J. Child Psychol. Psychiat. Vol. 39, No. 6, pp. 893–902, 1998
Cambridge University Press

0021–9630/98 $15.00+0.00

Are There Subgroups within the Autistic Spectrum? A Cluster Analysis of a Group of Children with Autistic Spectrum Disorders

Margot Prior
University of Melbourne, Australia

Richard Eisenmajer
La Trobe University, Melbourne, Australia

Susan Leekam
University of Kent, Canterbury, U.K.

Lorna Wing and Judith Gould
National Autistic Society, London, U.K.

Ben Ong
La Trobe University, Melbourne, Australia

David Dowe
Monash University, Melbourne, Australia

Comprehensive data on the developmental history and current behaviours of a large sample of high-functioning individuals with diagnoses of autism, Asperger's syndrome, or other related disorder were collected via parent interviews. This provided the basis for a taxonomic analysis to search for subgroups. Most participants also completed theory of mind tasks. Three clusters or subgroups were obtained; these differed on theory of mind performance and on verbal abilities. Although subgroups were identified which bore some relationship to clinical differentiation of autistic, Asperger syndrome, and Pervasive Developmental Disorder Not Otherwise Specified (PDD-NOS) cases, the nature of the differences between them appeared strongly related to ability variables. Examination of the kinds of behaviours that differentiated the groups suggested that a spectrum of autistic disorders on which children differ primarily in term of degrees of social and cognitive impairments could explain the findings.

Keywords: Asperger syndrome, autism, Pervasive Developmental Disorder, symptomatology, cluster analysis.

Abbreviations: AS: Asperger syndrome; BPVT: British Picture Vocabulary Test; CA: chronological age; HFA: high-functioning autism; PDD: Pervasive Developmental Disorder; PDD-NOS: Pervasive Developmental Disorder Not Otherwise Specified; PPVT-R: Peabody Picture Vocabulary Test-Revised; VMA: verbal mental age; WISC-III: Wechsler Intelligence Scale for Children-Version 3.

Introduction

Current psychiatric classification systems (e.g. DSM-IV, American Psychiatric Association, 1994; ICD-10, World Health Organisation, 1993) have attempted to provide a way to diagnose differentially children with Pervasive Developmental Disorders (PDD) into subcategories of autism, Asperger syndrome, and Pervasive Developmental Disorder Not Otherwise Specified (PDD-NOS). Although children with this broad class of disorders share social and communicative impairments and restricted, repetitive patterns of interests and behaviours, classification systems suggest that a diagnosis of Asperger syndrome applies when there is no clinically significant delay in language development, and intelligence is in the normal range. However, research indicates that suggested strategies for differential diagnosis may be meeting with little success, and clinicians are confused about how to diagnose these various conditions in a reliable and valid way, based on both current presentation and developmental history (see, e.g., Eisenmajer et al., 1996). Fombonne (in press) notes that at least four different diagnostic proposals have been suggested but no consensus has been reached about their validity; nevertheless, differential epidemiological estimates of Asperger syndrome and autism have been published (Ehlers & Gillberg, 1993).

There is, too, variety in the ways in which samples of autistic and Asperger individuals are selected for research studies on the basis of descriptions offered by Wing (1981), Gillberg (1992), and Szatmari (1992). This leads to difficulties in interpretation of the results of empirical studies that often seek to identify differences between subgroups.

Requests for reprints to: Professor Margot Prior, Department of Psychology, Royal Children's Hospital, Parkville, Victoria, 3052 Australia.

In the last few years, much research has focused on trying to identify similarities and differences between children with Asperger syndrome (AS) and those with high-functioning autism (HFA), whom they appear to resemble in many respects. It is of course amongst the better-functioning groups that the diagnostic dilemma is most salient, since most low-functioning children with autism would not be considered for an AS diagnosis.

A number of studies (e.g Gillberg, 1989; Klin, Volkmar, Sparrow, Cicchetti, & Rourke, 1995; Szatmari, Archer, Fisman, Streiner, & Wilson, 1995; Szatmari, Tuff, Finlayson, & Bartolucci, 1990; Waterhouse et al., 1996) have compared children diagnosed with autism or AS on symptom patterns, behavioural manifestations, cognitive profiles, and "mentalising" ability, i.e. the ability to take the mental perspective of another person (e.g. Ozonoff, Rogers, & Pennington, 1991), to try to find more objective ways of differentiating between the disorders. Ghaziuddin, Tsai, and Ghaziuddin (1992) have summarised some of these studies. Generally, it seems that children with AS are those who are at the higher levels of intellectual ability amongst a group of children with autistic type disorder. Yet Asperger himself (translated by Frith, 1991, p. 74) pointed out that some children with clinical features of the group he described had mild to severe learning disability. Schopler (1985) has argued that we have not yet demonstrated meaningful clinical distinctions between groups of children, and that using the AS label has increased confusion. In sum, attempts to differentiate these subgroups objectively have not yet led to any consensus.

One method of investigating whether there are any *empirically* "true" diagnostic differences between children with HFA and AS is to use statistical approaches to look at factors or clusters of symptoms that characterise putative distinguishable subgroups. There have been some previous attempts at subcategorising autistic children using cluster analytic or taxonomic techniques, most commonly using young children (e.g. Prior, Perry, & Gajzago, 1975). Recent studies of this genre have included that of Waterhouse and colleagues (1996), who identified two taxa or subgroups of young (3–7-year-old) children, using symptoms and behaviours as the database. Although the two groups had different behavioural patterns, the authors argued that their taxa were primarily differentiated by a variety of indicators of developmental status including chronological age and verbal IQ, which underpinned the symptom patterns.

