group ($F(1, 48) = 2{\cdot}28$, $P = 0{\cdot}14$). Observation of the means (see Table 2) show the clinical groups to be significantly worse on the Coherence condition relative to their own performance on the Temporal condition, whereas the normal control group showed no differential pattern of performance on the two conditions.

In order to determine whether the clinical groups' relative failure in rearranging sentences on the Coherence condition was determined by just a few individuals in each group or was more widespread among the group, a majority analysis was conducted. Thus, the number of participants in each of the clinical groups scoring above (and below) the control group mean was calculated. This was compared with the numbers of participants in the normal group scoring above (and below) their mean. The analysis revealed that the autism group differed significantly from the normal control group ($\chi^2(1) = 9{\cdot}07$, $P < 0{\cdot}01$) although the Asperger's syndrome group did not differ from the normal control group ($\chi^2(1) = 1{\cdot}07$, $P = 0{\cdot}30$).

Response times for rearranging sentences

The raw response time data were transformed into logarithms (base 10) to reduce skewness and stabilize the variances which were unequal due to outliers. Then a two-factor repeated measures ANOVA was performed on the response times for the three groups. This ANOVA had a between-participant variable of Group, and a within participant variable of Condition

Table 2. *Mean accuracy scores and response times for rearranging sentences*

Participant group ($N = 17$)	Rearranging		Response times	
	Coherence	Temporal	Coherence	Temporal
Normal				
Mean	6·77	7·24	29·92	25·53
S.D.	0·97	0·97	1·69	1·60
Range	(5–8)	(5–8)	(10·00–63·10)	(7·94–50·12)
Autism				
Mean	5·06	7·12	43·45	15·56
S.D.	1·25	0·70	1·47	1·35
Range	(3–7)	(6–8)	(25·12–79·43)	(7·94–25·12)
Asperger's syndrome				
Mean	5·82	7·35	37·15	14·62
S.D.	1·33	0·79	1·48	1·53
Range	(4–8)	(6–8)	(19·95–63·10)	(7·94–25·12)

Table 3. *Mean accuracy scores on each condition*

Participant group	G. Inf.	C. Goal	Comp.	Mem.
Normal ($N = 17$)				
Mean	9·24	9·29	9·77	9·65
S.D.	0·90	0·77	0·44	0·35
Range	(8–10)	(8–10)	(9–10)	(8·8–10)
Autism ($N = 17$)				
Mean	7·00	8·97	9·71	9·80
S.D.	1·70	1·35	0·59	0·24
Range	(4–9)	(5–10)	(8–10)	(9·3–10)
Asperger's syndrome ($N = 17$)				
Mean	8·00	9·24	9·88	9·71
S.D.	1·70	1·30	0·33	0·38
Range	(5–10)	(5–10)	(9–10)	(9·0–10)

G. Inf., Global inference question; C. Goal, Story character's goal or intention question; Comp., Comprehension question; Mem, Memory recall task.

(Coherence and Temporal). It revealed a non-significant main effect of Group ($F(2, 48) = 0{\cdot}82$, $P = 0{\cdot}45$) but a significant main effect of Condition ($F(1, 48) = 283{\cdot}72$, $P < 0{\cdot}001$). The higher-order interaction of Group by Condition was also significant ($F(2, 48) = 42{\cdot}81$, $P = 0{\cdot}001$).

Given that the Condition effect was significant, simple effects were examined that compared the two conditions for each group. Analysis of simple effects showed the effect of Condition to be significant for all three groups (F aut. $(1, 48) = 199{\cdot}45$, $P < 0{\cdot}001$; F Asp. $(1, 48) = 165{\cdot}02$, $P < 0{\cdot}001$; F norm. $(1, 48) = 4{\cdot}87$, $P < 0{\cdot}05$). Observation of the means (see Table 3) shows all three groups to be significantly slower at making their responses on the Coherence condition, relative to their own performance on the Temporal condition.

Although the Group effect was non-significant, the Group by Condition interaction was investigated further to see if there were different group effects on each condition. Simple effects showed the effect of Group to be significant for the Temporal condition (F Temporal $(2, 48) = 9{\cdot}56$, $P < 0{\cdot}001$) and the Coherence condition (F Coherence $(2, 48) = 3{\cdot}12$, $P < 0{\cdot}05$). Since predictions were not made about participants' response times, the source of the interaction effect was investigated further using *post hoc* Newman-Keul's tests. *Post hoc* comparisons revealed that response times on the Temporal condition for the clinical groups were significantly different from that of the normal control group ($P < 0{\cdot}05$; $P < 0{\cdot}01$ respectively) although the two clinical groups did not themselves differ ($P > 0{\cdot}05$). Similarly, *post hoc* comparisons revealed that response times on the Coherence

condition were for the autism group significantly different from that of the normal control group ($P < 0{\cdot}05$), though the Asperger's syndrome group did not differ significantly from the normal control group, and neither did the clinical groups differ ($P > 0{\cdot}05$).

Discussion

Relative to the control group the two clinical groups performed less well in rearranging sentences in the Coherence condition, but did not have any difficulty in rearranging sentences in the Temporal condition. Furthermore, in the Coherence condition the autism group tended to perform more poorly at rearranging sentences compared with the Asperger's syndrome group. The clinical groups' relative difficulty on the Coherence condition was also reflected in their response times, which were longer for both groups and significantly longer for the autism group. When looking at performance within groups it was found that the normal control participants performed equally well on both conditions. In contrast, the clinical groups performed worse on their Coherence condition relative to their own performance on the Temporal condition. The results of a majority analysis suggest that poor performance characterizes the majority of those with autism, but not the majority of those with Asperger's syndrome.

The stories with temporal information are easier to arrange since an individual can rely on lexicon-based knowledge, such as 'got up early', 'after lunch', 'by mid-afternoon' and 'evening'. This contrasts with the stories containing thematic information, where participants have to interpret the information. Thus, the sentence 'It was not long before Charlotte was able to pat and stroke the horse's mane', suggests the protagonist has a greater familiarity with the horse than the sentence 'The horse would gallop away every time Charlotte would walk towards him.' However, these differences are not explicit. Thus, this task requires a continual establishing of relationships between the different pieces of information and the context in which they are presented. The clinical groups' greater difficulty with these sentences may stem from the fact that they are less able to interpret and be sensitive to contextual information.

