Development and Psychopathology **21** (2009), 127–138
Copyright © 2009 Cambridge University Press
Printed in the United States of America
doi:10.1017/S095457940900008X

REGULAR ARTICLE

Developmental course of autistic social impairment in males

JOHN N. CONSTANTINO, ANNA M. ABBACCHI, PATRICIA D. LAVESSER, HANNAH REED, LEAH GIVENS, LILY CHIANG, TEDDI GRAY, MAGGIE GROSS, YI ZHANG, AND RICHARD D. TODD
Washington University School of Medicine

Abstract

Recent research has suggested that autistic social impairment (ASI) is continuously distributed in nature and that subtle autistic-like social impairments aggregate in the family members of children with pervasive developmental disorders (PDDs). This study examined the longitudinal course of quantitatively characterized ASI in 3- to 18-year-old boys with and without PDD. We obtained assessments of 95 epidemiologically ascertained male–male twin pairs and a clinical sample of 95 affected children using the Social Responsiveness Scale (SRS), at two time points, spaced 1–5 years apart. Longitudinal course was examined as a function of age, familial loading for PDD, and autistic severity at baseline. Interindividual variation in SRS scores was highly preserved over time, with test–retest correlation of 0.90 for the entire sample. SRS scores exhibited modest general improvement over the study period; individual trajectories varied as a function of severity at baseline and were highly familial. Quantitative measurements of ASI reflect heritable traitlike characteristics. Such measurements can serve as reliable indices of phenotypic severity for genetic and neurobiologic studies, and have potential utility for ascertaining incremental response to intervention.

The pervasive developmental disorders (PDDs) are a group of inherited conditions, primarily characterized by specific impairments in reciprocal social behavior, the most common of which are autistic disorder, Asperger disorder, and PDD—not otherwise specified (PDD-NOS). Research examining the longitudinal course of PDD has historically focused on whether or not case status (categorically defined as presence vs. absence) of each condition is preserved over time. Such studies have involved use of the 15-item Childhood Autism Rating Scale (CARS; Eaves & Ho, 1996; Mesibov, Schopler, Schaffer, & Michal, 1989), parental reports of *Diagnostic and Statistical Manual of Mental Disorders—Fourth Edition (DSM-IV;* American Psychiatric Association, 1994) criteria (Piven, Harper, Palmers, & Arndt, 1996), and the Autism Diagnostic Interview—Revised (ADI-R; McGovern & Sigman, 2005; Szatmari, Bryson, Boyle, Streiner, & Duku, 2003), each principally designed to establish categorical diagnoses. These studies have uniformly revealed a high degree of stability of diagnosis (ranging from 90% retention of diagnosis among higher functioning PDD subjects to 100% among low-functioning PDD

This work was supported by Grant HD042541 from the National Institute of Child Health and Human Development (to J.N.C.). We gratefully acknowledge the parents and families participating in the Washington University Social Developmental Studies Program and the Missouri Family Registry for their ongoing dedication to scientific research. This work is dedicated to the memory and continuing legacy of the late Richard D. Todd, PhD, MD, a pioneer in child psychiatric genetics and this manuscript's senior author (for whom it is published posthumously).

Address correspondence and reprint requests to: John N. Constantino, Washington University School of Medicine, 660 South Euclid Avenue, Campus Box 8134, St. Louis, MO 63110; E-mail: constantino@wustl.edu.

https://doi.org/10.1017/S095457940900008X Published online by Cambridge University Press

subjects), as well as evidence for subtle improvement with respect to the *DSM-IV* symptom counts over the course of childhood. The sample sizes of these studies, however, have not been adequate to establish norms by which to predict change in symptom counts over the course of development for children with various PDDs.

Aside from what can be inferred from symptom counts, very little is known about how the *severity* of autistic social impairment (ASI) changes over the course of development; until recently, valid quantitative measures of severity have not been available to researchers or clinicians. Quantitative indices of severity have become especially important in both clinical and research settings, as it has become increasingly apparent that clinical PDDs represent the extreme end of a continuous (quantitative) distribution of social impairments that occur in nature (Constantino & Todd, 2003; Ronald, Happe, Price, Baron-Cohen, & Plomin, 2006). The implications of such a distribution are (a) that subtle changes in an individual's level of symptomatology might occur and might affect his or her level of social adaptation and (b) that it may be arbitrary where to place cutoffs that distinguish "affected" from "unaffected" individuals. From a genetic standpoint, ASI whose severity falls below the (arbitrary) threshold for a clinical diagnosis has been found to aggregate in the unaffected family members of many children with autism (Constantino et al., 2006; Dawson et al., 2005; Pickles et al., 2000; Piven, Palmer, Jacohbi, Childress, & Arndt, 1997), and it is therefore possible that the cause of such impairment is related to the causes of autism itself. Furthermore, it has been observed that subclinical autistic impairments may operate to compound and worsen co-occurring psychiatric conditions (Constantino, Przbeck, Friesen, & Todd, 2000).

Given these implications of the quantitative structure of ASI, it is important to establish whether *measurements* of such quantitative traits are actually stable over time in the same way that categorical PDD diagnoses are stable over time. Understanding of the natural course of severity over time is critical for appropriately interpreting outcomes of interventions for PDD (McDougle et al., 2005; Pine, Luby, Abbacchi, & Constantino, 2006), which rarely, if ever, result in complete remission of disease status. Identifying congruence in the longitudinal course of ASI between fully affected and subclinically affected children would support a common underlying biological structure for both groups, and would justify the use of larger samples (representing a broader range of severity) to examine associations between behavior and underlying biological variables (genotype, neurobiological markers, etc.). As an example of this, we have recently demonstrated that linkage signals for autism susceptibility genes (Duvall et al., 2007) can be enhanced by utilizing quantitative trait data from *all* siblings in autism spectrum disorder (ASD)-affected families, in comparison to traditional methods that restrict such analyses to the fully affected children only.