Szatmari (1992), Castelloe and Dawson (1993), Eaves, Ho, and Eaves (1994), Siegel, Anders, Ciaranello, Bienenstock, and Kraemer (1986) and Sevin et al. (1995), have also reported subclassification attempts. It is probably fair to conclude that in most of this work, differences found are related to severity of impairment, and especially level of cognitive and adaptive functioning, rather than to distinctive diagnostic patterns of behaviours. That is, the most salient discriminating features within autistic samples are those related to the ability level of the individual rather than to particular behavioural patterns. Level of ability, though, is imperfectly related to behaviours in that, although higher-functioning children tend to be somewhat less behaviourally deviant, there are individuals of high ability whose behaviour is severely disturbed and very difficult to modify. The relationship between cognitive level and behaviours is complicated by the fact that cognitive factors influence different behaviours in different ways across various domains. For example, Prior and MacMillan (1973) found that higher-functioning children showed more cognitively sophisticated preoccupations and routines than did lower-functioning children (in-depth knowledge of Hopi Indian culture, vs. repetitive chair stacking, for example), although obsessive type behaviours were a strong feature of both groups.

While there is currently substantial interest in trying to discriminate between children with HFA, and those diagnosed with AS, a group named "Pervasive Developmental Disorder-Not Otherwise Specified" (PDD-NOS) in DSM-IV has been somewhat neglected. Children with this "default" diagnosis seem to be those for whom autistic-type symptomatology is atypical or below a threshold where clinicians feel confident in giving an autism diagnosis. They are like children with autism, but impairments are less severe or have a subtly different flavour to them. They may in some cases overlap with children with other diagnoses such as "semantic-pragmatic disorder" (Bishop, in press). Whether children with PDD-NOS are a subcategory of autism or just an "other" category for children with a variety of somewhat abnormal behaviours that resist a clear diagnosis is a further challenge for subclassification research.

This study reports the application of empirical clustering (taxonomic) methods to the reported symptoms and behaviours of a sample of children with autistic-type disorders, and in addition considers the influence of age and verbal abilities on identified patterns. The study is distinguished by the substantial sample size and the higher age and level of functioning of the subjects, by comparison with previous cluster analysis attempts.

Cantwell and Rutter (1994) have noted that it is important that diagnosis and classification should be validated through means other than the defining symptoms. A strategy that could help in deciding whether any identifiable subgroups do have external validity would be to find clear subgroup differences on important behavioural (or cognitive, or biological, etc.) markers. An obvious candidate for such a task is performance on mentalising or "theory of mind" tasks. Strong claims have been made over the past decade for the significance of this ability, or rather disability, in the social/cognitive profiles of autistic children. Indeed it is claimed that it is this "mind blindness" that is the core of the disorder (see, e.g., Baron-Cohen, 1989; Frith, 1989; Happé, 1994).

Theory of mind as a distinguishing feature of autism has been argued to provide an avenue to investigate theoretical and possibly neurological connections between specific brain systems or modules that may be damaged, and key behavioural deficits. In the case of autism it is suggested that there is a basic dysfunction in those systems serving mentalising functions (Baron-Cohen, 1989; Frith, 1989). Although originally it was believed that this ability was a "pure" functional brain module (Baron-Cohen, Leslie, & Frith, 1985, 1986) independent of verbal and cognitive capacities, recent research has suggested that this argument is hard to sustain. Studies of higher-functioning children, including both autistic and AS cases, have shown some capacity for mentalising in a proportion of cases, even if this is limited by comparison with that of normally developing children. In addition it has now become clear that there are discernible relationships between verbal cognitive abilities and the capacity to demonstrate theory of mind capacities (Bowler, 1992; Eisenmajer & Prior, 1991; Frith, Happé, & Siddons, 1994; Ozonoff et al., 1991). It is much less clear whether there are systematic relationships between particular behaviours or symptom clusters in autism, and mentalising ability.

Individuals who show mentalising ability may belong to a different category or subtype of autistic disorder; hence the role of this feature as a diagnostic marker may be relevant *within* an autistic spectrum of disorders. The research reported here was designed to: (a) investigate the possibility that empirically derived subgroups within a pool of children with autistic-type disorders could be identified on the basis of symptoms and behavioural patterns; (b) examine whether mentalising ability might provide one measure of external validity for the existence of any subgroups found, in differentiating between them; or (c) to investigate whether mentalising ability is primarily influenced by cognitive/language ability level rather than by symptom patterns.

In brief, our study was an attempt to discovery whether we could identify by empirical means a subgroup of AS children distinct from autistic children or those with other related disorders such as PDD-NOS; and whether performance on theory of mind tasks might meaningfully differentiate between any obtained groups of children, and therefore could provide some external validity for subtypes of PDD.

The first step in this project was to gather a large sample of children with autistic-type disorders via independent *clinical* diagnoses, and to attempt to separate them into empirically derived clusters, based on behaviour patterns and developmental history. This paper reports on the cluster analyses undertaken to meet this aim, and the relationship between performance on theory of mind experimental tasks and cluster group membership.

Method

Participants

A group of children diagnosed with AS, high-functioning autism (HFA), or related PDD, from Britain and from the South Eastern states of Australia, was initially enrolled in the study ($N = 135$). The sample came from three sources: (a) children assessed and diagnosed at the Centre for Social and Communication Disorders in Britain by the British-based authors; (b) children assessed and diagnosed at the Austin Hospital in Melbourne through the Developmental Assessment team; and (c) through referral for assessment to the first two authors at La Trobe University, Victoria or privately in Adelaide, South Australia. Only high-functioning children were included (i.e. those with intellectual capacities in the approximately normal range), since the study would involve theory of mind testing, which required a minimum verbal mental age of approximately 3–4 years. In any case, the diagnostic confusion that drives this research pertains primarily to higher-functioning children with autistic behaviours.