Further evidence for their difficulty comes from response times. Participants from all groups took longer to arrange the sentences in the Coherence condition relative to their own performance on the Temporal condition. This is explicable in terms of there being no temporal cues with which to assist them so participants had to rely on interpreting the information in context. Furthermore, the clinical groups took longer to respond on this condition than did their normal control group, although it was significant only for the autism group. This suggests that they had greater difficulty in trying to achieve coherence. In contrast, the clinical groups were significantly faster than the normal control group when it came to making their responses on the Temporal condition. This speed difference presumably reflects an over-reliance on a temporal strategy. Thus, the clinical groups tended to be drawn first to the temporal cues, and only later did they check whether their arrangement was coherent. In contrast, the normal group tended first to deliberate on meaning, and only then did they use the cues as a means of verifying their arrangements. The results on the Temporal condition seem to suggest that their good performance on this condition was made possible by the presence of temporal cues.

In the next part of this discussion we consider some factors that might be suggested as influencing the clinical groups' ability to arrange sentences coherently. At the end of this paper we return to consider other factors common to both the experiments reported in this paper. These factors include memory, comprehension, motivation, attention, fatigue, communication skills and omissions.

It is unlikely that the clinical groups' poor performance in arranging sentences coherently is a secondary consequence of impulsivity (motor or cognitive). This is because they responded slower on the arrangements that they could not identify easily and were slower in responding on the Coherence condition relative to the normal control group and relative to their Temporal condition. The clinical groups' longer response times suggests prolonged thinking times, rather than impulsivity. Neither is it likely that the clinical groups' poor performance is due to a tendency to fixate on details, process locally, or

due to a problem with mental disengagement. This is because on the Temporal condition, a condition on which they were unimpaired, they were required to disengage from one sentence (or part of a sentence) to the next in order to be able to arrange the sentences coherently. Also, it is unlikely that the differences were due to sequencing problems or reading problems since they had no difficulty with the Temporal condition. Moreover, the fact that groups were matched on verbal intelligence would make deficient reading ability less likely. Since the stories require only a minimal amount of social learning or general knowledge, it is unlikely that these normally intelligent adults had so little life experience that they would be disadvantaged in these respects. However, it is possible that they could have had a problem in using their knowledge. But it is unlikely that deficient reasoning ability is a factor, since these individuals are not known to have a reasoning impairment *per se*. This has been shown with a whole range of reasoning tasks such as the Raven's Matrices (Raven *et al*. 1992) and tests of logical, analogical and counterfactual syllogistic reasoning (Scott & Baron-Cohen, 1996; Scott *et al*. 2000).

It is possible that the clinical groups simply maintained poorer standards of coherence, i.e. were satisfied with less coherent arrangements – this is because the arrangement is determined at least in part by the degree of coherence sought after (Van den Broek *et al*. 1995; Britton & Graesser, 1996). However, if the clinical groups employed lower standards of coherence or failed to check for plausibility one might have expected their response times to be faster, not slower. The longer response times by clinical groups on the Coherence condition suggests more, rather than less processing, making the idea that participants employ weaker standards of coherence much less likely. Moreover, inspection of the data suggests the clinical groups were more likely to make their errors on those contexts that made the greatest demands on high-level integration. This tends to support a coherence deficit.

The results suggest that even when normal participants have no temporal cues to assist them, they have a drive to establish meaning. This is consistent with Frith's (1989) comment that in the normal cognitive system there is a built-in propensity to form coherence over as wide a range of stimuli as possible. At the same time the clinical groups' relatively poor performance on the Coherence condition supports Frith's suggestion that it is this capacity for coherence that is diminished in autism. Weak central coherence would predict difficulties when large amounts of information need to be pulled together. In addition, despite trying to design the stories to make sentences difficult to chain, the inevitable and frequent 'chaining' that did occur when participants were rearranging sentences, raises serious questions as to whether the clinical groups were able to integrate fully what they read. Problems in integrating linguistic information have been reported in several studies. These studies present findings that suggest that children and adults with autism of differing abilities find tests that require the integration of linguistic elements very difficult (Carpentieri & Morgan, 1994; Ozonoff & Miller, 1996; Rumsey & Hamburger, 1988, 1990). The next experiment looks at how the clinical groups interpret the meaning behind spoken material that needs to be integrated in context.

EXPERIMENT TWO: GLOBAL COHERENCE INFERENCES

This experiment examines whether participants can extract and integrate information from a story so as to make a global inference about why a particular action was carried out. The stories were short, each being just five to seven sentences long. Each story began by stating the main character's desire or goal, e.g. wanting to decorate a bare wall. This was followed by stating a subgoal, which the character hoped would achieve this goal, e.g. wanting to buy a painting. When this subgoal could not be achieved, or was found not to be plausible, e.g. not getting a painting, a new subgoal was introduced, e.g. that of buying a beautiful clock. On hearing these stories participants were required to do four things: answer a question as to why the character wanted to achieve the first subgoal, e.g. buy a painting; answer a question as to why the character wanted to achieve the second subgoal, e.g. buy a clock; answer a comprehension question, e.g. did the character do/get X?, e.g. a painting; and recall the story from memory. While even high-functioning

people with an autism spectrum condition have difficulties with complex mental state reasoning (Baron-Cohen *et al.* 1997) we chose the basic mental states of 'desire' and 'goal' because these are understood by people with autism (Baron-Cohen, 1991).

The question as to why the character wanted to achieve the first subgoal, e.g. buying a painting, determines whether the participant can understand the character's desire (to decorate a bare wall). The question as to why the character wanted to achieve the second subgoal, e.g. buying a beautiful clock, determines whether the participant can make a coherent global inference.

The ability to make global inferences has received a great deal of attention during the last two decades (Kintsch & van Dijk, 1978; Garrod & Sanford, 1988, 1990; Albrecht & O'Brien, 1995; Kreunz & MacNealy, 1996). There are different views on exactly how global inferences are made, and it is beyond the scope of this paper to discuss these views. However, many researchers have proposed that global inferences (or global coherence inferences as they are often called) serve to establish connections between widely separated pieces of textual information. Global coherence is believed to be established when an individual integrates what they have just heard or read with earlier sentences that they have heard or read. McKoon & Ratcliff (1992) proposed that when new information does not map directly onto old information, comprehenders must search the context for information that would increase coherence. Thus, when the character in the story decides to buy a beautiful clock as a result of not being able to buy a painting, a reader needs the global information that she is trying to decorate a bare wall to make sense of the scenario. This is because the relation between the problem (not being able to get a painting) and the substitute (a beautiful clock) cannot easily be determined at a local level.