In this study, we examined the first wave of data from an ongoing longitudinal study of ASI in both clinically ascertained and general population sibling pairs, which allowed exploration of whether longitudinal trajectories for subthreshold ASIs matched those for clinical level symptoms when equivalent measurement methods were used. The study encompassed the age range from 3 to 18 years and the full range of severity of ASI that occurs in nature (see Constantino & Todd, 2003).

To our knowledge, this is the first prospective study of ASI to incorporate validated *quantitative* measurement methods. We hypothesized that ASI would exhibit a high degree of stability over time across the entire range in which it manifests in nature.

Methods and Materials

Sample

Subjects for this study are participants in an ongoing longitudinal study of ASI. This report concerns the first two subject groups for whom systematic prospective data has been collected.

The first group of subjects were twins who were recruited for enrollment into this study from a sample of 232 male pairs age 8 to 15, who had originally been epidemiologically ascertained from the general population for the Missouri Twin Study (children who were nonverbal or who carried diagnoses of autistic

https://doi.org/10.1017/S095457940900008X Published online by Cambridge University Press

disorder or mental retardation were excluded from this sample), and assessed in 1999 using the Social Responsiveness Scale (SRS; for a more detailed description of the sample, see Constantino & Todd, 2000). It should be stated that there are numerous research groups following the twins in this sample. Research access to the twins is governed by an oversight committee whose charge is to minimize the risk of the families being overburdened by multiple research requests. Research contacts are allowed to occur no more frequently than every 6 months; thus, there are constraints on access to various sets of twins at various points in time.

Over the 2004–2005 year, we were able to contact the families of 123 of the original 232 male twin pairs to collect 5- to 6-year follow-up data. Of those, 28 declined participation, and 95 (40 identical twin pairs, 55 nonidentical twin pairs) enrolled in the current study. There were no statistically significant differences in mean age or SRS score between enrollees, those whose families declined participation, and the remainder of the previously assessed sample.

In addition to obtaining follow-up SRS assessments, we conducted additional laboratory assessments on available twins who had scored in the top 10% of the distribution (at the time of the original assessment) to verify that elevated scores in a general population "screening" were consistent with appreciable autistic symptomatology. PDDs are believed to affect up to 1% of males in the general population (Fombonne, 2005); however, children who carried formal diagnoses of autism had been excluded from the Missouri Twin Study Sample. The laboratory assessments were completed according to the Autism Diagnostic Observation Schedule (ADOS), an established semistructured diagnostic measure for autism (Lord, Rutter, Dilavore, & Risi, 1999) that measures social impairment, communicative impairment, and sterotypic behaviors that are specific to PDD. Of seven high-scoring SRS twins who participated in this observational assessment, all had appreciable deficits in at least one domain on the ADOS, four exceeded the cutoff score of 4 for deficits in reciprocal social interaction toward a clinical PDD diagnosis; the mean for the group was 4.3 ± 2.7.

The second subject group comprised 85 boys with PDDs consecutively recruited by their physicians (in 2003–2005) from either (a) the Washington University Child and Adolescent clinics or (b) from outpatient child psychiatry practices in the greater St. Louis metropolitan area. Any child with a PDD diagnosis documented by a child psychiatrist was eligible for inclusion in the study; for the Washington University subgroup, an additional inclusion criteria was that index PDD subjects had at least one male full sibling (whether coaffected or unaffected by PDD). Ten of the male siblings were coaffected with a PDD. Families were excluded from the study if the index PDD subject carried a diagnosis of a comorbid psychiatric disorder or if there was any sustained ambiguity with respect to a singular primary diagnosis. All index cases, affected sibs, and any undiagnosed male sibs were assessed using the SRS (see below) by parent and teacher report at baseline. All clinically affected subjects (85 index PDD cases, 10 affected male sibs) were subsequently reassessed with the SRS at 1-year follow-up. Table 1 summarizes the assessment schedule and selected sample characteristics, as a function of specific groupings of subjects.

For diagnostic confirmation all clinically affected subjects were assessed with the ADI-R and the ADOS. Of the total 95 PDD subjects, 71 scored at or above the clinical cutoff for a full *DSM-IV* diagnosis of autistic disorder on the ADI-R, the ADOS, or both; the remainder had substantially elevated scores on these measures and were diagnosed clinically with either Asperger disorder or PDD-NOS by their respective clinicians. There was no change in clinical diagnosis for any of the participants over the course of the 1-year follow-up period. Current IQ scores were available for 38 of the subjects; mean full-scale IQ for this group was 93.2 ($SD = 24.7$).

