Two-Year Outcome of Preschool Children With Autism or Asperger's Syndrome

Peter Szatmari, M.D.

Susan E. Bryson, Ph.D.

David L. Streiner, Ph.D.

Freda Wilson, B.A.

Lynda Archer, Ph.D.

Cynthia Ryerse, M.D.

Objective: DSM-IV specifies that Asperger's disorder is a type of pervasive developmental disorder without clinically significant cognitive or language delay. There are no data, however, on the outcome of children with Asperger's disorder or on whether their outcome differs from that of children with autism. The objectives of this study were to compare the outcome of groups of children with these disorders over a period of 2 years on variables independent of the defining criteria and to identify variables that might account for these differences.

Method: All children 4–6 years of age who came for assessment or were currently in treatment at a pervasive developmental disorder service of one of several centers in a large geographic region were identified. Children who received a diagnosis of autism (N=46) or Asperger's syndrome (N=20) on the basis of a diagnostic interview and had an IQ in the nonretarded range were given a battery of cognitive, language, and behavioral tests. Families were contacted roughly 2 years after the date of their enrollment in the study, and many of the tests were readministered.

Results: Children with Asperger's syndrome had better social skills and fewer autistic symptoms 2 years after study enrollment than the children with autism. The differences in outcome could not be explained by initial differences in IQ and language abilities. Children with autism who had developed verbal fluency at follow-up were very similar to the children with Asperger's syndrome at study enrollment.

Conclusions: Although the exact mechanism for the differences in outcome remain to be determined, it appears that Asperger's disorder and autism represent parallel but potentially overlapping developmental trajectories.

(Am J Psychiatry 2000; 157:1980–1987)

Children with pervasive developmental disorders demonstrate qualitative impairments in reciprocal social interaction and in verbal and nonverbal communication and engage in a pattern of repetitive, stereotypic activities (DSM-IV). "Pervasive developmental disorder" is a general term referring to a spectrum of disorders that differ with respect to the number or type of symptoms or age at onset. Autism is the best-known example of a pervasive developmental disorder subtype, but, through the years, a number of other subtypes have been proposed as well. These include atypical autism, Asperger's disorder, and disintegrative disorder (DSM-IV). The diagnostic boundary between autism and each of these other subtypes is unclear and varies according to the classification scheme (1).

Asperger's disorder is defined by marked difficulties in reciprocal social interaction, fluent but pragmatically impaired speech, and bizarre preoccupations and obsessions. The key features that distinguish autism from Asperger's disorder are that the latter is characterized by a lack of clinically significant language and cognitive delay (DSM-IV and ICD-10). However, the data on whether Asperger's disorder is a valid subtype of pervasive developmental disorder are both limited and inconsistent (2, 3).

Although there are several outcome studies of children with nonautistic pervasive developmental disorder (4–8), no study has specifically examined Asperger's disorder. Now that many more children are receiving this diagnosis, outcome studies are essential to evaluate the clinical usefulness of the category and to provide information to parents. In a previous publication based on cross-sectional data (9), we reported that 4–6-year-old children with Asperger's disorder have better language and socialization scores and fewer autistic behaviors than children with high-functioning autism, but whether this is a useful distinction depends, in part, on outcome evaluations.

In this study, we wanted to see if there were differences in outcome at 2 years between children with autism and children with Asperger's disorder. The primary outcome variables were measures of social competence and a composite measure of autistic symptoms, variables that were not used in the initial clinical distinction but are clinically relevant to long-term outcome. Two secondary outcome variables were also considered: measures of language and visual-spatial abilities. Although measures of expressive and receptive language may be related to the way the groups were defined at the beginning of the study (that is,

TABLE 1. Characteristics of Children With High-Functioning Autism or Asperger's Syndrome Participating in Outcome Study

Characteristic	Children With Autism (N=46)				Children With Asperger's Syndrome (N=20)			
	N	%	Mean	SD	N	%	Mean	SD
Age (months)								
At study enrollment			65.1	11.7			68.5	9.1
At 2-year follow-up			90.8	11.5			93.4	9.2
Female sex	3	6.5			4	20.0		
Class placement in school								
Special class	8	17.4			4	20.0		
Partial integration	14	30.4			6	30.0		
Regular class	24	52.2			10	50.0		
Leiter IQ								
At study enrollment			86.6	17.9			99.3[a]	16.9
At 2-year follow-up			82.4	17.9			92.6[b]	13.5

[a] Significant difference between groups (F=7.66, df=1, 66, p=0.007).
[b] Significant difference between groups (F=4.69, df=1, 64, p=0.04).

delayed and deviant language), it is still important to see whether the strong differences between the groups at 4–6 years of age were still present or had attenuated, as might be expected if the clinical distinction had little impact on outcome or was subject to large amounts of measurement error. Further, in view of the report by Klin et al. (10) that children with Asperger's disorder have less visual-spatial and motor ability than children with autism, we wanted to see whether this difference had emerged at 6–8 years of age, even though we had not seen this at 4–6 years.