Narrative texts are particularly suitable for testing the ability to make global coherence inferences, since they are like everyday experiences; involving people performing actions in pursuit of goals, the occurrence of obstacles to goals, and reactions to these obstacles. Moreover, despite the attempt to provide naturalistic scenarios that deal with simple desires, the emotional language within the stories was kept to a minimum. This was done so that clinical patients who might have deficits in the comprehension of affective states would not be disadvantaged. In keeping emotional language to a minimum, the extraction of meaning in context could be separated from the ability to understand affect.

Given that Frith (1989) believed that in autism there is a deficient capacity to extract higher-level 'meaning', the prediction was made that individuals with either autism or Asperger's syndrome would be less proficient at making globally coherent inferences.

Participants

The participants were the same as those who took part in the first experiment.

Materials

The stimuli presented consisted of short stories ranging from five to seven sentences in length. There was one trial story and 10 test stories. Each story had three accompanying questions and a memory task which appeared on an A4 score sheet along with a space to record what was said.

Two tape recorders were used. One of the tape recorders presented the stories, while the other recorded the participants' attempts at recalling stories, for transcribing at a later date.

The stories that were used in this experiment were behavioural and they were relatively easy to comprehend and recall. Examples of the stories and questions used in this experiment are shown in Appendix 1.

Procedure

The experimenter and participant were seated at a table in a room that was free from distractions.

The types of question

Each of the stories had three accompanying test questions and a memory task: the global inference question, which took the form of 'Why did X do A?'; the desire (character's goal) question, which took the form of 'Why did X want B?'; the comprehension question, which usually took the form of 'Did X do/get C?'; the memory task, which took the form of 'I want you to try and recall the story you have just heard!'.

The following story will illustrate these question types. Thus, in the story where Mrs Jones is trying to decorate her bare wall with a painting, but was unable to get the one she wanted so goes and buys a beautiful clock instead, the *global inference question* 'Why did Mrs Jones buy a clock?' and the statement that 'she bought the clock to brighten her wall', is correct for the story context. However, the global inference statement 'so she would be able to tell the time' would be incorrect for the context of the story. The *desire* (*character's goal*) *question* 'Why did Mrs Jones want a painting?' would require an answer like, 'so she could decorate her wall'. The *comprehension question*, 'Did Mrs Jones get a painting?' would require the participants answering 'no'. Finally, the *request to recall* the story from memory would require participants to recall the four propositional units of the story. Thus, participants had to recall correctly the character's desire (decorating wall), subgoal (buying a painting), problem with subgoal (not getting the desired painting), and the new subgoal (a beautiful clock) for each story type.

The global inference question was the critical question and required participants to be particularly sensitive to the story context (and the character's goal or desire) in order to make the correct inference for the story. The other two questions for the story were control questions. Thus, the desire (character's goal) question was aimed at assessing whether any difficulty with the global inference question could be due to a general problem in apprehending the story character's intention or desire. The comprehension question and memory task was to check whether any difficulty on the global inference question was due to a problem in comprehending what was happening in the story or in remembering the events of the story. The control questions/task also enabled the experimenter to address and possibly control for problems with concentration, motivation and fatigue.

Ordering of questions, stories and avoiding a response bias

Since the global inference question was the critical question it was necessary to have this question preceding the desire (character's goal) question, since the answers to both these questions were essentially the same and the desire (character's goal) question could cue the global inference question. By the same reasoning, this inevitably results in the situation where the global inference question could cue the desire (character's goal) question. If the clinical groups did have a problem with the global inference question, then this was thought to make the 'desire' question more difficult for them to pass, and thus this control question would be more rather than less conservative, since the clinical groups would then get less benefit from cueing. It was also necessary to have the desire (character's goal) question coming before the story recall task, again so as to prevent the memory task from cueing this important question.

The risk of cueing constrained the ordering of questions. In half of the stories the comprehension question came first and in the other half last. The upshot of this ordering was that the global inference and desire (character's goal) questions and the memory task, were always presented either first, second and third, or second, third and fourth. Therefore, five of the stories had their questions in one order, while five had their questions in another order. In addition, the two orders of questions were presented intermixed. Thus, participants were presented with stories that had the two orders or ways of asking the questions, intermixed. This provided some degree of control against order effects with participants answers to story questions. To control further for order effects, it was necessary to control for the order of the stories themselves having an effect as well. Therefore, participants in each group were randomly assigned to one of two story orders. Thus, one of the two orders of presenting the stories and the mixed ordering of questions were given to each participant.

In order to prevent a response bias in participants answers to the comprehension question, half of the comprehension questions required a 'yes' answer and half required a 'no' answer.

The trial item

The tape recorders were placed face up on the table directly in front of the participant. Participants were instructed to listen very carefully, as straight after they had heard each story they would be required to answer questions about it, as well as recall the gist of it from memory. The

stories were presented on one tape recorder, and the other machine was used to record the participants attempts at recalling the stories. The questions were not presented on a tape recorder, but were read from a score sheet. The experimenter played the trial story and then asked its accompanying questions and the participants had to recall the story from memory. Although a minority of the clinical participants had a problem in answering the global inference question they performed perfectly on the control questions. For participants failing the global inference question their performance was not corrected, as any attempt to try to point out why they were wrong was felt perhaps to give rise to them seeing a strategy which they could use to pass other subsequent and similar questions. Their perfect performance on the other questions left the experimenter feeling fully justified in not correcting participants' errors on the global inference question.

The test items

When the participant had completed the trial item, the experimenter then introduced the test items. The participants were told that the procedure and requirements was the same as for the trial item. They were instructed to listen very carefully because they would get to hear the story only once. The experimenter presented one of the two story orders to each participant and then asked the three test questions and requested participants to recall the story after each story had been presented. The participants were given as long they wanted to provide their answers and recall the story since the test was not timed. They were also permitted to make amendments, and were given frequent breaks between stories in order to avoid fatigue. Occasionally, it was necessary to prompt participants for further information. Thus, if participants in response to the global inference question (e.g. 'Why did Mrs Edwards get her knitting needles out?') gave a local inference (e.g. 'because she couldn't buy a jumper') they were prompted further, 'yes, she got her knitting needles out because she couldn't buy a jumper, but why did she get her knitting needles out?'. Participants were prompted only once. Similarly, if they gave an ambiguous answer, they were prompted for more information in the form of 'What do you mean by that?', but again only once. Thus, if in response to a global inference question such as 'Why did Mark go cycling?' a participant was to answer 'because he needed to', the participant would be prompted further ('What do you mean by that?') to see whether they understood that the context of the story was about getting more exercise. Such prompting prevented ambiguous and unscorable answers.