Measures

SRS. The SRS is a 65-item quantitative measure of ASI, which capitalizes on observations of children in naturalistic social contexts by either parent or teacher report. The instrument uses a 4-point Likert scale (*not true, sometimes true, often true, almost always true*) for each item.

https://doi.org/10.1017/S095457940900008X Published online by Cambridge University Press

Table 1. *Schedule of study assessments and selected characteristics of study groups*

Study Group	PDD Subjects	Undiagnosed Male Sibs	General Population Twins
n	95	66	190
Baseline assessment	SRS-p, SRS-t, ADI-R, ADOS	SRS-p, SRS-t	SRS-p
Age (*SD*) at baseline	8.0 ± 4.0	7.0 ± 3.4	11.6 ± 1.8
SRS-p mean at baseline	102.8 ± 26.2	35.0 ± 31.3	38.1 ± 24.8
Follow-up assessment	SRS-p, SRS-t, DIGS/FIGS[a]	—	SRS-p, ADOS[b]
Time of follow-up assessment	1 year	—	5 years

Note: SRS-p, Parent-Report Social Responsiveness Scale; SRS-t, Teacher-Report Social Responsiveness Scale; ADI-R, Autism Diagnostic Interview—Revised (McGovern & Sigman, 2005); ADOS, Autism Diagnostic Observation Schedule (Lord et al., 1999); DIGS, Diagnostic Interview for Genetic Studies (Nurnberger et al., 1994); FIGS, Family History Interview for Genetic Studies (Maxwell, 1992).
[a]Sixteen high functioning verbal adolescents.
[b]Subjects who scored in top 10% of distribution at baseline.

SRS items cover each of the three *DSM-IV* criterion domains (social, language, repetitive, stereotypic behaviors/restricted range of interest) for autism; for the social domain, separate items ascertain the extent to which a child is aware of social cues in his/her environment (social awareness), appropriately interprets those cues (social cognition), is capable of a reciprocal communicative response (social communication), and is *motivated* to engage socially. Scores on the SRS are highly heritable (Constantino & Todd, 2000, 2003, 2005), continuously distributed in the general population (Constantino & Todd, 2003), exhibit a unitary factor structure (Constantino et al., 2004), and distinguish children with autism spectrum conditions from those with other child psychiatric conditions (Constantino, Przybeck, Friesen, & Todd, 2000; Constantino & Gruber, 2005). Both the 3-year-old version of the SRS (Pine et al., 2006) and the 4- to 18-year-old version (Constantino & Gruber, 2005) were implemented in this study; the number of items and range of scores are identical for the two instruments; wording of the items differ only where developmentally appropriate.

The SRS was obtained exclusively by maternal report for the twins, and by both maternal report and teacher report for the clinically ascertained sample of sibling pairs (all at baseline and clinically affected children at 1-year follow-up). For teacher reports, parents were asked to request that an SRS form be completed by a current classroom teacher who had known the child for a minimum of 2 months and whom they felt knew their child best. The SRS generates a singular total score for ASI empirically validated via factor, cluster, and latent class analysis (Constantino et al., 2004), as well as treatment scale scores, which (although not empirically derived) represent relevant target domains for intervention. The SRS exhibits nonsignificant correlations with IQ (Constantino et al., 2006), substantial agreement with the ADI-R (Constantino et al., 2007), and an absence of age effects in the range of ages represented by this study (Constantino & Gruber 2005).

For the analyses in this study, raw (nontransformed) scores on the SRS were used; the scores range from 0 to 195, with higher scores indicating more severe levels of social impairment. For the clinically ascertained sibling sample, the correlation between parent- and teacher-report SRS score at baseline was robust with an intraclass correlation coefficient (ICC) of 0.66. The familiality of SRS reports in the clinical sample was documented by calculating the average of raters' ICC for sibling pairs (one pairing involving the index case and closest in age male sib for clinical subjects with one or more male sibs); the ICC was 0.25. This value was expected to be somewhat less than the sibling correlation of 0.35 that we have observed in the general population (Constantino & Todd, 2003), because the range of SRS scores encompassed by the clinically affected index cases

https://doi.org/10.1017/S095457940900008X Published online by Cambridge University Press

was restricted (within the pathological end of the distribution) in comparison to the range represented by the general population.

Intervention history. Extensive treatment records were available (at or near the time of enrollment) from the medical records of PDD subjects who had been recruited from the Washington University Child and Adolescent Psychiatry Service ($n = 40$). Data abstracted from the records included presence or absence of *history* of treatment with applied behavior analysis, presence or absence of current occupational therapy, and presence or absence of current pharmacotherapy (categorically subdivided into atypical neuroleptics, stimulants, selective serotonin reuptake inhibitors).

Familial loading for PDD. For each clinical subject, parents were interviewed according to a uniform protocol for inquiring about any known (diagnosed) or strongly suspected case of a PDD within the extended family. The pedigree was reviewed out to third-degree relatives; and the parent was asked whether a diagnosis of PDD, Asperger disorder, autism, ASD or schizoid personality disorder had ever been made for each family member. Next the parent was asked whether any undiagnosed relatives had constellations of behavior that he/she (the parent) recognized in retrospect as similar in nature to core PDD symptoms (for which the list of *DSM-IV* criteria for autism was reviewed with the informant); if such behaviors were viewed by the parent as significantly limiting that relative's functioning, that relative was designated as "suspected" of having a PDD. A reliability subsample of such suspected family members ($n = 35$) revealed that when the index subject's parent completed an SRS report on that family member, mean score was 81.8 ± 35.6. In this sample, when comparing PDD subjects who have at least one other family member affected to those with none, there was no significant difference in mean SRS score, $t\,(69) = 1.3, p = .20$.