Method

Subjects

All children 4–6 years of age who either came for assessment or were currently in treatment at a pervasive developmental disorder service of six different centers that serve preschool children with developmental disabilities in southern Ontario were identified. Children who received a clinical diagnosis of pervasive developmental disorder but were untestable or achieved a mental age score less than half their chronological age on psychometric testing were dropped from further consideration. The remaining children, including those without psychometric data, entered the study; written informed consent was obtained from the parents of each child. The children's assent was also obtained on the basis of their willingness to participate in the assessment.

Children who received a diagnosis of autism or Asperger's syndrome on the basis of data from the Autism Diagnostic Interview (1988 version) (11) and who had either a Leiter IQ above 68 or a Stanford-Binet IQ above 70 were included in the final study group. Because we used criteria that are different from those of DSM-IV and ICD-10, children in our study group are described as having Asperger's "syndrome" rather than Asperger's "disorder."

During the enrollment phase, 164 children 4–6 years of age were screened. Of these, 80 children were excluded because either they did not have pervasive developmental disorder, their behavior was too low functioning to get an IQ estimate, or previous psychometric testing revealed that they were functioning below the mental age criterion. The remaining 84 children with pervasive developmental disorder underwent the full psychometric battery. A further 16 were excluded because their Leiter or Stanford-Binet IQs were below the IQ cutoff for mental retardation (IQ of 68 or 70, respectively). Thus, the size of the study group at enrollment was 68 nonretarded preschool children with pervasive developmental disorder.

Information from the Autism Diagnostic Interview was used to classify children according to ICD-10/DSM-IV as having either Asperger's syndrome or autism. Because a diagnostic algorithm from the Autism Diagnostic Interview is not available for Asperger's disorder, we derived an algorithm on the basis of our previous work and reviews of the literature (3, 12, 13).

To qualify for a diagnosis of Asperger's syndrome, the child had to have at least one example of an impairment in reciprocal social interaction (as opposed to autism, which stipulates that there must be at least two examples in this area), at least one example of an impairment in verbal or nonverbal communication (same as for autism), and at least one example of a repetitive, stereotyped behavior (same as for autism). These qualifications were specified by using the items from the Autism Diagnostic Interview diagnostic algorithm for autism. Moreover, the diagnosis of Asperger's syndrome required an absence of clinically significant language delay (i.e., spontaneous phrase speech by 36 months of age) and an absence of marked or persistent deviant language development such as delayed echolalia, pronoun reversal, and neologisms (coded as 2 or 3) on these items from the Autism Diagnostic Interview. In fact, the majority of children with Asperger's syndrome had many examples of impairments in the three domains; their mean Autism Diagnostic Interview score was 12.5 (SD=3.5) on the social domain (autism cutoff=10), 11.8 (SD=2.9) on the communication domain (autism cutoff=8), and 6.0 (SD= 2.5) on the repetitive activities domain (autism cutoff=4).

In contrast, children with the diagnosis of autism spoke after 36 months and/or had evidence of marked deviance in language development (as defined for Asperger's syndrome) and met the diagnostic algorithm for autism as specified by Le Couteur et al. (11).

We used a hierarchical method to classify the 68 children into the two pervasive developmental disorder subtypes according to these criteria; that is, a diagnosis of Asperger's syndrome had precedence over a diagnosis of autism. Using these diagnostic rules, all 68 children could be classified; 47 met our criteria for autism, and 21 met our criteria for Asperger's syndrome.

Our criteria for Asperger's syndrome are consistent with, but not identical to, the ICD-10 and DSM-IV criteria for Asperger's disorder, which were published after data collection was completed. In fact, strictly applying the DSM-IV criteria (as operationalized in the Autism Diagnostic Interview algorithm) would have resulted in only one child receiving a DSM-IV diagnosis of Asperger's disorder. Of the 21 children who met our criteria for Asperger's syndrome, 15 also met criteria for autism on the Autism Diagnostic Interview and so failed to meet the DSM-IV definition for Asperger's disorder (i.e., a diagnosis of autism takes precedence over a diagnosis of Asperger's disorder). Of the remaining six children, five did not meet the Autism Diagnostic Interview cutoff for autism in the domain of impairments in social reciprocity, which is necessary for a diagnosis of Asperger's disorder according to DSM-IV. As a result, only one child met the social cutoff and did not meet criteria for autism. Since there is no empirical

basis for the rule that a diagnosis of autism takes precedence over a diagnosis of Asperger's disorder (although there may be a logical one), we stipulated that children with Asperger's syndrome could meet criteria for autism as long as they did not have clinically significant language delay as already defined.

