

Differentiating between Autism Spectrum Disorders and Other Developmental Disabilities in Children Who Failed a Screening Instrument for ASD

Pamela Ventola · Jamie Kleinman · Juhi Pandey · Leandra Wilson · Emma Esser · Hilary Boorstein · Thyde Dumont-Mathieu · Gail Marshia · Marianne Barton · Sarah Hodgson · James Green · Fred Volkmar · Katarzyna Chawarska · Tammy Babitz · Diana Robins · Deborah Fein

Published online: 1 August 2006
© Springer Science+Business Media, Inc. 2006

Abstract This study compared behavioral presentation of toddlers with autistic spectrum disorders (ASD) and toddlers with global developmental delay (DD) or developmental language disorder (DLD) who display some characteristics of ASD using the diagnostic algorithm items from the Autism Diagnostic Observation Schedule, Generic (ADOS), the Childhood Autism Rating Scale (CARS), and Modified Checklist for Autism in Toddlers (M-CHAT). To date, 195 children have failed the M-CHAT and have been diagnosed with ASD, DD or DLD. Children with ASD had prominent and consistent impairments in socialization skills, especially joint attention skills and were more impaired in some aspects of communication, play, and sensory processing. Children with ASD and children with DD/DLD shared common features, but certain behavioral markers differentiated the two groups.

Keywords Autism · Early detection · Global developmental delay · Developmental language disorder

Introduction

Young children with autistic spectrum disorders (ASD) share many features with children with other developmental delays, such as global developmental delay and developmental language delay (Charman et al., 1998; Landry & Loveland, 1988; Lord, 1995; Lord, Storoschuk, Rutter, & Pickles, 1993). These similarities contribute to difficulties in accurately diagnosing ASD in very young children. However, accurate diagnosis is crucial for children to receive specialized and appropriate intervention services tailored to their specific needs.

There is a relative scarcity of research that looks specifically at the behavioral differences between children with ASD and children with other early manifesting developmental delays, and most of the literature in the area compares these groups of children for the purpose of validating diagnostic instruments. The language, cognitive, and social differences between children with ASD and children with other delays has been examined, but the majority of this research focuses on older children and not on the behavioral differences that are seen in very young children (Adrien, Deletang, Martineau, Couturier, & Barthelemy, 2001; Bartak, Rutter, & Cox, 1977; Lord & Schopler, 1989). Few studies have focused specifically on the differential diagnosis of ASD in young children (Lord, 1995; Trillingsgaard, Sorensen, Nemeč, & Jorgensen, 2005), and currently, no published studies compare the behavioral differences in toddlers with ASD and toddlers with global developmental or developmental language delay based on standardized and widely used diagnostic instruments as well as a parent-report screening measure.

P. Ventola (✉) · J. Kleinman · J. Pandey · L. Wilson · E. Esser · H. Boorstein · T. Dumont-Mathieu · G. Marshia · M. Barton · S. Hodgson · J. Green · D. Fein
Department of Psychology, University of Connecticut, 406 Babbidge Rd., Storrs, CT 06269-1020, USA
e-mail: pamela.e.dixon@uconn.edu

F. Volkmar · K. Chawarska · T. Babitz
Yale University Child Study Center, New Haven, CT, USA

D. Robins
Georgia State University, Atlanta, GA, USA

In the studies comparing the social abilities and characteristics of children with ASD to those of children with other developmental delays, toddlers with ASD were found to be more impaired than children with other developmental disabilities in joint attention, imitation skills, empathic responding, pointing to express interest, interest in other children, and displaying a range of facial expressions (Charman et al., 1998; Landry & Loveland, 1988; Lord, 1995; Rogers, Hepburn, Stackhouse, & Wehner, 2003; Trillingsgaard et al., 2005). Similar results have been found with slightly older children; participants with ASD were more impaired in specific aspects of social interaction, such as shared enjoyment, pointing to indicate interest, offering comfort, offering to share, eye contact, peer relationships, and overall quality of social overtures (Lord et al., 1993; Lord & Pickles, 1996; Noterdaeme, Sitter, Mildenberger, & Amorosa, 2000). Although certain social behaviors and interaction styles do discriminate between children with ASD and children with other developmental delays, the differential diagnosis can still be difficult in very young children because even though the behaviors discriminate the groups as a whole, many children with global developmental or language delay display at least a few of the characteristic social impairments of children with ASD (Charman et al., 1998; Cox et al., 1999; Lord, 1995; Lord et al., 1993).