Diagnostic Interview for Genetic Studies/Family History Interview for Genetic Studies (DIGS/FIGS). Although it is common for children to carry PDD diagnoses and known that those diagnoses are highly stable over time, it is unexplainably rare for *adults* to carry PDD diagnoses. Conversely, it is relatively common for adults to carry diagnoses of schizoid personality disorder and rare for children to carry that diagnosis, even though retrospective histories of individuals with schizoid personality disorder indicate substantial precursor symptomatology in childhood. Given the tremendous overlap in symptomatology between these two conditions, we were interested in the question of whether milder PDD syndromes in adolescence constitute a prodrome for schizoid personality disorder in adulthood. For this reason, adolescent PDD subjects who were willing and able to complete the assessment were asked to complete the schizoid personality disorder section of the DIGS (self-report; Nurnberger et al., 1994). Their mothers completed the analogous section of the FIGS (parent report; Maxwell, 1992), which is a complementary measure in which a parent or other adult informant reports on the subject, regarding the same items ascertained in the DIGS. Sixteen subjects completed the assessments (mean age $= 15.2 \pm 2.5$ years); one of these subjects carried a clinician diagnosis of high-functioning autistic disorder; the remainder carried clinical diagnoses of Asperger disorder or PDD-NOS.

Data analysis

Univariate analyses of the stability of SRS scores over time were conducted separately for the twin and clinical samples using (a) ICC between baseline and follow-up scores and (b) paired *t* tests.

To examine the genetic–environmental structure of change over time in the twin data, variance in change scores for first-born twins, variance for second-born twins, and within-pair covariance were calculated separately for the respective groups of monozygotic (MZ; $n = 40$ pairs) and dizygotic (DZ; $n = 55$ pairs) male twins. The resulting covariance matrices were fit to models of genetic and environmental causation using the structural equation modeling software Mx (Neale & Cardon, 1992). Univariate models followed the classic twin design, as adopted and described in our previous work (Constantino & Todd, 2003).

For the clinical sample, linear regression analysis was used to assess the effects of familial

https://doi.org/10.1017/S095457940900008X Published online by Cambridge University Press

loading (presence or absence of a second PDD-affected individual among first- to third-degree relatives), baseline severity, and age on the change in SRS scores. We separately examined effects of specific treatments (which have significant associations with age and baseline severity) on change in SRS scores. Principal components factor analysis was subsequently implemented to determine whether specific subsets of symptoms might exhibit independent patterns of change over time, among clinically affected subjects.

Finally, data common to the twin and clinical samples (maternal SRS scores at baseline and follow-up) were pooled and analyzed for within-subject contrasts using a generalized linear modeling repeated-measures analysis executed with SPSS statistical software (SPSS, Chicago).

Results

General population twins: Longitudinal analysis

We first describe the univariate analyses of the longitudinal course of SRS scores in the twin sample. Over time, interindividual differences were highly preserved, as depicted for maternal-report SRS scores in the scatter plot in Figure 1. The ICC between baseline and follow-up maternal SRS score was 0.71 (selecting one twin per family at random).

Despite a lack of cross-sectional age effects on SRS scores in the twin sample, there were modest improvements in mean scores over time, on the order of 0.5 *SD* over the course of the 5-year follow-up. When examining differences between baseline and follow-up for the twins, the most conservative approach is to include one twin per family ($n = 95$): for Twin 1 mean change = 9.7 points on the SRS, $t\ (94) = 5.15$, $p < .001$; for Twin 2 mean change = 13.0, $t\ (94) = 6.44$, $p < .001$.

General population twins: Genetic structure of time-rated change

We next explored the genetic and environmental structure of time-rated change in SRS scores in twins. The difference in SRS total score between baseline assessment and 5-year follow-up was calculated for each twin, covariance matrices were separately constructed for MZ and DZ twins, and the data were subjected to univariate structural equation modeling using the statistical software Mx (Graphic User Interface; Neale, 2004). Structural equation modeling incorporates variance/covariance data into models that consider causal influence on latent variables that underlie (but do not equate with) phenotypic measurements (for a description of this method of analysis, see Constantino & Todd, 2003; Neale & Cardon, 1992). The results revealed that the most parsimonious model for causal influence on change over time was a model that included additive genetic influences (A, explaining 73% of variance) and modest unique environmental influences (E, explaining 27% of variance; E subsumes measurement error) on time-rated change in the latent variable represented by SRS scores. Results of model fitting and parameter estimates derived from the AE model are summarized in Table 2.

It is important to note that any common environmental influence interacting with genetic factors to bring about a reduction in SRS scores would be subsumed under the parameter for additive genetic influences. The presence of such an interaction would be consistent with the finding that the most pronounced improvements occurred among children with the most impairment at baseline.

General population twins: Do separate genetic factors influence time-rated change in children versus adolescents?

Given the results of pronounced effects of additive genetic influence on change, as well as our previously published results documenting substantial heritability of cross sectional SRS ratings (Constantino & Todd, 2003), we proceeded with an additional analysis to explore whether separate sets of genetic influences might be operating during earlier versus later stages of development (within the same children). To do this, we attempted a set of bivariate analyses (the two variables being SRS score at baseline and SRS score at follow-up) involving children in the sample who were 8–12 years of age at baseline and greater than 13 years at follow-up ($n = 30$ MZ pairs and 41 DZ pairs from

https://doi.org/10.1017/S095457940900008X Published online by Cambridge University Press