Children accepted into the study had originally been diagnosed by numerous clinicians from various agencies as HFA ($N = 48$), AS ($N = 69$), PDD-NOS ($N = 7$), or as having autistic features ($N = 11$). This was not a population sample but a clinical one, recruited through contact with experienced clinicians working constantly with these children in their daily practice, and using DSM-III-R criteria for diagnosis at the time of this sampling. Hence its representativeness of a total population of high-functioning autistic cases is unknown.

The ages of the children and adolescents ranged from 3 to 21 years, with a mean age of 10.22 years. Verbal mental age (VMA), obtained for 110 of the subjects as assessed with the Peabody Picture Vocabulary Test (PPVT-R) (Dunn & Dunn, 1981) or the British Picture Vocabulary Test (BPVT) (Dunn, Dunn, Whetton, & Pintilie, 1982), ranged from 2.5 to 33 years, with a mean of 9.8 years. Twenty-five cases were not assessed by the PPVT-R for reasons such as refusal to cooperate or time constraints. Data from testing with the Wechsler Intelligence Scale for Children (WISC-III) were available for 66 of the children. The mean Full Scale IQ score for the subgroup was 96 (Verbal scale IQ 95; Performance scale IQ 97). The correlation between the PPVT-R and WISC verbal IQ estimates for children on whom we had both sets of scores was .74 ($N = 63$). There were 114 males and 21 females.

Measures

The major measure used to provide a detailed account of the developmental history and behavioural characteristics of the subjects, and used as the data bank for the cluster analyses, was a comprehensive diagnostic interview that was developed by the authors based on earlier work by Wing and Rapin with the Autism Spectrum Disorders Checklist (see Rapin, 1996). This checklist covers the symptoms required for diagnosing according to DSM-IIIR, DSM-IV, and ICD-10 systems. It also contains items from descriptions by Kanner, Asperger, and Wing and Gould's triad of social impairments. Empirically derived diagnostic clusters based on data from this instrument could thus be compared with the clinical diagnosis by any one of the "official" systems (Eisenmajer et al., 1996).

Inter-rater and test–retest reliability data are not yet available for this instrument. However, interviewers in this study were trained to present the checklist in a standardised and highly structured way, and checks were made at intervals throughout the research to identify and resolve any procedural or scoring queries. Most participant interviews were carried out by RE. Interviewers were not blind to original child diagnosis. although the results of the clustering analysis were of course not known until all data were analysed. Information was collected on both behaviours and background history.

Behaviours. The checklist covers the domains of: *impairments in social interaction* (including use of body language, greeting behaviour, comfort seeking and giving, awareness of feelings of others, friendships, awareness of social rules; imitation and play, including joint referencing and interactive play; pretend play, and imitation; *impairments in communication and imagination*, including comprehension and use of language, speech characteristics, nonverbal communication, imagination, and pretence; *restrictions and repetition in self-chosen behaviour*, including stereotyped movements, pre-occupations with objects and with patterns of interests, maintenance of sameness.

Responses to the questions are in the form of a "Yes" (positive score) if the behaviours are present, or a "No" (not scored) if they are not present. If the variable was not applicable for the child (e.g. owing to age) or they could not be ascertained from the parents, they were recorded as missing. The set of analyses reported here focuses on currently present behaviours.

Background history. For most cases, data were gathered on pregnancy and birth history, developmental milestones, health problems, family history of any disorders, onset of the disorder (or when first noticed), treatment, and school or work placement, etc. (these data were not available for older individuals not living at home with parents).

The PPVT-R (in Britain, the BPVT), was administered to each child to provide a measure of verbal mental age and standard score. These tests provide measures of single word receptive vocabulary. The PPVT-R has a median correlation of .68 with the WISC-R (Sattler, 1988), and correlations are better with the verbal than with the performance scale of the WISC-R.

Theory of mind measures included the now famous "Sally-Anne" task (Baron-Cohen et al., 1985; Wimmer & Perner, 1983); and the "Box of Smarties" task (Perner, Leekam, & Wimmer, 1987), as tests of "first order" theory of mind; in this case as measures of false belief. Performance on these tasks indicates whether children have the ability to recognise that other people may have false beliefs about a situation, which can lead them to behave in a particular way.

We also administered a test of "second order" theory of mind, i.e. understanding related to a belief about another person's belief. For the Australian children a version of Bowler's (1992) shopping story was used; and in Britain, a version of Perner and Wimmer's (1985) Ice Cream story was used. Both stories involve two characters who want to buy something, and through a series of events they develop differing knowledge states. The subject is asked to solve the problem of the kind "X *thinks* Y *thinks* that…", and then to predict a character's behaviour on the basis of his/her false belief. The two stories were used to allow for cultural differences between British and Australian children. For example the notion of a village, which is in the British story, is not part of Australian culture, and we wanted to make the stories as consistent with children's experience as possible. Comparisons across the two samples and the two stories showed no significant differences in the numbers, or in the characteristics of children passing or failing the theory of mind questions. Moreover, Bowler (1992), in his study of older autistic individuals, found no differences on performance with two stories of this kind.

Procedure

The checklist was completed during interview with parents in the family home or in the clinic. In some cases, questions are only relevant to either early (e.g. babbling) or current history (e.g. maintaining friendships at school or work); hence developmentally relevant data were incorporated in the schedule. The interview and child assessment measures took, on average, about 3 hours for each child and family.