Children with ASD also display a specific pattern of impairments in communication relative to children with other developmental delays. In general, the findings indicate that when compared to children with other developmental delays (global developmental delay or developmental language delay), young children with ASD use fewer conventional gestures, especially nodding and shaking their head (Lord et al., 1993; Lord, Rutter, & Le Couteur, 1994), display more echolalia and stereotyped phrases (when speech is present) (Landry & Loveland, 1988; Mildenberger, Sitter, Noterdaeme, & Amorosa, 2001; Noterdaeme et al., 2000), and are less likely to initiate or respond to verbal communication (Lord, 1995; Loveland et al., 1988; Trillingsgaard et al., 2005). Toddlers with ASD and toddlers with developmental delays both have been found to have impairments in their pretend play skills (Baron Cohen et al., 1996; Charman et al., 1998). However, in studies using slightly older children (3½–5 years old), children with ASD show significantly less pretend play than children with other developmental disabilities (Cox et al., 1999; Lord et al., 1994; Noterdaeme et al., 2000; Wainwright & Fein, 1996). Wainwright and Fein (1996) found that even high-functioning preschool children with ASD showed significantly less symbolic play than children

with developmental language disorders, and this difference increased over time during a play session. Even though significant differences in communication have been found in children with ASD and children with other developmental delays, it can still be difficult to differentiate the two groups because some young children with global developmental or developmental language delay display communication characteristics that are common in children with ASD (Charman et al., 1998; Lord, 1995; Lord et al., 1993; Stone et al., 1999; Trillingsgaard et al., 2005).

Several studies have investigated sensory processing in children with ASD, children with global developmental delay, and children with language delays, but the results of the studies in this area have been variable. Some have found that young children with Autistic Disorder display sensory processing atypicalities (i.e. under- or over-responsive to sensory stimulation) and repetitive behaviors (Cox et al., 1999; Lord et al., 1993; Rogers et al., 2003). However, Cox et al. (1999) found that young children with PDD-NOS do not necessarily display these symptoms, and Lord et al. (1993) found that many young children with language or developmental delay do present with sensory symptoms and repetitive behaviors, although perhaps not to the same degree as seen in children with Autistic Disorder (Rogers et al., 2003). Some studies found that repetitive behaviors and sensory processing abnormalities can be used to distinguish children with ASD from children with other developmental delays (Mayes, Volkmar, Hooks, & Cicchetti, 1993; Rogers et al., 2003), whereas other studies found that they could not (Cox et al., 1999; Saemundsen, Magnusson, Smari, & Sigurdardottir, 2003). The apparent variability in the findings of studies on the sensory and repetitive behavior symptoms suggests particular variability in this domain of functioning among children with ASD and developmental delays.

In sum, identifying ASD and differentiating it from other developmental disabilities, such as developmental language delay and global developmental delay, can be challenging, especially in young children. Children with ASD and children with other developmental delays show similar features (Charman et al., 1998; Lord, 1995; Trillingsgaard et al., 2005), many of the current diagnostic instruments do not accurately differentiate ASD and other disabilities in very young children (Cox et al., 1999; Lord et al., 1993; Saemundsen et al., 2003), there is often an overlap between ASD and global developmental delay (Bartak et al., 1977; Vig & Jedrysek, 1999), and the behavioral presentation of young children with delays changes as they mature and develop (Vig & Jedrysek, 1999).