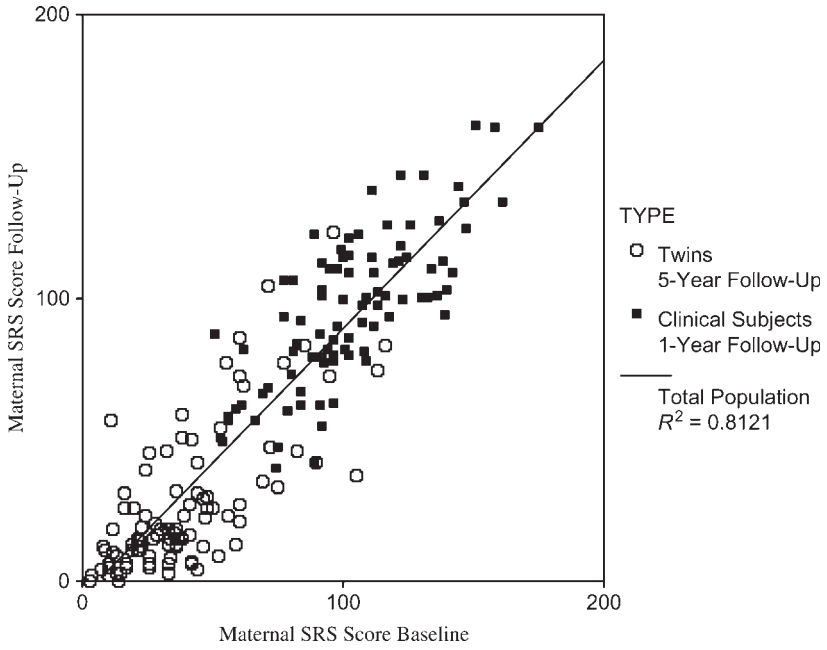


Figure 1. A scatter plot of the maternal SRS scores at baseline and follow-up. This plot incorporates both groups of study subjects (general population and clinic subjects) to represent the full range of SRS scores that occur in nature. When calculating the ICC for the whole sample, ICC = 0.90; when calculated separately for each study group, twins (one per family, as incorporated in the scatter plot) ICC = 0.71, and clinic subjects, ICC = 0.76. Lower coefficients of correlation are expected when the range of trait variation within a subsample is narrower.

Table 2. *Genetic and environmental influences on change over time*

	Max. Likelihood Analysis					Param. Estim. for Causal Influence		
Model	χ^2	*df*	*P*	AIC	RMSEA	a^2 (95% CI)	c^2	e^2 (95% CI)
ACE	6.08	3	0.07	0.83	0.13			
AE	7.06	4	0.13	−0.94	0.10	0.73 (0.64–0.79)	—	0.27 (0.20–0.37)
CE	12.7	4	0.01	4.70	0.20			

Note: a^2, additive genetic influence (also subsumes interactions between unmeasured environmental influences and genetic factors); c^2, common environmental influence; e^2, unique environmental influence (also subsumes measurement error); CI, confidence interval; AIC, Akaike information criterion; RMSEA, root mean square error approximation. For both of these indices, lower values represent improved fit to the data. Models incorporate effects of additive genetic influences (A), common or shared environmental influences (C), and unique or nonshared environmental influences (E).

the total of 95 male pairs in the sample). We compared models for complete overlap, partial overlap (Cholesky decomposition), and nonoverlap of genetic influences using bivariate (SRS baseline, SRS follow-up) models depicting each respective level of genetic overlap, in the manner described in Constantino, Hudziak, and Todd, 2003. These models incorporated parameters for additive genetic (A) and unique environmental influence (E) as substantiated by the results of univariate analysis of change scores (described above) and by previously published univariate analyses of baseline SRS scores in males (Constantino & Todd, 2000).

Although the power of the subsample is extremely limited for differentiating the three

https://doi.org/10.1017/S095457940900008X Published online by Cambridge University Press

bivariate models, the goodness of fit of the non-overlap model was much poorer (Akaike information criterion [AIC] = 95.6) than that of either the complete overlap model (AIC = 26.1) or the partial overlap model (AIC = 1.8, 95% confidence interval −10.2 to +21.4); lower value indicates superior fit. Parameter estimates derived from the partial overlap (best-fitting) model indicated that although a majority of additive genetic influences on ASI in childhood overlapped with those operating in adolescence, up to 33% (based on 95% confidence limits for parameter estimation) of the total additive genetic influence (0.80) were specific to developmental stage (childhood vs. adolescence).

Clinical sample: Longitudinal analysis

As was the case for twins in the general population, interindividual differences among PDD subjects were highly preserved over time by maternal SRS report (ICC = 0.76). There was substantial agreement between mothers and teachers on SRS score at baseline (ICC = 0.66) and follow-up (ICC = 0.63). When averaging SRS scores from parent and teacher report at baseline and follow-up, the coefficient of correlation between baseline and follow-up scores was 0.63.

There were modest improvements in mean scores over time, which reached statistical significance by maternal report (as observed in twins) but not by teacher report, as depicted in Table 3. In most cases, teacher reports at follow-up were provided by different teachers than those who provided baseline reports, which may, in part, explain the discrepancy in magnitude of trends for change over time, between parent- and teacher-report data. None of the children in the clinical group experienced a magnitude of reduction in maternal SRS scores over the 1-year period that would have been consummate with a "loss" of a PDD diagnosis.

Clinical sample: Predictors of change

Linear regression analysis, examining the effects of age, baseline SRS, and familial loading, revealed only an effect of severity at baseline on time-rated improvements in SRS scores, F (3.79) = 3.31, R^2 = .11; baseline SRS effect, t = −2.85, p = .006, β = 18.2. Higher level of severity at baseline was associated with treatment with atypical neuroleptic medication and occupational therapy, so it is possible that these contributed to the subsequent trend toward improvement (in more severely affected subjects in the clinical group); however, when the effects of the various treatment modalities were examined in relation to change over time, no statistically significant predictors emerged among the 40 subjects for whom data were available, F (7, 32) = 0.94.