Table 1
Number and Percentage of Particular Clinical Diagnoses in Each Cluster Group

Cluster group	Clinical diagnosis: Autism	Asperger syndrome	Other
Cluster A	22 (45.8%)	15 (21.7%)	7 (38.8%)
Cluster B	14 (29.2%)	40 (58.0%)	1 (5.6%)
Cluster C	12 (25.0%)	14 (20.3%)	10 (55.6%)

Table 2
Cluster Group Differences

Variable/ Cluster group	Mean (*SD*)	
CA ($N = 135$)		
A ($N = 44$)	8.94 (4.7)	
B ($N = 55$)	11.46 (3.2)	$F(2,132) = 4.97$, $p < .01$
C ($N = 36$)	9.91 (4.3)	
VMA ($N = 110$)		
A ($N = 34$)	7.29 (3.6)	
B ($N = 53$)	10.37 (4.9)	$F(2,107) = 8.39$, $p < .001$
C ($N = 23$)	12.26 (5.9)	
PPVT-R[a] ($N = 110$)		
A ($N = 34$)	75.8 (28.5)	
B ($N = 53$)	91.4 (18.5)	$F(2,107) = 7.57$, $p < .001$
C ($N = 23$)	96.9 (18.5)	
Age[b] ($N = 135$)		
A ($N = 44$)	6.55 (3.6)	
B ($N = 55$)	8.54 (4.0)	$F(3,132) = 3.18$, $p < .05$
C ($N = 36$)	7.55 (4.1)	

[a] PPVT-R standard score.
[b] Age in years when diagnosed.

Results

Methods of Analysis

A cluster analysis was performed to classify participants into groups on the basis of autistic behaviour derived from their scores from the Autism Spectrum Disorders Checklist. Snob is a cluster analysis program that uses an information theoretic decision statistic, called the message length, as an objective function to be minimised. The Snob program implements the Minimum Message Length (MML) principle by endeavouring to allocate, divide, and merge subjects into clusters (Wallace & Dowe, 1994; Wallace & Freeman, 1987). Snob holds two advantages over other cluster analytic techniques. First, the Snob algorithm determines the number of clusters that best classifies the data, so there is no need to specify the number of clusters a priori. The message length index can be used to identify the most parsimonious solution. Second, Snob can handle missing data by simply assuming a fixed constant cost for each missing attribute value, thereby rendering the missing values an irrelevancy to the minimisation of the message length. Snob has been used for taxonomic purposes in autism (Prior et al., 1975), and other areas of psychiatric research such as depression and family functioning (see Kissane et al., 1996, for details). Data from the checklist were used for the Snob cluster analysis. Five variables from the checklist for which less than 10% of Yes responses occurred were eliminated from the Snob procedure. These included: no babbling, no spoken language, no response to communication, lack of spontaneous activity, and smearing faeces.

Having obtained the clusters from the Snob analysis, chi-square tests were conducted on all items to determine the manner in which the clusters differed in terms of autistic behaviour, developmental history, and family variables. Indeed, Snob also produces log-likelihood ratio estimates, which can be used to indicate the variables that best discriminate between the clusters. Results of these estimates and those of chi-square tests are usually consistent with each other.

Table 3
Significantly Differentiating Variables ($p < .001$)

Variable	Percentage in cluster groups Cluster A ($N = 44$)	Cluster B ($N = 55$)	Cluster C ($N = 36$)
Social Domain			
No anticipation of being held	84.1	36.5	30.6
Dislikes physical affection	79.5	47.3	30.6
Impaired use of nonverbal signals during social interaction	68.2	90.7	38.9
Does not spontaneously wave goodbye	54.5	76.4	30.6
Does not spontaneously say hello	81.8	74.5	44.4
Inappropriate selection of person to whom to show affection	55.0[a]	51.0[f]	16.7
No peer friendship	63.6	61.8	19.4
Wants friends (but has a poor grasp of friendship)	54.5	87.3	41.7
Has one "friend" with same circumscribed interest	13.6	55.6[e]	27.8
Makes embarrassing remarks in public	100.0[b]	70.0[g]	46.4[i]
No reciprocation in simple games	56.8	25.5	13.9
Does not point things out to share pleasure/interest	81.8	50.9	16.7
Does not bring toys, objects etc. for shared pleasure/interest	79.5	50.9	13.9
Communication Domain			
Conversation is one-sided, repetitive, without appropriate turn-taking	65.9	85.5	41.7
No response to instructions except in familiar context	72.7	20.0	33.3
Fails to use context in comprehension	72.7	40.0	19.4
Literal understanding of phrases	97.1[c]	87.3	56.7[j]
Immediate echolalia/delayed echolalia	56.8	36.4	16.1[k]
Idiosyncratic gestures	41.0[d]	59.6[h]	16.2
Repetitive/Stereotyped Behaviours Domain			
Unusual sensory responses (e.g. smelling inert objects, scratch surfaces, fascination with sounds, lights etc.)	63.6	70.9	16.7
Collecting, or unusual fascination with particular objects	29.5	72.7	16.7
Maintenance of sameness of environment	54.5	74.5	16.7
Maintenance of sameness of routines	45.5	60.0	13.9
Limited pattern of self-chosen activities	75.0	80.0	22.2
Repetitive questions/talk on repetitive themes	68.2	90.9	52.8
Special skills	27.3	78.2	25.0

Variations in sample size, either because behaviour not relevant (e.g. child has insufficient language to make embarrasing remarks in public), or parents felt unable to provide responses (e.g. did not know if literal responses were made): [a]$N = 40$; [b]$N = 29$; [c]$N = 35$; [d]$N = 39$; [e]$N = 54$; [f]$N = 51$; [g]$N = 50$; [h]$N = 47$; [i]$N = 28$; [j]$N = 30$; [k]$N = 31$.

SNOB Three Cluster Solution

An exhaustive Snob search produced a convergent three cluster solution (Table 1). Cluster A contained 44 individuals, Cluster B contained 55 individuals, and Cluster C 36 individuals. Almost half (45.8%) of the children with a clinical diagnosis of autism were grouped in Cluster A, with the remaining subjects roughly evenly spread in Cluster B (29.1%) and Cluster C (25%). More than half of the originally diagnosed AS children were grouped in Cluster B (58%), whereas most of the children diagnosed with "Other" diagnoses (such as PDD-NOS) were found in Cluster C (55%) (see Table 1).