The purpose of the current study is to investigate further the behavioral differences between toddlers with ASD and toddlers with other developmental delays. As will be discussed below, all of the children in the current study failed a screening instrument for ASD (M-CHAT), so they all displayed some characteristics of ASD. Therefore, the DD/DLD sample is probably not representative of all children with delays, but is more representative of children referred for “possible autism.” Clinicians are increasingly asked to diagnose toddlers who display some characteristics of ASD and who are referred for “possible autism,” so studying this group of children who failed the M-CHAT will likely provide useful information on differential diagnosis of ASD in very young children. The data are based on items from two widely-used and standardized diagnostic instruments, ADOS and CARS, and a parent-report screening instrument for ASD, the Modified Checklist for Autism in Toddlers (M-CHAT). The aim of the study is to help with the challenge of differential diagnosis of ASD and other developmental delays in toddlers by identifying behaviors that may differentiate children with ASD from children with other developmental delays, who also display some characteristics of ASD, during the first 2 years of life.

Method

Participants

Participants were 195 children¹ (152 male, 43 female) aged 16–32 months diagnosed with either an ASD (Autistic Disorder or Pervasive Developmental Disorder, Not Otherwise Specified) ($n = 150$), global developmental delay (DD) ($n = 15$) or developmental language disorder (DLD) ($n = 30$). For most of the analyses, the children with DD and the children with DLD were grouped together to form a non-autistic group. All of the children failed the Modified Checklist for Autism in Toddlers (M-CHAT), a parent-report checklist designed to screen for Autistic Spectrum Disorders in 16–30 month old children (Robins, Fein, Barton, & Green, 2001).

The mean chronological age of the entire sample at the time of screening was 24 months, with a range of 16–30 months, and the mean chronological age for the sample at the time of the diagnostic evaluation was 27 months, with a range of 16–32 months. Forty-six of

the children were evaluated before age 24 months. At the time of the diagnostic evaluation, the mean chronological age of the children with ASD was 26.7 months with a standard deviation of 4.4 months, and the mean chronological age of the children with DD/DLD was 27.2 months with a standard deviation of 4.5 months. The mean mental age for the children with ASD was 17 months for non-verbal skills and 14 months for verbal skills, and for the children with DD/DLD, it was 20 months for non-verbal skills and 15 months for verbal skills (Refer to Table 1 for additional sample characteristics). The sample was 86% Caucasian, 8% Latino, 4% Asian, and 2% other. None of the children had been diagnosed with a DSM-IV disorder (APA, 1994) prior to completing the screening instrument.

Instruments

The M-CHAT (Robins et al, 2001) is a 23-item yes–no parent report screening instrument for autistic spectrum disorders (ASD). Initial failure on the screening instrument is defined as any 3 items failed, or any 2 critical items failed. The critical items were identified by discriminant function analysis of children with and without a disorder on the autism spectrum (Robins et al., 2001) and include items concerning joint attention (proto-declarative pointing, bringing to show, following a point), interest in other children, responding to name, and imitation.

The Autism Diagnostic Observation Schedule-Generic (ADOS) (Lord, Rutter, DiLavore, & Risi, 1999) Module 1 is a semi-structured assessment of communication, social interactions and relatedness, play, and imagination. On this measure, the child receives a score in the social domain, in the communication domain, and in the combined social and communication domains. Diagnostic classification is made by exceeding cut-off scores in these three areas (social, communication, and combined). A child can be classified as having Autistic Disorder, Pervasive Developmental Disorder-Not Otherwise Specified, or as non-autistic.

The Autism Diagnostic Interview-Revised (ADI-R) (Lord et al., 1994) is a semi-structured clinician-based interview for parents or caregivers that evaluate the child’s communication, social development, play, and restricted, repetitive, and stereotyped behaviors. The ADI-R has a scoring algorithm that is based on the DSM-IV criteria for autism that yields a classification of either Autistic Disorder or non-autistic; it does not consider PDD-NOS as a possible diagnosis.

The Childhood Autism Rating Scale (CARS) (Schopler, Reichler, & Renner, 1988) consists of 15

¹ Due to changes in the protocol and missing data, the sample sizes for the individual measures are different. Differences are noted when applicable in the methods and results sections.