Table 3. *Mean SRS scores by rater*

	PDD Subjects (n = 95)			
	Mother		Teacher	
	Mean	*SD*	Mean	*SD*
Baseline SRS	102.8	26.2	92.4	26.9
1-year follow-up	95.6	26.7	89.7	32.4
Paired t	3.87		0.73	
df	94		92	
p (two tailed)	0.0002		0.47	

Note: SRS, Social Responsiveness Scale (Constantino & Todd, 2000); PDD, pervasive development disorder.

To determine whether specific subsets of SRS items might exhibit independent patterns of correlated changes over time, principal components factor analysis was conducted on change scores (at the item level) and revealed no evidence of independent change in specific symptom clusters. This supported the validity of examining naturalistic change on the basis of a singular (total) SRS score.

Clinical sample: Overlap of autistic and schizoid symptomatology

We next explored the possible continuity between ASIs and schizoid personality disorder symptoms in higher functioning adolescent ASD subjects at 1-year follow-up. On average, parents reported 2.8 schizoid personality disorder symptoms in their ASD children, which was substantially higher than what was endorsed by self-report among the subjects themselves (4 symptoms are required for a *DSM-IV* diagnosis of schizoid personality disorder). The most

https://doi.org/10.1017/S095457940900008X Published online by Cambridge University Press

common symptoms endorsed by parents were "has no one to be really close to or confide in, or just one person, outside of the immediate family" and "acts cold or distant, hardly ever smiles or nods back at people." By parent report, 5 of 16 ASD subjects met full *DSM-IV* criteria for schizoid personality disorder; an additional 3 subjects met three of four criteria required for the diagnosis. Quantitative trait scores on the SRS did not significantly differ between those who met criteria for schizoid personality disorder and those who did not.

Longitudinal analysis of pooled clinical and epidemiologic data

Given the high level of congruence between the respective developmental trajectories of these traits in our clinically and epidemiologically ascertained samples, we pooled maternal SRS report data from the clinical and twin samples to derive the largest available sample for longitudinal data analysis. Test–retest reliability (ICC) for total SRS scores across the entire range of scores represented by the pooled sample was 0.90. We conducted a generalized linear modeling repeated-measures analysis for within-subject contrasts, which revealed that time-rated improvements as reported by parents, although modest in magnitude, achieved a high level of statistical significance ($p < .001$). The trend for improvement, although most pronounced in more severely affected children (see above regression analysis) spanned the entire range of the SRS distribution, and therefore was not entirely explainable as a straightforward function of regression toward the mean.

Discussion

Consistent with the results of previous longitudinal studies of autism, we observed a high degree of preservation of interindividual differences in ASI over time, but also subtle improvements over the course of the study period in both clinically affected and "unaffected" children. In this first wave analysis of our longitudinal study of *quantitative* ASI, interindividual variation, as reported by mothers using the SRS, exhibited a baseline to follow-up correlation exceeding 0.70 for both twins from the general population and for a clinically ascertained sample of boys with PDD. Congruence in the longitudinal course of autistic traits across the range of severity observed in nature extends findings from our previous family/genetic studies supporting a biological link between clinical and subthreshold levels of symptomatology (Constantino et al., 2006; Duvall et al., 2007). Furthermore, the level of stability in interindividual differences over time indicates that quantitative measurements of ASI can serve as extremely reliable markers of symptomatology, which might relate to neurobiological and genetic determinants of autism, and which might be useful for ascertaining the response of core symptoms to successful intervention.

The changes that we observed over time were very gradual and, on average, would only be construed as clinically significant when allowed to accumulate over years of time. There was a clear (statistically significant) tendency in both the clinical and general population samples for more severely affected children to exhibit greater reduction in impairment scores over time. Although the prediction of improvement by baseline severity is consistent with some influence of regression toward the mean on our results, baseline severity accounted for only 11% of the variance in change in our clinical sample, and the overarching trend for improvement extended to even the most socially competent group of children at baseline. Moreover, the nature of that improvement in the clinical group encompassed the entire constellation of symptoms observed in autism. As we have observed for the factor structure of autism itself (Constantino et al., 2004), the factor structure of improvement over time is consistent with a singular underlying component, best characterized quantitatively by a total impairment score.

If it is true (as our data suggest) that general improvements are occurring in the absence of age (developmental) effects on cross-sectional measurements of ASI, in essence that all children, irrespective of age, are getting better (albeit slowly), this would raise the possibility of subtle *period effects* on ASI. Period effects are influences on an entire population (in this

https://doi.org/10.1017/S095457940900008X Published online by Cambridge University Press

case, children across the entire range of ages represented by the sample) for a specific period of time (in this case, the years over which the study was conducted). The identification of true period effects (possible examples would be beneficial effects of new treatments across the age distribution, or recent improvements in the capacity of the educational system to assist children affected by a wide range of autistic traits) would have potentially important implications for public health strategies for improving the outcomes of children with PDDs.

One would expect that period effects, if present, would result from environmental (not genetic) influences on the course of ASI over time. Our analysis of the genetic and environmental structure of change over time among twins revealed strong evidence for the importance of genetic factors in determining the longitudinal course of symptoms and the interesting possibility that different sets of genetic factors might account for some of the heritable influences on ASI in childhood versus adolescence. The caveat, however, is that in the analysis of twin data, any interaction (or correlation) between genetic factors and *unmeasured* environmental factors is subsumed under the parameter for genetic influence. Thus, if there is a socioenvironmental change that is resulting in the observed improvements in symptomatology over time, it is likely interacting with genetic factors to exert such an influence on outcome. A possible candidate for such an interaction with the environment is inherited deficiency in social behavior itself; in this study, we directly observed that the children who experienced the most improvement over time were the ones who were most affected by ASI at baseline.