Scoring

Participants' original responses were noted, along with any amendments. Their answers were recorded in full, so that they could be analysed at a later date. Participants were given a score of one for each correct answer. In the case of the memory control task, participants were given a score of four, if all four of the propositional units were recalled. The total score for each participant on the memory control task was then divided by four prior to the analysis, so that each condition was given an equal weighting in the analysis.

Whereas the comprehension question gave rise to a straightforward 'yes' or 'no' answer, judging the appropriateness of explanations and to a much lesser extent evaluating the propositional units recalled, is clearly subjective. A second rater was therefore employed, who was blind to the hypothesis being tested and the diagnostic status of the participants. She rated all the answers and the propositional units recalled for nine of the participants in each experimental group. As expected, inter-rater agreement was extremely high. There was 99% agreement on each of the Memory control, Desire (character's goal) and Global Inference tasks.

Results

The scores for each type of question or task, were the number of correct answers, which could range from 0 to 10. The mean accuracy scores for the three groups, by task, are illustrated in Table 3.

Accuracy on the four conditions

A two-factor repeated measures ANOVA was performed on the mean accuracy scores for each of the three groups. This ANOVA had a between-participant variable of Group, and a within participant variable of Task (Global

Inference; Story Character's goal; Memory; Comprehension). The ANOVA revealed a significant main effect of Group ($F(1, 48) = 3{\cdot}96$, $P < 0{\cdot}05$) and a significant main effect of Task ($F(2, 48) = 43{\cdot}45$, $P < 0{\cdot}001$), the latter of which can only be meaningfully interpreted as the result of its higher-order interaction with Group, which was as predicted significant ($F(2, 48) = 7{\cdot}05$, $P < 0{\cdot}001$).

Given that the Task effect was significant, then the Group by Task interaction must at least in part be due to there being different task effects for the three groups. Simple effects were examined which compared the different tasks for each group. These showed the effect of Task to be significant for the clinical groups (F aut. $(1, 48) = 39{\cdot}12$, $P < 0{\cdot}001$; F Asp. $(1, 48) = 185$, $P < 0{\cdot}001$), but not the normal control group ($F(1, 48) = 1{\cdot}57$, $P = 0{\cdot}20$). Observation of the means (see Table 3) show the clinical groups to be worse on the Global Inference task relative to the other three tasks, whereas the normal control group showed no differential pattern of performance.

The significant Group effect and this Group by Task interaction means that this interaction must again be due, at least in part, to there being different Group effects for each task. Simple effects were examined which compared the different groups on each task. These showed the effect of Group to be non-significant for the memory control task ($F(2, 48) = 0{\cdot}84$, $P = 0{\cdot}44$), the comprehension question ($F(2, 48) = 0{\cdot}64$, $P = 0{\cdot}53$) and the story character's desire question ($F(2, 48) = 0{\cdot}45$, $P = 0{\cdot}64$), but significant for the global inference question ($F(2, 48) = 9{\cdot}74$, $P < 0{\cdot}001$). The global inference question was investigated further using t tests. Planned contrasts indicated that the global inference means for the clinical groups, were as predicted, significantly lower than that of the normal control group (t aut. $(48) = 4{\cdot}41$, $P < 0{\cdot}001$; t Asp. $(48) = 2{\cdot}43$, $P < 0{\cdot}05$) and the global inference mean for the clinical groups did not differ, although it approached significance ($t(48) = 1{\cdot}97$, $P = 0{\cdot}06$).

In order to determine whether the clinical groups' relative difficulty with the global inference question was determined by just a few individuals in each group or was more widespread among the participants, a majority analysis was conducted. Thus, the number of participants in each of the clinical groups scoring above (and below) the control group mean was calculated. This was compared to the numbers of participants in the normal group scoring above (and below) their mean. The analysis revealed that the clinical groups differed significantly from the normal control group (χ^2 Asp. $(1) = 4{\cdot}64$, $P = 0{\cdot}031$; χ^2 aut. $(1) = 9{\cdot}67$, $P = 0{\cdot}002$, Yates' Continuity Correction to correct for expected frequencies < 5).

Omissions

The number of omissions made on the global inference question was examined. Because of the small number of omissions involved, three 2×2 chi-squared tests (with Yates' Continuity Correction to correct for expected frequencies < 5) were conducted to see whether the groups differed in the proportion of their errors which could be considered to be due to omissions. The analysis revealed no significant difference between any of the groups ($P < 0{\cdot}62$).

Does sentence rearranging and making global inferences correlate?

Since both experiments were designed to tap weak central coherence it was useful to see whether accuracy on the critical conditions correlated with one another. Pearson correlations found the two critical conditions to be positively correlated (Arrang. $\times$ Inf. $= 0{\cdot}56$; $P < 0{\cdot}01$).

Discussion

The three groups did not differ on the memory, comprehension and desire tasks. However, they did differ on the global inference question. In comparison to the normal control group, both groups of clinical participants performed significantly less well on this question, where they were required to give the context-appropriate inference (global coherence inference) for a story character's action. Although the clinical groups themselves did not differ on the global inference question, the autism group's tendency to be less accurate than the Asperger's syndrome group approached significance. Furthermore, the clinical groups performed significantly worse on the global inference task relative to their own performance on the other tasks, whereas the normal group showed no differential pattern of

Table 4. *Number of participants correctly answering the different global inference questions*

Participant group ($N = 17$)	Normal	Autism	Asperger's syndrome
Stories			
Mr Robinson	17	16	16
Mr Owen	17	16	16
Sandra	15	14	14
Chris	17	15	16
Mrs Edwards	17	14	14
Mark	16	12	13
Mrs Jones	17	12	14
Tom	15	9	12
Kevin	13	7	11
Mrs Green	13	4	8

performance on any of their tasks. The results of the majority analysis suggest that poor performance on the global inference question characterizes the majority of both clinical groups.

Given that the clinical groups were able to perform normally on the desire question, it suggests that the findings on this experiment might be explicable in terms of weak central coherence. Thus, the results on this test seem to suggest that coherence factors may be more pervasive than low-level sociocognitive deficits, such as reasoning about desires. Bruner & Feldman (1993) report that able adolescents with autism recognized the psychological motivations in stories, such as attempts to mislead, and yet gave relatively fragmented narratives in a story-telling task.