Group Differences

Table 2 shows the differentiating variables from the taxonomic analysis that were significant at an alpha level of .001 and less. This alpha level was set because of the large numbers of comparisons made and the consequent risk of chance differences. Significance levels are based on chi-square comparisons. ANOVA tests were conducted

Table 4
Results of Theory of Mind Comparisons

Theory of mind study	Cluster A	Cluster B	Cluster C	Significance
Sally-Anne (1st order)				
Number of subjects tested	34	54	27	
Pass	15	50	25	
Fail	13	3	2	$\chi^2(2) = 24.14$, $p < .00$
Fail control questions	6	1	0	
Number of subjects not tested	10	1	9	
Smarties (1st order)				
Number of subjects tested	33	54	27	
Pass	15	49	24	$\chi^2(2) = 16.08$, $p < .000$
Fail	12	5	3	
Fail control questions	6	0	0	
Number of subjects not tested	11	1	9	
Overcoat/Ice-cream (2nd order)				
Number of subjects tested	15	49	23	
Pass	7	28	14	n.s.
Fail	6	12	7	
Fail control questions	2	9	2	
Number of subjects not tested	2	1	2	

on chronological age (CA), VMA, standard score, and age at diagnosis data.

Background variables including developmental history and family factors. There were no significant cluster group differences on: child gender, history of language delay or deviance; presence/absence of delays in sitting or walking; family history of cognitive or emotional disorder. Cluster A cases were somewhat more likely to have had a delay in crawling when compared to the other two clusters ($p < .05$). Cluster A cases were also more likely to have been diagnosed at an earlier age than the other two clusters, and were significantly younger than Cluster B but not Cluster C cases. Cluster A cases had significantly lower PPVT, VMA, and standard scores than Clusters B and C.

Diagnostic variables. The aim of these analyses was to identify any behaviours that discriminated significantly between the three cluster groups and that therefore might have differential diagnostic import.

From Table 3 it can be seen that on most of the variables the most frequent "autism-like" responses were observed in Cluster A, followed by B and then least frequent, Cluster C. The Cluster A group were more likely to be different on variables which rely on an "awareness" of others, such as anticipating being held, greeting behaviours, making embarrassing remarks in public, pointing out things of interest, and showing interest in reciprocal games. Cluster A cases were also more likely to show problems in the communication domain, such as lack of responding to instructions, echolalia, echopraxia, and to have a literal understanding of language. In the area of repetitive, restricted, and stereotyped behaviours, Cluster A were more similar to Cluster B than to Cluster C, with a greater likelihood of showing sensory disturbances (such as sound and light sensitivities), preference for sameness of environment and routines, and a limited pattern of self-chosen activities.

Cluster B were more likely to desire friends but to have a poor grasp of friendship; to make communicative approaches that are one-sided and repetitive, to have sensory disturbances, to collect objects, to seek to maintain the sameness of their environment and routines, and to possess a special skill. Cluster B was like Cluster A on behaviours such as problems with greeting, poor discrimination of person to whom to show affection, no real peer friendships, having a literal understanding of language, using idiosyncratic gestures, and having a limited pattern of self-chosen activities.

Cluster C was most different from Cluster A and more behaviourally similar to Cluster B. On almost all behaviours they were less impaired than the other cluster groups. They were similar to Cluster A on behaviours such as asking repetitive questions and rarely having a marked special skill. They were more similar to Cluster B on behaviours such as anticipating being held, responding to affection, and were more likely than Cluster A to be able to respond to instructions and engage in more spontaneous activities.

Since the age range of the sample was very wide it was decided to run the Snob analyses a second time, excluding the very young children (less than 5 years of age), giving a sample of 125 cases. Three cluster groups emerged in this second analysis with similar characteristics to those derived from the total group. These groups did not differ on CA, but Cluster A was of lower VMA than Clusters B and C. In comparing the differentiating behavioural features in this second analysis, some of the originally significant variables could not be considered significant because of very small cell numbers, rendering chi-square comparisons illegal (Howell, 1992). However, they were the same variables that distinguished the groups in the original total sample. It was therefore decided to retain the original Snob analysis.

Theory of mind comparisons. It was not possible to test all subjects on the theory of mind tasks. Twenty-one subjects were not tested on the Smarties task and 20 were not tested on the Sally-Anne task. These children were not tested because they were noncompliant. None of these was tested on the second order theory of mind tasks. The children who were not tested on theory of mind measures proved to be the youngest of the sample, with a mean CA of 5.47 years ($SD = 2.9$).

Subjects who passed the first order theory of mind tasks (i.e. answered the test and control questions correctly) were tested on a second order theory of mind task. Failure on these tasks was defined as failing the test question or failing any of the control questions (regardless of whether the test question was answered correctly). The results show that on the first order tasks, Cluster A cases were significantly less likely to pass compared to Clusters B and C (Table 4). Univariate comparisons conducted between these three clusters on age and mental age measures were conducted. No difference was found between groups on CA. Cluster A was found to have a significantly lower mean VMA compared to Cluster C, and a significantly lower mean PPVT standard score compared to Clusters B and C [$F(1,99) = 5.7$, $p < .04$].