Finally, we observed that among higher functioning adolescents with PDD, a high proportion (50%) meet or approach *DSM-IV* diagnostic criteria for schizoid personality disorder. It is rare for nonretarded adults to carry PDD diagnoses, despite the fact that these are common childhood conditions with enduring symptomatology over time; this suggests a possible lack of systematic recognition of the impairments that are carried forward into adulthood by individuals who were diagnosed with PDD-NOS, Asperger disorder or "high-functioning autism" in childhood. One hypothesis supported by our exploratory accrual of data on schizoid personality disorder symptoms is that a sizeable portion of higher functioning PDD subjects become diagnosable with schizoid personality disorder in adulthood. Elucidating the continuity and discontinuity between these two conditions warrants further study and remains a goal of our ongoing longitudinal research. Because the subjects themselves had minimal insight into the presence of schizoid personality disorder symptoms (that their parents identified in them), it will be important for future research efforts to incorporate observations of parents or other informants, rather than relying exclusively on reports by the subjects themselves.

There are some limitations of this analysis of first-wave data from our ongoing longitudinal study; the sample size, although the largest to date for tracking the longitudinal course of quantitative ASI, warrants efforts at replication in still larger samples. In addition, our clinical subjects (for whom there has been minimal attrition in the sample to date) have been studied at only two time points thus far. Nevertheless, the convergence of findings across two disparate study groups, with respect to *both* stability of interindividual differences *and* change over time, supports the reliability of the findings in each respective sample. Reliance on maternal report SRS scores is potentially vulnerable to rater bias; however, the absence of age effects at either baseline or follow-up make it less likely that the findings regarding change are substantively influenced by systematic maternal reporting bias. We have previously reported an absence of evidence for maternal rating bias on SRS scores obtained in considerably larger samples of twins (Constantino & Todd, 2003). Continued follow-up of this sample, which will include yearly reassessments of the clinical subjects described herein, and every other year assessments of the clinical subjects' male siblings (see Constantino et al., 2006), by both parent and teacher report, will help elucidate any ongoing effects of reporting bias. An additional limitation was that current IQ data was available for only a minority of clinical subjects and none of the twins. Furthermore, this report is limited to an analysis of male

https://doi.org/10.1017/S095457940900008X Published online by Cambridge University Press

sibling pairs; parallel data collections involving female subjects are in progress.

In this study we showed that the average SRS score recorded by parent and teacher exhibits very high levels of consistency ($r = 0.63$) over time even when two different teachers provide ratings at respective baseline and follow-up assessments. Use of multiple-informant data from parent and teacher report on a given child may more comprehensively represent the social functioning of children across home and school settings than when relying on information from a single source (Constantino et al., 2007).

Conclusions

Quantitative measurements of ASI are both highly heritable and extremely stable over time; they reflect traitlike characteristics that can serve as reliable markers of core components of the autistic syndrome for genetic and biological studies, and for evaluation of the effects of intervention. Our observation of subtle improvement over time is consistent with findings of a number of previous research studies. In addition, these data indicate that naturalistic improvements over time exhibit a unitary factor structure (across all symptom domains in the autistic syndrome) and involve children whose social impairments fall above or below the threshold for a clinical diagnosis. The improvements appear most pronounced among children with more severe levels of ASI at baseline. It will be important for longitudinal studies, treatment studies, and genetic studies to continue to explore causal influences on change over time, as this will have important implications for educational and public health interventions for affected children. Such studies will be greatly enhanced by comprehensive coverage of the array and timing of *all* interventions received by affected children over the course of their development. The possibility that different sets of genetic factors might predominate in childhood versus adolescence in influencing ASI warrants further study and warrants consideration of segregating child and adolescent subjects in studies of gene expression in autism.

Further exploration of the continuity between a childhood diagnosis of PDD-NOS and an adult diagnosis of schizoid personality disorder requires additional study. This may lead to better understanding those patterns of social behavior in childhood that predict enduring conditions construed as personality disorder in our current nomenclature. It is always important to note that quantitative characteristics (traits) of any psychiatric condition that fall below the threshold for clinical diagnosis may for many individuals, and under many conditions, be adaptive. The continued study of such traits may lead to greater insights not only into disease processes but also into the biology of normal human social development.

References

American Psychiatric Association. (1994). *Diagnostic and Statistical Manual of Mental Disorders—Fourth Edition*. Washington, DC: Author.

Constantino, J. N., & Gruber, C. P. (2005). *The Social Responsiveness Scale manual*. Los Angeles: Western Psychological Services.

Constantino, J. N., Gruber, C. P., Davis, S., Hayes, S., Passanante, N., & Przybeck, T. (2004). The factor structure of autistic traits. *Journal of Child Psychology and Psychiatry*, *45*, 719–726.

Constantino, J. N., Hudziak, J. J., & Todd, R. D. (2003). Deficits in reciprocal social behavior in male twins: Evidence for a genetically independent domain of psychopathology. *Journal of the American Academy of Child & Adolescent Psychiatry*, *42*, 458–467.

Constantino, J. N., Lajonchere, C., Lutz, M., Gray, T., Abbacchi, A., McKenna, K., et al. (2006). Autistic social impairment in the siblings of children with pervasive developmental disorders. *American Journal of Psychiatry*, *163*, 294–296.