The errors made on the global inference question were quite striking. For example, in the story where a lady got her knitting needles out in order to make her grandson a jumper, one clinical participant described her action of getting knitting needles out as her wanting to 'keep herself occupied', another as her 'trying to make something to keep herself warm'. In the story where a man decided to save money, by stopping going to the theatre, one clinical participant described this action as being a consequence of the theatre having 'nothing good on any more'.

Although the clinical participants were found to make errors on all types of story, observation of the scores on individual stories (see Table 4) shows that they were not unable to make or generate global inferences. Rather, inspection of the scores for individual stories suggests that the clinical groups have a relative inefficiency in making global coherence inferences and tended to perform worse on some stories than others. The stories that they tended to perform worse on were those where the causal link between the story character's action and the context was less obvious, such as was the case where plants could not be found to make a lady's garden look nice, so the lawn-mower was brought out instead. In contrast, their performance was better when the causal link between the story character's action and the context was more obvious, as was the case where a rope was not long enough to rescue a girl from a pothole, so another entrance to the cave was sought. The crucial difference seems to be that when the degree of elaboration required is less, the clinical participants are more likely to be able to give a coherent global inference. However, when the degree of elaboration required is greater, the clinical participants (and the normal participants to a lesser extent) are less likely to be able to give a coherent global inference.

Making the global coherence inference requires integrating the action in context and being able to extract and isolate the key elements from the context, and see how these elements relate to and hence can be integrated with the global inference question. The patients were able to understand the characters' desire in the stories, but had difficulty extracting this information from the context when the reason behind an action must be inferred from the context. Being unable to use context, results in the clinical participants having to make an assessment based only on the action itself. Therefore, when a participant's statement is considered without respect to the context it might seem locally coherent, but when the wider context is taken into account their idiosyncratic interpretation stands out.

GENERAL DISCUSSION

This paper sought to assess linguistic coherence at the global level. Evidence from the Global Integration test suggests that in comparison to their normal controls the clinical groups were less able to integrate sentences with each other

and with the theme to provide the most coherent arrangement of the sentences. Evidence from the Global Coherence Inferences test suggests that in comparison to their normal controls the clinical groups provided significantly more context-inappropriate explanations for a story character's action. The results from the two experiments suggest that individuals with an autism spectrum condition have difficulty in achieving global coherence.

In the Global Coherence Inferences test the majority of the participants in both clinical groups were impaired on the global inference question. In the Global Integration test only the autism group demonstrated a majority impairment in arranging sentences coherently. Although it is not entirely clear why the Asperger's syndrome group was not similarly impaired, it is possible that on this measure, they (and the autism group) were able to enjoy some benefit from 'chaining'.

The ability to integrate linguistic information at the global level, did not vary that widely in the normal group, and this was the case whatever the material employed. However, there was wider variability both between and within clinical groups. This suggests that the ability to integrate global (linguistic) information coherently may lie on a continuum in the way that it appears to do on tasks assessing local (linguistic) coherence (Jolliffe & Baron-Cohen, 1999), and in the way that it seems to do with the EFT and Block Design scores (see Jolliffe, 1997).

Could the difficulty for the clinical groups be one of comprehension rather than coherence, such as a failure to comprehend the story even at the sentence level? For the first experiment, the subtleties of errors suggests a coherence rather than a comprehension problem *per se*. Thus, the predicted effect of a comprehension deficit might be errors, which were based simply upon local words such as 'start a new school' or 'builders arrived'. However, most of the clinical groups' errors tended to more global rather than local in nature. That is, their errors tended to reflect either a lack of plausibility in relation to other sentences in the sequence, or tended to be insensitive to the context (title). For example, for the story about a man building his own home, it would be implausible to have the sentence 'John's wife ordered a new suite of furniture for their living room', before the foundations of John's house was built. Moreover, on the Temporal condition a failure to comprehend at the sentence level, would not have allowed participants to arrange the stories where the temporal cues referred to more than one day, such that a morning cue came last rather than first, because the sentence was referring to the next day. To have performed so well on these stories participants must have been comprehending at the sentence level, otherwise sentences with morning cues would have been more likely to have been placed first. Similarly, on the second experiment the clinical groups good performance on the comprehension question, despite not having the passage to refer to, suggests that participants had some ability to comprehend the story. Interestingly, other research, which employs similarly able participants, has found the comprehension of stories to be intact, but have found failures to integrate the information so as to make a coherent global inference (Rumsey & Hamburger, 1988). Moreover, the fact that the clinical groups were of normal verbal IQ, and the fact that many had passed public examinations, would make failures of comprehension at the sentence level much less likely. Despite the evidence for a coherence, rather than comprehension deficit, one cannot rule out that a deficit in comprehension did not have some impact on performance. However, what the evidence seems to suggest is that failures to comprehend at the level of the sentence is unlikely to account for the clinical groups' significant difficulties on both experiments.

It is possible that ceiling effects on the control measures may have affected our ability to detect group differences on these measures. However, in the second experiment, the clinical groups tended to out perform controls on the memory control task, tending to remember not only the gist of the stories but more in the way of details. It is also unlikely that the clinical groups' poor performance is a secondary consequence of motivational, or attentional difficulties, for a number of reasons: first, because they did so well on the control questions; secondly, on both tests the clinical groups were observed to respond to the varying degrees of difficulty of the stimulus materials employed in much the same way as their control group. These findings also suggest that fatigue is unlikely. Thus, conditions,

questions and tasks were intermixed, with the result that fatigue should have affected all conditions, questions and tasks equally, and yet the clinical groups were only impaired on the critical measures. Also, given participants had the opportunity to amend their verbal responses it is unlikely that the clinical groups' relative inefficiency is due to poor communication skills. Neither was there any evidence to suggest that poor performance could be due to a lack of willingness or ability to make a response or have a guess in comparison to their normal control group, since there were no group differences for omissions. Finally, the fact that the clinical groups were able to make inferences suggests that their difficulty cannot be due to generating information *per se*, but rather to generating contextually appropriate information.