All but two of the 68 children (one with autism and one with Asperger's syndrome) were assessed at follow-up. The families of the two children had moved out of the area. The characteristics of the remaining 66 children are outlined in Table 1. The mean age at follow-up was 90.8 months in the autistic group and 93.4 months in the Asperger's syndrome group. A full description of the characteristics of the groups at the beginning of the study is given elsewhere (9).

Procedure

The time 1 assessment consisted of two IQ tests, tests of language and nonverbal abilities, a semistructured interview (the Autism Diagnostic Interview [11]), the Autism Behavior Checklist [14, 15], and the Vineland Adaptive Behavior Scales [16, 17]. Families were contacted roughly 2 years after the date of their enrollment in the study (mean=26 months, range=21–30). The assessment conducted at follow-up included a full psychometric battery, parent-completed questionnaires, and structured interviews. All follow-up assessments were administered by a different research assistant than the one who did the original assessments, thus maintaining blindness. Some children who failed the initial IQ screen were recalled to see whether we missed any truly nonretarded children with pervasive developmental disorder. None of the recalled children scored above the IQ cutoff at the follow-up evaluation.

Instruments

Autism Diagnostic Interview (11). This semistructured interview, administered to parents, was designed to make a diagnosis of autism according to both draft ICD-10 and DSM-III-R criteria.

Autism Behavior Checklist (14, 15). This self-administered checklist, filled out by parents, measures a wide range of pervasive developmental disorder symptoms. Interrater, test-retest, and validity data for total scores are excellent (interrater reliability is 0.85 [14], sensitivity is 0.80, and specificity is 0.70 [15]). Parents completed the Autism Behavior Checklist at both the beginning of the study and at follow-up. This measure was chosen as a primary outcome measure because it measures a clinically important aspect of outcome and was independent of the way the groups were originally defined.

Vineland Adaptive Behavior Scales (16, 17). The Vineland Adaptive Behavior Scales, a semistructured interview, is designed to assess adaptive behavior in socialization, communication, motor, and daily living skills. Scores derived from this instrument have a mean of 100 and a standard deviation of 15. The interview has demonstrated excellent reliability and validity (16) and is sensitive to severity of impairment in autism (17). The Vineland Adaptive Behavior Scales was administered to parents at the beginning of the study and at follow-up. The socialization score was chosen as one of the primary outcome measures because social skills are a significant predictor of adult outcome and this variable was not used in the original distinction between the groups; that is, it does not measure impairments in social reciprocity, part of the diagnostic criteria for pervasive developmental disorder.

Leiter International Performance Scale (18). This is a standard measure of nonverbal problem solving. The Leiter is widely used with children with pervasive developmental disorder and other language impairments. It is especially appropriate for use with the children in this study because it does not require verbal instructions for administration and correlates highly with WISC-R IQs (18). The Leiter was administered at both the beginning of the study and at 2-year follow-up.

Stanford-Binet Intelligence Scale, 4th ed. (19). The Stanford-Binet measures overall cognitive development as well as four different cognitive domains—verbal reasoning, quantitative reasoning, abstract/visual reasoning, and short-term memory skills (mean=100, SD=15). The verbal reasoning domain consists of tests of vocabulary, comprehension, and absurdities. All eight subtests were given at enrollment, but only the pattern analysis subtest was given at follow-up to measure nonverbal problem-solving skills.

Language measures. All of the children were evaluated with several measures of expressive and receptive language. The Test of Language Development—2 (20) was administered at both assessment points. The grammatic understanding and grammatic completion subtests of this instrument were used to measure grammatic comprehension and usage. Standard scores (mean=10, SD=3) were calculated for each child. The McCarthy Scales of Children's Abilities (21) were also administered. In the oral vocabulary section of the McCarthy Scales, the child is asked to define 10 words, and a total score is calculated from all 10 items. The test largely assesses the child's language fluency. This test was administered at the start of the study and at follow-up, and raw scores were used. The Peabody Picture Vocabulary Test—Revised (22), a measure of single word comprehension, was administered at follow-up. We used the grammatic morphemes subtest of the Test for Auditory Comprehension of Language—Revised (23), which measures a child's understanding of grammatic structures. A standard score is obtained for each child. Studies show that the Test of Auditory Comprehension of Language produces highly reliable scores, and content, construct, and criterion-related validity have also been demonstrated (23).

Nonverbal measures. The Beery Visual-Motor Integration Test (24) is an untimed copying task that measures visual-motor integration. A standard score (mean=10, SD=3) for each child was obtained at both enrollment and at follow-up. The Annett Pegs (25), a motor dexterity task, measures the time required to move 10 pegs as quickly as possible. The block design subtest from the WISC-R (26) was also used as a measure of visual-analytic skills.