On face value, no differences were found between cluster groups on passing second order theory of mind tasks. There were no differences on CA, VMA, and standard scores between the groups of children who took part in the second order theory of mind tests. This is due to the fact that the children in Cluster A who were able to be tested on second order theory of mind tasks and pass the control questions had significantly higher mean VMA and standard scores. Since it is the case that children of higher VMA are more likely to be eligible to be tested on second order theory of mind (Baron-Cohen, 1989), only the most able Cluster A cases were included. These children constituted less than a third of possible Cluster A cases. If we make the reasonable assumption that children who failed first order theory of mind tasks would also have failed the more complex second order tasks, then we find that Cluster A (8 out of 34 cases, or 23.5%) are significantly less likely to pass than Clusters B (28 out of 54 or 51.9%) and C (14 out of 27 or 51.9%) [$\chi^2(2) = 7.82$, $p = .02$].

Discussion

The first aim of this study was to determine whether subgroups of autistic-type children could be empirically derived from cluster analysis of current behavioural characteristics of a large group of diagnosed cases. Results showed that it was possible to differentiate three groups or clusters of autistic-type individuals roughly corresponding to those familiar to us through clinical experience (A—autistic-like, B—Asperger-like, and C—mild PDD or PDD-NOS). Nevertheless it should be noted that within this sample of high-functioning children, those with the independently obtained original *clinical* diagnosis of autism were as likely to be empirically subgrouped with those who had received an AS or "other" PDD diagnosis, as they were to be members of Cluster A. Moreover, almost 30% of diagnosed AS individuals were in Cluster A.

The notable differentiating features of Cluster B were those relating to somewhat higher levels of social and communicative development. Many of these children looked for friendships, albeit in clumsy and not very successful ways; and they were more likely to have pedantic-style, egocentric conversations, usually focused on their own current preoccupations or special interests. Such behaviours were not absent in the other groups but were significantly less common. We would argue that the presence of these characteristics reflects the capacity of knowing about and wanting friendships with other children, and of being able to develop special interests, albeit to a rather obsessive extent, thus reflecting somewhat higher levels of functioning by comparison particularly with Cluster A. Half of the Cluster B children and most of Cluster C children had shown joint attention skills, compared with only 18% of Cluster A, again indicating more normal development in this domain. One interpretation of the group comparisons is that severity of symptoms, rather than distinctive symptom patterns, was an underlying factor in the clustering.

Alternatively, these findings could argue for differential diagnosis based on the greater likelihood of particular symptoms, such as those noted above, in one group compared with the other. One might say that Cluster B (Asperger-like) children had some awareness of, and drive towards, friendship and social relationships; and that they were keen to communicate their interests, whether in joint attention behaviour or in conversations about their preoccupations, and they were also more likely to have special skills and interests related to higher cognitive capacities (e.g. architectural drawing, computer skills). This made them different from the more self-isolated Cluster A children and could support the concept of a subcategory of children who are not so profoundly socially impaired and whose higher level of language development enabled them to contact other people more readily, even if their language was egocentric (86% with one-sided and repetitive conversations in Cluster B). Table 3 illustrates the less abnormal nature of the language in Cluster B. By this argument, Cluster C children were differentiated as a group on the basis of much fewer problems across all domains, perhaps supporting a PDD-NOS category on the basis of subthreshold symptoms (Bishop, in press). These characteristics are also integral to a higher ability level and this too differentiated Clusters B and C from Cluster A.

One could also argue that the results support the spectrum concept of autistic-type disorders, rather than clearly distinct categories, with the spectrum based on severity of behavioural and cognitive/communicative impairment. Severity of impairment here refers to the *nature* of the differentiating symptoms, which in the case of Cluster B and C appear less distant from normal behaviours. For example, knowing about and seeking friendship represents less impairment than ignoring or isolating oneself from others, as does the ability to respond to instructions in unfamiliar contexts (Table 3). In other words, the group comparison data could be interpreted on the basis of relative severity of symptoms and levels of cognitive functioning (which are no doubt related factors). The symptom pattern was associated with the higher level of abilities evident in Clusters B and C, both in verbal comprehension and mentalising performance.

Since all of the children with AS also met conventional diagnostic criteria (DSM-III-R, ICD-10) for autism, this too perhaps supports a spectrum concept. That is, these may be conceptualised as all being autistic-type children, with the B and C Clusters the upper end of the spectrum in terms of behavioural, cognitive, and communicative functioning. It is also pertinent to note that in each domain many behaviours did *not* discriminate significantly between the groups (60 % of social domain items, 50 % of communication items, 52 % of repetitive behaviour items). Some of the items on which the groups did not differ in the social domain were: impaired eye contact, problem with recognising other person's personal space, poor relationships with peers, inappropriate responses to other people's emotions, lack of imaginative ability; in the communication domain: unusual tone of voice, poor nonverbal communication during speech, use of idiosyncratic language; and in the repetitive behaviour domain: self-injurious behaviour, clinging to objects, marked arranging of objects, marked repetitive actions with objects, marked interest in the abstract properties of objects.

Early history characteristics of this sample including language development did not provide evidence for group differentiation either (see also Eisenmajer et al., in press). The absence of cluster group differences on history of language delay or deviance needs to be highlighted. It suggests that this may not be a differentiating feature that can reliably be used in diagnosis.

Comparison of clinician diagnosis with cluster group membership suggests that Cluster B children were more likely to have received an AS diagnosis, but that Cluster A children were as likely to have received an AS diagnosis as an autism one. This could be related particularly to the fact that there has been a marked increase in clinician's use of the AS diagnosis for higher-functioning children in recent years. Chi-squared comparisons showed that there were no systematic differences between clinicians (nor between Australian and British children) in the likelihood of an AS rather than an autism diagnosis, however. The sample in this study was a relatively high-functioning one, excluding the substantial proportion of children who are resistant to assessment on standard tests; hence we sampled only the upper part of the spectrum, and our conclusions must be limited to high-functioning individuals.