Our findings reflects other research, which also shows these individuals to have impaired processing of context. Thus, their poor ability to use language in social and communicative contexts (pragmatic deficit) manifests itself in numerous reported problems. One of these problems is an impaired ability to appreciate an utterance's intended meaning, and this is particularly the case when communicative intent differs from the literal form of the words used, such as occurs with indirect requests, sarcasm and metaphorical expressions (Rumsey & Hanahan, 1990; Happé, 1993, 1994*a*; Ozonoff & Miller, 1996). This results in them tending to take utterances literally (Baron-Cohen, 1988; Happé, 1994*b*; Ozonoff & Miller, 1996) and they have trouble understanding the main point of conversations and stories (Rumsey & Hanahan, 1990). Not surprisingly, they can make socially inappropriate remarks and be tangential and verbose (Dewey & Everard, 1974) and their prosody and intonation are often affected (Baltaxe, 1977; Baltaxe & Simmons, 1992; Goodhart & Baron-Cohen, 1993). There are frequently abnormalities of gesture and facial expression (Attwood *et al*. 1988; Yirmiya *et al*. 1989) and they have an impaired ability to read emotional expression in language and faces (Shields *et al*. 1996).

Since we attempted to assess the impact that early language development had on the clinical groups' ability to achieve global coherence, it is useful to compare the performance of both groups. Although there were no qualitative differences, individuals with autism were more impaired (although not significantly) across both experiments than those with Asperger's syndrome. This is not that surprising since this group tends to be characterized by greater problems in childhood and occasionally greater problems in adulthood. What was surprising was not the fact that the Asperger's syndrome group had milder difficulties with integrating linguistic material than those with a history of classical autism, but the finding that they performed significantly below that of the normal group on both experiments, and that the majority were impaired on one of these experiments. This difference is despite the fact that these individuals did not have the early language delay that characterized our individuals with autism. This finding would seem to suggest that early language development may not play a major role on tasks that assess linguistic coherence and that future research on linguistic coherence could combine individuals with and without language delay. Moreover, of interest is the pragmatic deficits of these individuals, which do not become noticeable until after the first few years of life, when the use of language in context becomes more critical. Therefore, in this respect it may not be that surprising that the language tasks used here pick up a problem in processing language in context. However, what must be mentioned in connection with the particular Asperger's syndrome group recruited, is that it consisted of a high-proportion of clinic cases, and it is possible that clinic cases are more impaired than non-clinic cases.

Since the difference between the clinical groups is just one of degree, it is useful to explore what these groups could, and could not, achieve in terms of comprehension. It is important to clarify that our clinical groups were not unable to process for meaning. It seemed that they had some ability to process meaningfully across the experiments. Moreover, what this paper demonstrated was not an inability, but a relative inefficiency. The clinical groups had some ability to link distant sentences and make global inferences, but they had difficulty linking distant sentences and generating contextually appropriate inferences when a greater degree of elaboration was required, or when the story made very great demands on high-level integration. Connecting distant sentences and

making global inferences requires one to be able to extract, isolate and integrate key elements from a context. The patients were able to comprehend a character's desire in a story, but had difficulty extracting this information from a context when the reason behind an action must be inferred from a context.

This paper prompts future research. Future research will need to address whether the clinical groups' difficulty stems from a lack of general knowledge or a difficulty in using it, a failure to check the plausibility of the context generated, or maintaining poorer standards of coherence. Also, if their standard of coherence is the factor behind their relatively poor ability, one is led to question what causes their poorer standard of coherence and how this relates to their behavioural characteristics. Finally, future research needs to investigate the neurological basis for the difficulty in integrating linguistic information. It is noteworthy that individuals with acquired right hemisphere damage (RHD) or dysfunction have similar difficulties to individuals with an autism spectrum disorder. This might suggest that there is some right hemisphere involvement in autism and Asperger's syndrome. However, there is little neurological evidence to suggest that individuals with an autism spectrum condition have specific neurological abnormalities of the right hemisphere. However, despite this lack of neurological evidence, the wealth of similarities between these two conditions would seem to provide at least preliminary support for a right hemisphere involvement in autism and Asperger's syndrome. Such an involvement might be the disruption to circuits stemming from the right hemisphere. The striking similarities between those with right hemisphere deficits and those with an autism spectrum condition would seem to suggest a biological framework in which the clinical problems of autism and Asperger's syndrome can be interpreted and investigated.

The findings from the two experiments reported lend support to Frith's (1989) weak central coherence hypothesis of autism. Moreover, the correlation on the critical measures suggests that central coherence may be a unitary force in these different tasks. If individuals with either autism or Asperger's syndrome fail to interpret information within a wider context then their ability to appreciate situations would be impaired. Without making coherent global inferences, comprehension will be difficult and information will remain as disconnected bits and pieces. This might partly explain some of their difficulties with the pragmatic aspects of communication. The results coverge to suggest that with linguistic material, individuals with autism or Asperger's syndrome have a cognitive deficit in establishing global coherence.

We are grateful to a large number of people for helping us to recruit subjects: Lorna Wing, Wendy Phillips, the National Autistic Society, Pat Howlin, Clive Robinson, Ben Sacks and Pam Yates. We are grateful to Ian Nimmo-Smith for advice on methodology. The first author was supported by an MRC Studentship during the period of this work. The Harold Hyam Wingate Foundation also provided her with valuable financial support. Both studies reported here were submitted in part fulfilment of her Ph.D. degree at the University of Cambridge.

APPENDIX 1

The Global Integration Test

Coherence test stories

(1) *Making friends with a horse or taming a wild horse*
   (i) Charlotte found a wild horse whilst living near some moorlands.
   (ii) The horse would gallop away every time Charlotte would walk toward him.
   (iii) The horse slowly began to stay still when Charlotte approached him.
   (iv) The horse then let Charlotte touch him for a second.
   (v) It was not long before Charlotte was able to pat and stroke the horse's mane.

(2) *A man building his own home*
   (i) At a local auction John bought a very large piece of land on which to build a house.
   (ii) Builders arrived to dig the foundations of John's house and soon there was a huge mound of earth.
   (iii) The builders then set about building the walls and putting the windows in.
   (iv) It was not long before John could put the central heating in himself.
   (v) John's wife ordered a new suite of furniture for their living room.

(3) *A child starting school for the first time*
   (i) Amy was getting ready for her first day at school.
   (ii) Amy started to cry when her mother left her with the teacher.
   (iii) The teacher was kind to Amy and gave her lots of new books.
   (iv) Amy soon began to make friends with all the other children and started to join in their games.
   (v) When Amy was older she had to start at a new school.