If the child was mute or was unable to reach a baseline level on the language tests, he or she was given an imputed score equal to 1 point below the lowest score obtained by a child who was able to complete the test. This was done to ensure that missing data were kept to a minimum and provided a reasonably valid, if conservative, estimate of abilities. If a child was uncooperative during testing on several occasions, the score was recorded as missing for that test.

Analysis

First, the competence in socialization (Vineland Adaptive Behavior Scales socialization scale) and the number of autistic/pervasive developmental disorder symptoms (the Autism Behavior Checklist) of the two groups at follow-up were compared by using a simple analysis of variance (ANOVA). To control for differences in these variables at the beginning of the study, the baseline socialization and Autism Behavior Checklist scores were then entered as covariates along with the interaction term to see whether the groups with Asperger's syndrome and autism differed in the relationship between baseline and follow-up scores.

To see whether the differences observed at follow-up could be explained by initial differences in nonverbal IQ and measures of language, a multiple regression analysis was performed with the socialization and Autism Behavior Checklist scores as dependent variables, the Asperger's syndrome-autism distinction as independent variables, and the Leiter nonverbal IQ, McCarthy oral vocabulary test score, and Stanford-Binet verbal reasoning score from the beginning of the study as covariates (all covariates were correlated with the outcome: $r=0.29$ to $r=0.56$, $p<0.05$ for all). The

TABLE 2. Scores on Outcome Measures of Children 6–8 Years Old With High-Functioning Autism or Asperger's Syndrome 2 Years After Study Enrollment

Outcome Measure[a]	Score at 2-Year Follow-Up						
	Children With Autism (N=46)		Children With Asperger's Syndrome (N=20)		Analysis of Variance		
	Mean	SD	Mean	SD	F	df	p
Primary measures							
Socialization	61.8	13.6	75.6	15.7	12.94	1, 63	0.001
Autism Behavior Checklist symptom subscale	58.8	25.8	37.0	20.3	11.61	1, 64	0.001
Language measures							
Communication (Vineland Adaptive Behavior Scales)	66.1	25.5	87.7	22.0	10.80	1, 63	0.002
Grammatic understanding (Test of Language Development—2)	4.4	3.6	7.8	3.9	11.90	1, 64	0.001
Peabody Picture Vocabulary Test	56.9	21.7	83.7	29.9	16.80	1, 64	0.001
Grammatic completion (Test of Language Development—2)	3.8	2.1	6.4	3.5	15.50	1, 64	0.001
Test of Auditory Comprehension of Language	70.6	9.9	84.5	16.1	18.40	1, 63	0.001
McCarthy oral vocabulary	3.6	4.6	11.4	5.1	39.50	1, 63	0.001
Visual-spatial and motor measures							
Stanford-Binet pattern analysis	42.6	13.2	46.8	9.6	1.54	1, 62	0.22
Beery Developmental Test of Visual-Motor Integration	5.4	3.0	6.1	3.0	0.95	1, 64	0.33
WISC-R block design	8.9	5.9	10.0	4.1	0.51	1, 62	0.48
Annett Pegs							
Right hand	19.3	4.4	16.6	2.6	6.63	1, 63	0.02
Left hand	20.7	4.2	18.8	4.7	2.60	1, 63	0.12

[a] Higher scores indicate better functioning on all measures except the Autism Behavior Checklist symptom subscale and the Annett Pegs.

assumptions of homogeneity of variances and of slopes were satisfied for the covariate analyses.

In an exploratory analysis, we also wanted to see whether the findings could be understood in terms of the timing of the onset of fluent language. The largest difference between the groups at the beginning of the study was in language fluency as measured by the McCarthy oral vocabulary test (21). The mean score on this measure at study enrollment among the children with Asperger's syndrome was 5.2; the 95% confidence interval (CI) was 3.1–7.3. The autistic children were then stratified into fluent and nonfluent groups depending on whether they had reached the same level of language fluency at follow-up as the children with Asperger's syndrome had shown at the beginning of the study (a score greater than or equal to 3, the lower limit of the 95% CI). Our hypothesis was that the scores of the fluent autistic group at follow-up and of the children with Asperger's syndrome at the beginning of the study would be similar to each other but different from those of the remaining (nonfluent) autistic children. This model was tested with a simple one-way ANOVA with post hoc Tukey's honestly significant difference tests. Raw scores from the Vineland Adaptive Behavior Scales, rather than scores standardized for age, were used in this analysis for more accurate comparisons over time.