The second aim of the study was to establish whether performance on theory of mind tasks might differentiate empirically derived subgroups. Results confirmed the importance of ability/age variables, which have increasingly been emerging in studies of this genre (e.g. Happé, 1994). Few Cluster B and C children failed first order tasks, and more than half could also pass the second order task. Comparison with results from Bowler's (1992) study with British adolescents and adults suggests that a similar proportion of his sample with AS passed the second order task. It is also noteworthy that 54 % of Cluster A (autistic-like) individuals passed first order theory of mind tasks, and of that passing subsample who were presented with the second order task, 54 % passed. If first order failers were also assumed to fail second order, significant differences were found between Cluster A and Clusters B and C. We argue that this is probably a better representation of the group differences in ability to pass a theory of mind task, since limiting the second order results to only children who could pass first order would severely restrict the composition of the groups to the most able children.

The results also suggest that when verbal ability is not too far from an average level, theory of mind deficits are less likely to be evident. We suggest that mentalising deficits may be primarily associated with developmental cognitive and language delay rather than simply autistic disorders (for comparable results with hearing impaired children see Peterson & Siegel, 1995). That is, a child's level of cognitive and language competence has a general influence on all domains of functioning, of which mentalising ability is but one. The same influences moderate the nature and severity of behavioural symptoms. Hence, we argue that the results of this research support the concept of a spectrum of autistic disorders in which severity of social and communicative impairments underlie individual differences in the cognitive, behavioural, and adaptive functioning deficits observed.

In their study comparing two taxonomically derived groups (which broadly comprised a core autistic cluster and an "other PDD" cluster) with DSM and ICD systems, and with Wing and Gould's (1979) categorisation of "active but odd", "passive", and "aloof", Waterhouse et al. (1996) considered whether the obtained grouping reflected differentiation on the basis of level of functioning, or was based on a severity continuum or spectrum. This question is also of importance to our interpretations. Bearing in mind that we had an older sample, and three groups rather than two (compared with Waterhouse et al., 1996, our "less autistic-like" cluster subdivided into AS-like and "mild", or PDD-NOS), our data seem to support an interpretation based on what Waterhouse et al. (1996) called "severity of developmental compromise". We would also argue that this is not inconsistent with the spectrum conceptualisation of autism. The theory of mind results demonstrated that mental age in the verbal domain was important in assisting children with considerable social impairments to pass theory of mind tasks. We have gone one step further than Waterhouse et al. in showing that a sample of children with autism spectrum diagnoses can be empirically subdivided according to severity of selected social impairments and theory of mind ability, and the latter is related to verbal capacity. It is probable that the distinguishing symptoms of our Cluster B (limited friendships, pedantic speech, and circumscribed interests) are also related to level of cognitive functioning (see Tsai, 1992; Prior & McMillan, 1973), suggesting that it is this that is often primary in influencing an AS diagnosis in clinical practice (Eisenmajer et al., 1996).

The fact that the developmental history variables did not discriminate between the subgroups in any diagnostically meaningful way (particularly the language development variable) suggests the need for caution in using such data for differential diagnosis. It raises the possibility that aetiology too may support a continuum concept. Although our taxonomic study has produced three clusters or subgroups that show some differences on the marker we selected to assess validity, viz. theory of mind, we have argued that a close look at the findings suggests a spectrum of autistic disorders with severity of social and cognitive impairment being the primary basis for any group differences. The results also clearly demonstrate, as have other studies (Bishop, in press; Ehlers & Gillberg, 1993; Wing & Gould, 1979) that subgroupings within the autistic spectrum are not confined only to the subtypes of autism or AS but also include children with autism spectrum disorder who do not fit these clinical descriptions.

References

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

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., Leslie, A. M., & Frith, U. (1985). Does the autistic child have a "theory of mind"? *Cognition*, *21*, 37–46.

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1986). Mechanical, behavioural and intentional understanding of picture stories in autistic children. *British Journal of Developmental Psychology*, *4*, 113–125.

Bishop, D. (in press). What's so special about Asperger's Disorder? The need for further exploration of the borderlands of autism. In A. Klin, F. Volkmar, & S. Sparrow (Eds.), *Asperger syndrome*. New York: Guilford Press.

Bowler, D. (1992). "Theory of mind" in Asperger's syndrome. *Journal of Child Psychology and Psychiatry*, *33*, 877–894.

Cantwell, D., & Rutter, M. (1994). Classification: Conceptual issues and substantive findings. In M. Rutter, E. Taylor, & L. Hersov, *Child and adolescent psychiatry: Modern approaches* (3rd edn.) (pp. 2–21). Oxford: Blackwell.

Castelloe, P., & Dawson, G. (1993). Subclassification of children with autism and pervasive developmental disorders: A questionnaire based on Wing's subgrouping scheme: *Journal of Autism and Developmental Disorders*, *23*, 229–241.

Dunn, L. M., & Dunn, L. M. (1981). *Manual for the Peabody Picture Vocabulary Test-Revised*. Circle Pines, MN: American Guidance Service.

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

Eaves, L. C., Ho, H. H., & Eaves, D. M. (1994). Subtypes of autism by cluster analysis. *Journal of Autism and Developmental Disorders*, *24*, 3–22.

Ehlers, S., & Gillberg, C. (1993). The epidemiology of Asperger syndrome. A total population study. *Journal of Child Psychology and Psychiatry*, *34*, 1327–1350.

Eisenmajer, R. J., & Prior, M. (1991). Cognitive-linguistic correlates of "theory of mind" in autistic children. *British Journal of Developmental Psychology*, *9*, 351–364.

Eisenmajer, R. J., Prior, M., Leekam, S., Wing, L., Gould, J., Welham, M., & Ong, B. (1996). A comparison of clinical symptoms in autism and Asperger's Disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, *35*, 1523–1531.