Temporal test stories

(1) *Making a garden look nice*
   (i) Late one summer morning Tina set about tidying up her back garden, which was in a terrible mess
   (ii) By lunch time all the weeds and rubbish had been cleared and Tina was looking in the shed for the lawnmower.
   (iii) By mid-afternoon Tina had made the lawn look beautiful.
   (iv) Later that afternoon Tina went to the shops to buy some garden ornaments.
   (v) By evening Tina had made her garden look beautiful once again.

(2) *A fireman's shift*
   (i) The fireman left his home in the late afternoon to go on duty.
   (ii) The fireman ate his dinner with the other men, whilst they were all exchanging news.
   (iii) Soon after dinner the siren sounded, so the men hurried to their fire engines.
   (iv) All night they fought to put out the fire which had engulfed the council buildings.
   (v) In the early hours of the morning the firemen returned to the fire station.

(3) *Preparing for a dinner or preparing and cooking a dinner*
   (i) Mrs Smith set out at lunch time to buy all the food for her family's dinner.
   (ii) That afternoon Mrs Smith got home from the shop with her shopping bags full.
   (iii) Towards dinner time, Mrs Smith began chopping the carrots and onion.
   (iv) By early evening the dinner was almost cooked and ready to serve.
   (v) That evening Mr Smith returned home and found the dinner ready to eat.

The Global Inferences Test

Test items

(1) Mrs Green wanted to make her front garden look nice, so she went out to buy some plants to plant in her garden. The garden shop assistant sold her some plants and a magazine to go with them. When Mrs Green got home she had a cup of tea and read all about what she had to do with her plants. Afterwards she went into the garden to plant them, but couldn't find where she had put them. She just couldn't find her plants, so got out the lawnmower instead.

Global Inference:
Why did Mrs Green get the lawnmower out?
Character's Goal:
Why did Mrs Green want plants?
Memory:
Can you now recall the gist of the story.
Comprehension:
Did Mrs Green mislay her plants?

(2) Mrs Jones wanted to decorate her very bare wall. In the local paper, she saw a shop was advertising paintings in a sale. So she went to buy a painting to decorate her wall. When she arrived at the shop she hurried in to look around. She came across a painting that she liked very much, but was disappointed to discover it had just been sold. All she could find in the sale was work by modern artists. Mrs Jones didn't like modern paintings, so she went and bought a beautiful clock instead.

Global Inference:
Why did Mrs Jones buy a clock?
Character's Goal:
Why did Mrs Jones want a painting?
Memory:
Can you now recall the gist of the story.
Comprehension:
Did Mrs Jones buy a painting?

(3) Mr Owen was trying to rescue his daughter Amy. She had fallen a short distance down a pothole and needed a rope to haul her out. Mr Owen told his daughter that he was going to get a rope. Soon he came across a runner who

offered to run to a small cottage a mile or so away. After 20 minutes the runner came back with a rope and gave it to Mr Owen. The rope wasn't long enough, but they went to find an entrance to one side of the cave.

Comprehension:

Did the runner give Mr Owen a rope?

Global Inference:

Why did they go to find an entrance to one side of the cave?

Character's Goal:

Why did Mr Owen want rope?

Memory:

Can you now recall the gist of the story?

REFERENCES

Albrecht, J. E. & O'Brien, E. J. (1995). Goal processing and the maintenance of global coherence. In *Sources of Coherence in Reading* (ed. R. F. Lorch and E. J. O'Brien), pp. 263–278. Lawrence Erlbaum Associates: Hillsdale, NJ.

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

Attwood, A. H., Frith, U. & Hermelin, B. (1988). The understanding and use of interpersonal gestures by autistic and Down's syndrome children. *Journal of Autism and Developmental Disorders* **18**, 241–257.

Baltaxe, C. A. M. (1977). Pragmatic deficits in the language of autistic adolescents. *Journal of Pediatric Psychology* **2**, 176–180.

Baltaxe, C. A. M. & Simmons, J. Q. (eds.) (1992). *A Comparison of Language Issues in High-functioning Autism and Related Disorders with Onset in Childhood and Adolescence*. Plenum Press: New York.

Baron-Cohen, S. (1988). Social and pragmatic deficits in autism: cognitive or affective? *Journal of Autism and Developmental Disorders* **18**, 379–402.

Baron-Cohen, S. (1989). The autistic child's theory of mind: a case of specific developmental delay. *Journal of Child Psychology and Psychiatry* **30**, 285–297.

Baron-Cohen, S. (1991). Do people with autism understand what causes emotion? *Child Development* **62**, 385–395.

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.

Bartolucci, G., Pierce, S., Streiner, D. & Eppel, D. (1976). Phonological investigation of verbal autistic and mentally retarded children. *Journal of Autism and Childhood Schizophrenia* **6**, 303–316.

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

Brian, J. A. & Bryson, S. E. (1996). Disembedding performance and recognition memory in autism/PDD. *Journal of Child Psychology and Psychiatry* **37**, 865–872.

Britton, B. K. & Graesser, A. C. (1996). *Models of Understanding Text*. Lawrence Erlbaum Associates: Hillsdale, NJ.

Bruner, J. & Feldman, C. (1993). *Theories of Mind and the Problem of Autism*. Oxford University Press: Oxford.

Carpentieri, S. C. & Morgan, S. B. (1994). A comparison of patterns of cognitive functioning of autistic and non-autistic retarded children on the Stanford-Binet-Fourth Edition. *Journal of Autism and Developmental Disorders* **24**, 215–223.

Dewey, P. & Everard, P. (1974). The near-normal autistic adolescent. *Journal of Autism and Child Schizophrenia* **4**, 348–356.

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

Frith, U. & Snowling, M. (1983). Reading for meaning and reading for sound in autistic and dyslexic children. *British Journal of Developmental Psychology* **1**, 329–342.

Garrod, S. & Sanford, A. J. (1988). Thematic subjecthood and cognitive constraints on discourse structure. *Journal of Pragmatics* **12**, 519–534.

Garrod, S. & Sanford, A. J. (ed.) (1990). *Referential Processing in Reading: Focusing on Roles and Individuals*. Lawrence Erlbaum Associates: Hillsdale, NJ.

Goodhart, F. & Baron-Cohen, S. (1993). How many ways can the point be made? Evidence from children with and without autism. *First Language* **13**, 225–233.

Graesser, A. C., Singer, M. & Trabasso, T. (1994). Constructing inferences during narrative text comprehension. *Psychological Review* **101**, 371–395.