To test for differences in the secondary outcome variables, a multivariate analysis of variance (MANOVA) was run on the language tests, the Vineland communication scale, and the visual-spatial measures at follow-up. If there were significant differences between the groups at follow-up, the baseline value for that variable was included in the ANOVA (because the Peabody Picture Vocabulary Test and the Test of Auditory Comprehension of Language were administered only at follow-up, the verbal reasoning score from the Stanford-Binet was entered as the baseline covariate). The analyses were based on all available data.

Results

Comparisons on the Primary Outcome Measures

Table 2 presents the results of the follow-up evaluations. The groups of children with autism and Asperger's syndrome were compared first on the socialization score from the Vineland Adaptive Behavior Scales. Both groups had low scores on the socialization measure, but the children with Asperger's syndrome had higher scores than the children with autism (75.6 versus 61.8). In fact, the magnitude of the difference amounted to a full standard deviation, a highly significant difference (Table 2).

Two possible explanations could account for this difference at follow-up. The first is that the difference is simply a reflection of the variation observed in this measure at the beginning of the study. In fact, the correlations between baseline and follow-up scores on this measure were very high (r=0.68 for children with autism and r=0.82 for children with Asperger's syndrome, p<0.001 for both). Accordingly, the socialization score from the beginning of the study was first entered as a covariate; in this case the differences between the groups became nonsignificant (partial t=0.40, p=0.69), as was the interaction term (partial t=0.34, p=0.73). This suggests that the difference at follow-up was a function of the difference at baseline (partial t=6.22, p<0.001), indicating parallel trajectories on this measure.

The second possibility is that the difference at follow-up is entirely explained by the initial variation in nonverbal IQ and/or language competence between the groups. A second multiple regression model was built predicting socialization scores at follow-up from baseline scores on nonverbal IQ, oral vocabulary, and verbal reasoning as well as the clinical diagnosis. In this model, the clinical distinction between Asperger's syndrome and autism was still significant in predicting follow-up socialization scores (partial t=2.08, p=0.04). The only covariate significantly associated with outcome was the Leiter nonverbal IQ (partial t=2.67, p=0.01). In other words, the clinical distinction had predictive ability when initial differences between the groups in language and nonverbal IQ were controlled for.

TABLE 3. Scores for Autistic Symptoms and Socialization of Children 6–8 Years Old With Autism Who Were or Were Not Verbally Fluent at 2-Year Follow-Up and Scores of Children 4–6 Years Old With Asperger's Syndrome at Study Enrollment

| | Score of Children With Autism at 2-Year Follow-Up | | | | Score of Children With Asperger's Syndrome at Study Enrollment (N=21) | |
| | Nonfluent (N=26) | | Fluent (N=18) | | | |
Measure	Mean	95% CI	Mean	95% CI	Mean	95% CI
Autism Behavior Checklist[a]	66.0	56.8–75.1	45.8	33.6–58.0	49.7	38.5–61.0
Raw score from Vineland Adaptive Behavior Scales socialization scale[b]	48.8	43.0–54.5	69.8	62.7–76.9	64.3	58.2–70.5

[a] At follow-up, fluent children with autism had significantly lower scores than nonfluent children with autism (F=4.61, df=2, 62, p=0.02).
[b] Fluent children with autism at 2 years and children with Asperger's syndrome at enrollment had comparable scores, which were significantly higher than those of nonfluent children with autism at 2 years (F=13.55, df=2, 61, p<0.001).

Children with Asperger's syndrome also demonstrated fewer total autistic symptoms at follow-up on the Autism Behavior Checklist than the children with autism. The magnitude of the difference again amounted to a full standard deviation, indicating a large effect size and a highly significant difference (Table 2). Correlations between baseline and follow-up scores on this variable were very high for both groups (r=0.64 for children with autism and r=0.74 for children with Asperger's syndrome, p<0.001 for both). Again, the difference at follow-up between the children with Asperger's syndrome and the children with autism was not significant when initial differences on this variable at the beginning of the study were controlled for (partial t=0.84, p=0.40), and the interaction term was also not significant (partial t=0.04, p=0.97). This suggests that the difference at follow-up was a reflection of the variation at baseline (partial t=5.79, p<0.001), and that the groups were following parallel developmental trajectories on this measure as well.

In the multiple regression model, the clinical distinction of Asperger's syndrome from autism was predictive of outcome scores on the Autism Behavior Checklist (partial t= 2.21, p=0.03), even when baseline nonverbal IQs, verbal reasoning scores, and oral vocabulary scores were controlled for; none of the covariates was associated with the outcome.

Comparisons on Secondary Outcome Measures

On the communication scale and the language tests (Table 2, the children with Asperger's syndrome usually obtained scores within a standard deviation of the population mean, whereas the children with autism usually obtained scores at least two standard deviations below normal. The overall MANOVA was highly significant (Hotelling's t=6.53, df=5, 56, p<0.001); however, the statistical differences disappeared (except for the oral vocabulary test) when the baseline values entered the model, indicating that the follow-up scores were better predicted by initial verbal abilities than by the clinical distinction (Asperger's syndrome versus autism).