Eisenmajer, R. J., Prior, M., Leekam, S., Wing, L., Ong, B., Gould, J., & Welham, M. (in press). Delayed language onset as a predictor of clinical symptoms in pervasive developmental disorders. *Journal of Autism and Developmental Disorders*.

Fombonne, E. (in press). Epidemiology of autism and related conditions. In F. Volkmar (Ed.), *Autism and the pervasive developmental disorders*. Cambridge: Cambridge University Press.

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

Frith, U. (1991). *Autism and Asperger syndrome*. Cambridge: Cambridge University Press.

Frith, U., Happé, F., & Siddons, F. (1994). Autism and theory of mind in everyday life. *Social Development*, *3*, 108–124.

Ghaziuddin, M., Tsai, L. Y., & Ghaziuddin, N. (1992). Brief report: A comparison of the diagnostic criteria for Asperger syndrome. *Journal of Autism and Developmental Disorders*, *22*, 643–649.

Gillberg, C. (1989). Asperger's syndrome in 23 Swedish children. *Developmental Medicine and Child Neurology*, *31*, 520–531.

Gillberg, C. (1992). Autism and autistic-like conditions: Subclasses among disorders of empathy. *Journal of Child Psychology and Psychiatry*, *33*, 813–842.

Happé, F. G. E. (1994). Annotation: Current psychological theories of autism: The "theory of mind" account and rival theories. *Journal of Child Psychology and Psychiatry*, *35*, 215–229.

Howell, D. C. (1992). *Statistical methods in psychology* (3rd Edn.). Boston, MA: Duxbury Press.

Kissane, D. W., Bloch, S., Dowe, D. L., Snyder, R. D., Onghena, P., McKenzie, D. P., & Wallace, C. (1996). The Melbourne family grief study, 1: Perceptions of family functioning in bereavement. *American Journal of Psychiatry*, *153*, 650–658.

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

Manjiviona, J., & Prior, M. (1996). *Neuropsychological profiles of autistic and Asperger syndrome children*. Paper presented at the Congress Autism-Europe, Barcelona, May 1996.

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

Perner, J., Leekam, S., & Wimmer, H. (1987). Three-year-old's difficulty with false belief: The case for a conceptual deficit. *British Journal of Developmental Psychology*, *5*, 125–137.

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

Peterson, C., & Siegal, M. (1995). Deafness, conversation and theory of mind. *Journal of Child Psychology and Psychiatry*, *36*, 459–474.

Prior, M. R., & MacMillan, M. B. (1973). Maintenance of sameness in children with Kanner's syndrome. *Journal of Autism and Childhood Schizophrenia*, *3*, 154–167.

Prior, M., Perry, D., & Gajzago, C. (1975). Kanner's syndrome or early onset psychosis: A taxonomic analysis of 142 cases. *Journal of Autism and Childhood Schizophrenia*, *5*, 71–80.

Rapin, I. (1996). Preschool children with inadequate communication: Developmental language disorders, autism or low IQ. *Clinics in Developmental Medicine*, No. 139. London: Mac Keith Press.

Sattler, J. (1988). *Assessment of children* (3rd Edn.). San Diego, CA: Jerome Sattler Publishers.

Schopler, E. (1985). Convergence of learning disability, higher-level autism, and Asperger's syndrome. *Journal of Autism and Developmental Disorders*, *15*, 359.

Sevin, J. A., Matson, J. L., Coe, D., Love, S. R., Matese, M. J., & Benavidez, D. A. (1995). Empirically derived subtypes of pervasive developmental disorders: A cluster analytic study. *Journal of Autism and Developmental Disorders*, *25*, 561–578.

Siegel, B., Anders, T. F., Ciarenello, R. D., Bienenstock, B., & Kraemer, H. C. (1986). Empirically derived subclassification of the autistic syndrome. *Journal of Autism and Developmental Disorders*, *16*, 275–294.

Szatmari, P. (1992). The validity of autistic spectrum disorders: A literature review. *Journal of Autism and Development Disorders*, *22*, 583–600.

Szatmari, P., Archer, L., Fisman, S., Streiner, D. L., & Wilson, F. (1995). Asperger's syndrome and autism: Differences in behaviour, cognition and adaptive functioning. *Journal of the American Academy of Child and Adolescent Psychiatry*, *34*, 1662–1671.

Szatmari, P., Tuff, L., Finlayson, A. J., & Bartolucci, G. (1990). Asperger's syndrome and autism: Neurocognitive aspects. *Journal of the American Academy of Child and Adolescent Psychiatry*, *29*, 130–136.

Tsai, L. Y. (1992). Diagnostic issues in high-functioning autism. In E. Schopler & G. B. Mesibov (Eds.), *High-functioning individuals with autism* (pp. 11–40). New York: Plenum Press.

Wallace, C. S., & Dowe, D. L. (1994). Intrinsic classification by MML—the Snob program. In C. Zhang & J. Debenham (Eds.), *Artificial intelligence, sowing seeds for the future* (pp. 37–44). Singapore: World Scientific.

Wallace, C. S., & Freeman, P. R. (1987). Estimation and inference by compact coding. *Journal of the Royal Statistical Society, Series B, Methodology*, *49*, 223–265.

Waterhouse, L., Morris, R., Allen, D., Dunn, M., Fein, D., Feinstein, C., Rapin, I., & Wing, L. (1996). Diagnosis and classification in autism. *Journal of Autism and Developmental Disorders*, *26*, 59–86.

World Health Organisation. (1993). *International classification of diseases and disorders* (*ICD-10*). Geneva: Author.

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

Wing, L., & Gould, J. (1979). Severe impairments of social interaction and associated abnormalities in children: Epidemiology and classification. *Journal of Autism and Developmental Disorders*, *9*, 11–29.

Manuscript accepted 19 December 1997