Happé, F. G. E. (1993). Communicative competence and theory of mind in autism: a test of relevance theory. *Cognition* **48**, 101–119.

Happé, F. G. E. (1994*a*). 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 Developmental Disorders* **24**, 129–154.

Happé, F. G. E. (1994*b*). Annotation: current psychological theories of autism: the 'Theory of Mind' account and rival theories. *Journal of Child Psychology and Psychiatry* **35**, 215–229.

Happé, F. G. E. (1996). Studying weak central coherence at low levels: children with autism do not succumb to visual illusions. A research note. *Journal of Child Psychology and Psychiatry* **37**, 873–877.

Jolliffe, T. (1997). *Central Coherence Dysfunction in Autistic Spectrum Disorder*. Ph.D. thesis, University of Cambridge.

Jolliffe, T. & Baron-Cohen, S. (1997). Are people with autism and Asperger syndrome faster than normal on the embedded figures text? *Journal of Child Psychology and Psychiatry* **38**, 527–534.

Jolliffe, T. & Baron-Cohen, S. (1999). Linguistic processing in high-functioning adults with autism or Asperger syndrome: is local coherence impaired? *Cognition* **71**, 149–185.

Kintsch, W. & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review* **85**, 363–394.

Klin, A., Volkmar, F. R., Sparrow, S. S., Cicchetti, D. V. & Rourke, B. P. (1995). Validity and neuropsychological characterisation of Asperger syndrome: convergence with nonverbal learning disabilities syndrome. *Journal of Child Psychology and Psychiatry* **36**, 1127–1140.

Kreunz, R. J. & MacNealy, M. S. (1996). *Empirical Approaches to Literature and Aesthetics. Advances in Discourse Processes*. Mahwah, NJ: Lawrence Erlbaum Associates.

Long, D. L., Golding, J. M. & Graesser, A. C. (1992). A test of the on-line status of goal-related inferences. *Journal of Memory and Language* **31**, 634–647.

Loveland, K. A., McEvoy, R. E., Kelley, M. L. & Tunali, B. (1990). Narrative story telling in autism and Down's syndrome. *British Journal of Developmental Psychology* **8**, 9–23.

McKoon, G. & Ratcliff, R. (1992). Inference during reading. *Psychological Review* **99**, 440–466.

Minshew, N. J., Goldstein, G. & Siegel, D. J. (1995). Speech and language in high-functioning autistic individuals. *Neuropsychology* **9**, 255–261.

Mottron, L. & Belleville, S. (1993). Study of perceptual analysis in a high level autistic subject with exceptional graphic abilities. *Brain and Cognition* **23**, 279–309.

O'Brien, E. J. (1995). *Automatic Components of Discourse Comprehension*. Lawrence Erlbaum Associates: Hillsdale, NJ.

Ozonoff, S. & Miller, J. N. (1996). An exploration of right-hemisphere contributions to the pragmatic impairments of autism. *Brain and Language* **52**, 411–434.

Ozonoff, S., Strayer, D. L., McMahon, W. M. & Filloux, F. (1994). Executive function abilities in autism and Tourette syndrome: an information processing approach. *Journal of Child Psychology and Psychiatry* **35**, 1015–1032.

Perner, J., Frith, U., Leslie, A. M. & Leekam, S. R. (1989). Exploration of the autistic child's theory of mind: knowledge, belief and communication. *Child Development* **60**, 689–700.

Raven, J., Court, J. & Raven, J. (1992). *The Standard Progressive Matrices*. Oxford Psychologists Press: Oxford.

Rumsey, J. & Hamburger, S. (1988). Neuropsychological findings in high-functioning men with infantile autism, residual state. *Journal of Clinical and Experimental Neuropsychology* **10**, 201–221.

Rumsey, J. M. & Hamburger, S. D. (1990). Neuropsychological divergence of high-level autism and severe dyslexia. *Journal of Autism and Developmental Disorders* **20**, 155–168.

Rumsey, J. M. & Hanahan, A. P. (1990). Getting it 'right': performance of high-functioning autistic adults on a right-hemisphere battery. *Journal of Clinical and Experimental Neuropsychology* **12**, 81.

Scott, F. & Baron-Cohen, S. (1996). Logical, analogical, and psychological reasoning in autism: a test of Cosmides theory. *Development and Psychopathology* **8**, 400–411.

Scott, F., Baron-Cohen, S. & Leslie, A. (2000). 'If pigs could fly': an examination of imagination and counterfactual reasoning in autism. *British Journal of Developmental Psychology* **17**, 349–362.

Shah, A. & Frith, U. (1983). An islet of ability in autistic children: a research note. *Journal of Child Psychology and Psychiatry* **24**, 613–620.

Shields, J., Varley, R., Broks, P. & Simpson, A. (1996). Social cognition in developmental language disorders and high-level autism. *Developmental Medicine and Child Neurology* **38**, 487–495.

Tager-Flusberg, H. (1981*a*). On the nature of linguistic functioning in early infantile autism. *Journal of Autism and Developmental Disorders* **11**, 45–56.

Tager-Flusberg, H. (1981*b*). Sentence comprehension in autistic children. *Applied Psycholinguistics* **2**, 5–24.

Tager-Flusberg, H. (1988). On the nature of a language acquisition disorder: the example of autism. In *The Development of Language and Language Researchers* (ed. F. Kessel), pp. 249–267. Lawrence Erlbaum Associates: Hillsdale, NJ.

Tager-Flusberg, H. (1995). 'Once upon a ribbit': stories narrated by autistic children. *British Journal of Developmental Psychology* **13**, 45–59.

Van den Broek, P., Risden, K. & Husebye-Hartmann, E. (1995). The role of readers' standards for coherence in the generation of inferences during reading. In *Sources of Coherence in Reading* (ed. R. F. Lorch and E. J. O'Brien), pp. 353–373. Lawrence Erlbaum Associates: Hillsdale, NJ.

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale – Revised.* Psychological Corporation: New York.

Wing, L. (1981). Asperger's syndrome: a clinical account. *Psychological Medicine* **11**, 115–129.

World Health Organization (1992). *International Classification of Diseases and Related Health Problems* (*ICD-10*) (*10th edn*). World Health Organization: Geneva.

Yirmiya, N., Kasari, C., Sigman, M. D. & Mundy, P. (1989). Facial expressions of affect in autistic, mentally retarded and normal children. *Journal of Child Psychology and Psychiatry* **30**, 725–735.