When the groups were compared on the measures of visual-motor and visual-spatial abilities, no significant differences emerged even without controlling for measures taken at study enrollment (except for one subtest of the Annett Pegs).

Developmental Trajectories of Fluent and Nonfluent Autistic Children and Children With Asperger's Syndrome

The children with autism who reached the same level of language fluency at follow-up as shown by the children with Asperger's syndrome at study enrollment (i.e., a McCarthy oral vocabulary score of 3 or more) were considered "fluent." Children with autism who did not reach this level of language ability were considered "nonfluent." Table 3 shows the mean values and the 95% CIs of the fluent and nonfluent children with autism at follow-up and of the children with Asperger's syndrome at enrollment. One-way ANOVA on the Autism Behavior Checklist scores showed that the three groups were significantly different (Table 3). Post hoc tests demonstrated that at follow-up, the fluent autistic group was different from the nonfluent group (p=0.02) but not different from the group with Asperger's syndrome at baseline (p=0.87).

Similar results were obtained when we used raw scores from the Vineland Adaptive Behavior socialization scale. The overall ANOVA was significant (Table 3). The fluent group had higher scores than the nonfluent group (p<0.001) but was no different from the group with Asperger's syndrome (p=0.50). In other words, some fluent children with autism may have joined the pathway of the children with Asperger's syndrome and left the nonfluent children with autism behind. The difference is that the fluent group now resembled the group with Asperger's syndrome at the beginning of the study; i.e., they were on the same trajectory but about 2 years behind the children with Asperger's syndrome.

Discussion

Our study was originally designed to overcome several methodologic problems in the outcome literature on autism and provide new data on the outcome of children with Asperger's syndrome. We studied children with pervasive developmental disorder in a narrow age range soon after a reliable diagnosis can be made (i.e., at 4 years). As much as possible, this allows us to have an inception cohort of children at the same early stage of their disorder (27). It is virtually impossible to date age at onset reliably, but we can be sure that each child was enrolled in the

study within 1 or 2 years of definite diagnosis of pervasive developmental disorder. So far, we have had little attrition in our study group at follow-up (two children). In addition, a structured interview was used to make a diagnosis of autism and Asperger's syndrome, outcome assessments were conducted blind to predictor variables, there was a clear distinction between primary and secondary outcome measures, and a wide range of outcomes was assessed, including some measures that could be administered at both the beginning of the study and at 2-year follow-up.

The results reported here indicate, for the first time to our knowledge, that children with Asperger's syndrome and children with autism identified at 4–6 years of age demonstrate differences in social competence and autistic symptoms 2 years later. These differences are not clinically trivial and are evident even after initial differences in nonverbal IQ, expressive language, and verbal reasoning are controlled for. If the clinical distinction of Asperger's syndrome from autism were of dubious value, better predicted by initial variation in IQ and language abilities, or subject to large amounts of measurement error, the differences between the groups might attenuate over time and the groups might become more similar. Neither are the results circular, since the outcome measures were not used in the original distinction between the groups.

It is true, however, that the outcome differences can be explained by initial differences on Vineland Adaptive Behavior Scales socialization scores and the composite measure of autistic symptoms. In large part, the differences in outcome were already apparent at the beginning of the study, and the correlations between baseline and follow-up scores were very high and similar for both groups. This suggests that both subtypes of pervasive developmental disorder are following similar developmental pathways but that the autistic children are simply behind the children with Asperger's syndrome; that is, the pathways are parallel.

A closer look at the data, however, suggests that the variation in outcome seen in autistic children and those with Asperger's syndrome may be better explained by the observed changes in the level of language fluency, as measured by the oral vocabulary test, than by initial levels of this or other variables. There were large differences between the groups with Asperger's syndrome and autism on oral vocabulary at both the beginning of the study and at follow-up. Once children with autism develop a certain level of language fluency, they resemble children with Asperger's syndrome but at an earlier stage of development. This was demonstrated by our finding that autistic children who were fluent at follow-up were not statistically different from the children with Asperger's syndrome at the beginning of the study.

The present findings suggest that the differences between Asperger's syndrome and autism may be largely a matter of timing. The groups seem to be on different but parallel developmental trajectories initially, but there is the possibility that some children with autism may join the trajectory of children with Asperger's syndrome, once they develop a certain level of fluent language. The important implication of this result is that the clinical pathway (or subtype of pervasive developmental disorder) that a child follows largely depends on if, and when, the child develops fluent language. This model must be seen as exploratory at this point. The limitation of the analysis is that data are available at only two points in time; the outcome assessments were taken at the same time as the classification on the independent variable (language fluency). The proposed model, if replicated with other longitudinal assessments, may provide a more plausible explanation of the data than trying to determine whether Asperger's syndrome is on the spectrum of autism or is a different disorder altogether.