Constantino, J. N., LaVesser, P. D., Zhang, Y., Abbacchi, A. M., Gray, T., & Todd, R.D. (2007). Rapid quantitative assessment of autistic social impairment by classroom teachers. *Journal of the American Academy of Child & Adolescent Psychiatry*, *46*, 1668–1676.

Constantino, J. N., Przybeck, T., Friesen, D., & Todd, R. D. (2000). Reciprocal social behavior in children with and without pervasive developmental disorders. *Developmental and Behavioral Pediatrics*, *21*, 2–11.

Constantino, J. N., & Todd, R. D. (2000). Genetic structure of reciprocal social behavior. *American Journal of Psychiatry*, *157*, 2043–2045.

Constantino, J. N., & Todd, R. D. (2003). Autistic traits in the general population: A twin study. *Archives of General Psychiatry*, *60*, 524–530.

Constantino, J. N., & Todd, R. D. (2005). Intergenerational transmission of subthreshold autistic traits in the general population. *Biological Psychiatry*, *57*, 655–660.

Dawson, G., Webb, S., Wijsman, E., Schellenberg, G., Estes, A., Munson, J., et al. (2005). Neurocognitive and

https://doi.org/10.1017/S095457940900008X Published online by Cambridge University Press

electrophysiological evidence of altered face processing in parents of children with autism: Implications for a model of abnormal development of social brain circuitry in autism. *Development and Psychopathology*, *17*, 679–697.

Duvall, J. A., Lu, A., Cantor, R. M., Todd, R. D., Constantino, J. N., & Geschwind, D. H. (2007). A quantitative trait locus analysis of social responsiveness in multiplex autism families. *American Journal of Psychiatry*, *164*, 656–662.

Eaves, L. C., & Ho, H. H., (1996). Brief report: Stability and change in cognitive and behavioral characteristics of autism through childhood. *Journal of Autism and Developmental Disorders*, *26*, 557–569.

Fombonne, E. (2005). Epidemiology of autistic disorder and other pervasive developmental disorders. *Journal of Clinical Psychiatry*, *66*, 3–8.

Lord, C., Rutter, M., Dilavore, P. C., & Risi, S. (1999). *Autism Diagnostic Observation Schedule—WPS* (WPS ed.). Los Angeles: Western Psychological Services.

Maxwell, M. (1992). *Family interview for genetic studies*. Bethesda, MD: National Institute of Mental Health Clinical Neurogenetic Branch, Intramural Research Program.

McDougle, C. J., Scahill, L., Aman, M. G., McCracken, J. T., Tierney, E., Davies, M., et al. (2005). Risperidone for the core symptom domains of autism: Results from the study by the autism network of the research units on pediatric psychopharmacology. *American Journal of Psychiatry*, *162*, 1142–1148.

McGovern, C., & Sigman, M. (2005). Continuity and change from early childhood to adolescence in autism. *Journal of Child Psychology and Psychiatry*, *46*, 401–408.

Mesibov, G. B., Schopler, E., Schaffer, B., & Michal, N. (1989). Use of the Childhood Autism Rating Scale (CARS) with autistic adolescents and adults. *Journal of the American Academy of Child & Adolescent Psychiatry*, *28*, 538–541.

Neale, M. C. (2004). *Mx statistical modeling* (6th ed.). Richmond, VA: Virginia Commonwealth University, Department of Psychiatry. Retrieved from http://www.vcu.edu/mx

Neale, M. C. & Cardon, L. R. (1992). *Methodology for genetic studies of twins and families*. Boston: Kluwer Academic.

Nurnberger, J., Blehar, M., Kaufman, C., York-Cooler, C., Simpson, S., Harkavy-Friedman, J., et al. (1994). Diagnostic Interview for Genetic Studies. Rationale, unique features, and training NIMH Genetics Initiative. *Archives of General Psychiatry*, *51*, 849–859.

Pickles, A., Starr, E., Kazak, S., Bolton, P., Papanikolaou, K., Bailey, A., et al. (2000). Variable expression of the autism broader phenotype: Findings from extended pedigrees. *Journal of Child Psychology and Psychiatry*, *41*, 491–502.

Pine, E., Luby, J., Abbacchi, A., & Constantino, J. N. (2006). Quantitative assessment of autistic symptomatology in preschoolers. *Autism: The International Journal of Research and Practice*, *10*, 344–352.

Piven, J., Harper, J., Palmers, P., & Arndt, S. (1996). Course of behavioral change in autism: A retrospective study of high IQ adolescents and adults. *Journal of the American Academy of Child & Adolescent Psychiatry*, *35*, 523–529.

Piven, J., Palmer, P., Jacohbi, D., Childress, D., & Arndt, S. (1997). Broader autism phenotype: Evidence from a family history study of multiple-incidence autism families. *American Journal of Psychiatry*, *154*, 185–197.

Ronald, A., Happe, F., Price, T. S., Baron-Cohen, S., & Plomin, R. (2006). Phenotypic and genetic overlap between autistic traits at the extremes of the general population. *Journal of the American Academy of Child & Adolescent Psychiatry*, *45*, 1206–1214.

Szatmari, P., Bryson, S. E., Boyle, M. H., Streiner, D. L., & Duku, E. (2003). Predictors of outcome among high functioning children with autism and Asperger syndrome. *Journal of Child Psychology and Psychiatry*, *44*, 520–528.

https://doi.org/10.1017/S095457940900008X Published online by Cambridge University Press