The diagnoses of both autism and Asperger's syndrome used in this study are based on the draft ICD-10 criteria as operationalized in the Autism Diagnostic Interview (1988 version). Insofar as ICD-10 was the basis for DSM-IV, we would argue that the criteria for Asperger's syndrome are consistent with the spirit, if not the letter, of DSM-IV. Our approach differs from DSM-IV in two ways. First, our earlier work (12) and a review of the literature (3) suggested the importance of delayed and deviant language development (i.e., delayed echolalia) in distinguishing Asperger's syndrome from autism. The ICD-10 and DSM-IV criteria do not mention deviant language development as a distinguishing feature. This is not to say that the children with Asperger's syndrome did not show any mild language delays or some transient deviance in language. Some certainly did, but not enough to merit a coding of 2 or 3 on the Autism Diagnostic Interview. Moreover, many children with Asperger's syndrome obtained below-normal scores on the standardized measures of receptive language. Similar findings have been reported by Eisenmajer et al. (28). Close inspection of the data, however, suggests that the feature of deviant language adds little to the more important criterion of the timing of fluent and spontaneous speech.

Second, we reversed the hierarchy rule, thus allowing a diagnosis of Asperger's syndrome to take precedence over a diagnosis of autism. The difficulty is that DSM-IV requires that children with Asperger's disorder meet the same threshold level of social impairment and repetitive activity as is necessary for autism but must not also meet criteria for autism. To assemble an inception cohort of children with pervasive developmental disorder, we had to identify preschool children with Asperger's syndrome. It turns out that many of these children will meet DSM-IV criteria for autism, as we and others (29) have found. This baseline study group might be considered a group of children with relatively more severe Asperger's syndrome than those ascertained at a later age.

Five other studies (28–32) have now reported that the DSM-IV criteria for Asperger's disorder are essentially un-

workable, largely because the children with a clinical diagnosis of Asperger's disorder also meet DSM-IV criteria for autism. If our more inclusive approach to the diagnosis of Asperger's syndrome were a serious problem, its effect would be to attenuate the differences between the groups at follow-up. Thus, if anything, the diagnostic bias in our study is conservative. However, large effect sizes were observed on both primary outcome measures.

To some extent, it may be more appropriate to see this as a study that compares the outcome of children with high-functioning pervasive developmental disorder with and without spontaneous and fluent language at 3 years of age and not to become distracted about names. We would argue, however, that as defined in DSM-IV, Asperger's disorder is so rare (one out of 68 children with high-functioning pervasive developmental disorder according to the Autism Diagnostic Interview algorithm) as to be not very useful clinically.

The classification of children with autism has a long a varied history (33). Perhaps the most useful distinction is between autistic children above and below the IQ cutoff for mental retardation. The data reported in this study indicate that another useful clinical distinction may be between children with high-functioning pervasive developmental disorder with different levels of language development. This study was not intended to make the claim that this distinction has etiological significance or that the diagnostic validity of Asperger's syndrome has been established. No doubt this will have to await a clearer understanding of the genetics of pervasive developmental disorder and the mechanism for the variation in language deficits. Rather, the results reported here indicate that the clinical distinction between autism and Asperger's syndrome does have prognostic importance, although it may be more useful to think of the pervasive developmental disorder subtypes in terms of different developmental trajectories than to argue that they represent different disorders or are on a continuum of severity.

References

1. Hertzig M, Snow ME, New E, Shapiro T: DSM-III and DSM-III-R: diagnosis of autism and PDD in nursery school children. J Am Acad Child Adolesc Psychiatry 1989; 29:123–136

2. Waterhouse L, Wing L, Fein D: Re-evaluating the syndrome of autism in the light of empirical research, in Autism: Nature, Diagnosis and Treatment. Edited by Dawson G. New York, Guilford Press, 1989, pp 203–281

3. Szatmari P: The validity of autistic spectrum disorders: a literature review. J Autism Dev Disord 1992; 22:583–600

4. Wolff S, Chick J: Schizoid personality in childhood: a controlled follow-up study. Psychol Med 1980; 10:85–100

5. Sparrow SS, Rescorla LA, Provence S, Cordonson SO, Goudreau D, Cicchetti DV: Follow-up of atypical children. J Am Acad Child Adolesc Psychiatry 1986; 25:180–185

6. Demb HB, Weintraub AG: A five-year follow-up of preschool children diagnosed as having an atypical pervasive developmental disorder. J Dev Behav Pediatr 1989; 10:292–298

7. Brown JL: Long-term follow-up of 100 "atypical" children of normal intelligence, in Autism. Edited by Rutter M, Schopler E. New York, Plenum, 1978, pp 463–474

8. Gillberg C, Steffenburg S: Outcome and prognostic factors in infantile autism and similar conditions: a population-based study of 46 cases followed through puberty. J Autism Dev Disord 1987; 17:273–288

9. Szatmari P, Archer L, Fisman S, Streiner DL: Asperger syndrome and autism: differences in behavior, cognition and adaptive functioning. J Am Acad Child Adolesc Psychiatry 1995; 34:1662–1671

10. Klin A, Volkmar FR, Sparrow SS, Cicchetti DV, Rourke BP: Validity and neuropsychological characterization of Asperger syndrome. J Child Psychol Psychiatry 1995; 36:1127–1140

11. Le Couteur A, Rutter M, Lord C, Rios P, Robertson S, Holdgrater M, McLennan J: Autism Diagnostic Interview: a standardized investigator-based instrument. J Autism Dev Disord 1989; 19:363–387

12. Szatmari P, Bartolucci G, Bremner R: Asperger's syndrome and autism: comparisons on early history and outcome. Dev Med Child Neurol 1989; 31:709–720

13. Szatmari P, Tuff L, Finlayson MAJ: Asperger's syndrome and autism: neurocognitive aspects. J Am Acad Child Adolesc Psychiatry 1990; 29:130–136

14. Krug DA, Arik J, Almond P: Behavior checklist for identifying severely handicapped individuals with high levels of autistic behavior. J Child Psychol Psychiatry 1980; 21:221–229

15. Volkmar F, Cicchetti DV, Dykens E, Sparrow SS, Leckman JF, Cohen DJ: An evaluation of the Autism Behavior Checklist. J Autism Dev Disord 1988; 18:81–98

16. Sparrow S, Balla D, Cicchetti D: Vineland Adaptive Behavior Scales (Survey Form). Circle Pines, Minn, American Guidance Service, 1984

17. Volkmar FR, Sparrow SS, Goudreau D, Cicchetti D, Paul R, Cohen DJ: Social deficits in autism: an operational approach using the Vineland Adaptive Behavior Scales. J Am Acad Child Psychiatry 1987; 26:156–161

18. Levine MN: Leiter International Performance Scale: A Handbook. Montreal, Institute of Psychological Research, 1986

19. Thorndike RL, Hagen E, Sattler J: Stanford-Binet Intelligence Scale, 4th ed. Itasca, Ill, Riverside, 1985

20. Newcomer PL, Hammill DD: Test of Language Development—2. Toronto, Psycan, 1988

21. McCarthy D: Manual for the McCarthy Scales of Children's Ability. New York, Psychological Corp, 1972

22. Dunn L, Dunn L: Peabody Picture Vocabulary Test—Revised Manual for Forms L and M. Circle Pines, Minn, American Guidance Service, 1981

23. Carrow-Woolfolk E: Test of Auditory Comprehension of Language (TACL-R). Allen, Tex, DLM Teaching Resources, 1985

24. Beery KE: Developmental Test of Visual-Motor Integration. Chicago, Follett, 1987

25. Annett M: The distribution of manual asymmetry. Br J Psychol 1972; 63:343–358

26. Wechsler D: Manual for the Wechsler Intelligence Scale for Children—Revised. New York, Psychological Corp, 1974

27. Sackett DL, Haynes RB, Tugwell P: Clinical Epidemiology: A Basic Science for Medicine. Toronto, Little, Brown, 1985

28. Eisenmajer R, Prior M, Leekam S, Wing L, Gould J, Welham M, Ong B: Comparison of clinical symptoms in autism and Asperger's disorder. J Am Acad Child Adolesc Psychiatry 1996; 35:1523–1531

29. Miller JN, Ozonoff S: Did Asperger's cases have Asperger disorder? a research note. J Child Psychol Psychiatry 1997; 38:247–251

30. Ghaziuddin M, Tsai LY, Ghaziuddin N: A comparison of the diagnostic criteria for Asperger syndrome. J Autism Dev Disord 1992; 22:643–649

31. Leekam S, Libby S, Wing L, Gould J, Gillberg C: Comparison of ICD-10 and Gillberg's criteria for Asperger syndrome. Autism 2000; 4:11–28

32. Fecteau S, Mottron L, Berthiaume C: Using the ADI-R to distinguish autistic disorder and Asperger's syndrome, in Scientific Proceedings of the 1999 Annual Meeting of the American Academy of Child and Adolescent Psychiatry. Washington, DC, AACAP, 1999, p 129

33. Cohen D, Volkmar FR, Paul R: Issues in the classification of pervasive and other developmental disorders. J Am Acad Child Psychiatry 1986; 25:158–161