Table 1 Characteristics of sample

	Gender	Chronologic age in months	Verbal mental age in months ^a	Non-verbal mental age in months ^a	Mental age in months ^b
ASD (<i>n</i> = 150)	82% Male	M = 26.73 SD = 4.47	M = 14.90	M = 17.38	M = 15.14
DD/DLD (<i>n</i> = 45)	76% Male	M = 27.22 SD = 4.54	M = 15.59	M = 20.77	M = 20.00
DD (<i>n</i> = 15)	79% Male	M = 27.09 SD = 4.72	M = 12.37	M = 13.95	M = 20.25
DLD (<i>n</i> = 30)	76% Male	M = 27.51 SD = 4.25	M = 16.41	M = 23.14	M = 21.06

^a Only includes children who received the Mullen (ASD *n* = 89; DD *n* = 7; DLD *n* = 15)

^b Only includes children who received the Bayley (ASD *n* = 45; DD *n* = 12; DLD *n* = 8)

items intended to measure the presence and severity of pervasive developmental disorders. The child is rated on each item based on the clinician's observation of the child's behavior throughout the testing as well as on the parent's report. The CARS includes items on socialization, communication, emotional responses, and sensory sensitivities. The child is classified with Mild-Moderate autism (total score 30–36.5), Severe autism (total score 37 or higher), or as non-autistic (total score 15–30).

Clinical judgment by experienced clinicians is considered to be the “gold standard” for autism diagnosis (Klin, Lang, Cicchetti, & Volkmar, 2000; Spitzer & Siegel, 1990). In this study, the clinicians used the DSM-IV criteria for Autistic Disorder (APA, 1994) on which to base their clinical judgments. A diagnosis of Autistic Disorder was given if the child met the DSM-IV criteria. DSM-IV criteria were also used for PDD-NOS. In most cases, a diagnosis of PDD-NOS was given if the child did not have any repetitive or stereotyped behaviors or if the child's atypical behaviors were not severe or consistent enough to warrant a diagnosis of Autistic Disorder.

Vineland Adaptive Behavior Scales (Sparrow, Balla, & Cicchetti, 1984) is a widely used parent interview scale that assesses adaptive functioning in the areas of socialization, communication, daily living, and motor skills.

Mullen Scales of Early Learning (Mullen, 1989) is a test that is given to the child that measures ability in five domains: gross and fine motor, receptive and expressive language, and visual problem solving. The gross motor domain was not tested as part of this study.

Bayley Scales of Infant Development, Second Edition (Bayley, 1993) is an instrument that measures mental and psychomotor development. It yields a developmental index score of the child's overall development.

Procedure

The children in the current study failed the Modified Checklist for Autism in Toddlers (M-CHAT; Robins et al., 2001). No child had previously received a diagnosis of ASD or any other developmental disorder. If the children were already receiving early intervention, the services were minimal (no more than 1–2 h per week). Once a child failed the screening instrument, the family was contacted for a telephone follow-up. This conversation followed a script with specific examples in which all failed items were reviewed with a parent. If the child still failed the M-CHAT after the telephone follow-up and the family agreed to an evaluation (*n* = 210), the child was given a developmental and diagnostic evaluation. Only children who were diagnosed at this developmental evaluation with either an ASD, global developmental delay, or developmental language disorder were included in the current study. Fifteen children were excluded due to less frequent diagnoses (e.g. 4 with severe motor delays, 1 with cerebral palsy).

Evaluation Procedure

Since all of the children presented for evaluation because of failing the M-CHAT, some degree of risk for developmental disorder was present, so blind assessment was not possible. The evaluations took place at the Psychological Services Clinic at the University of Connecticut (*n* = 147), the Yale Child Study Center (*n* = 20), in the child's home (*n* = 14), or at the early intervention office (*n* = 14). Evaluations were completed by a team of clinicians consisting of one licensed psychologist or developmental pediatrician who specializes in autism and one graduate or post-doctoral student. All children were observed and tested for at least 3 h, and their parents were inter-

viewed extensively using standardized interviews. Parents also completed a questionnaire asking about medical complications and developmental milestones. Due to changes in study protocol, not all of the children received the same battery. The first 74 children received the CARS, clinical judgment diagnosis, the Vineland Adaptive Behavior Scales, and the Bayley Scales of Infant Development. The next 121 children received the ADOS, the ADI-R, CARS, clinical judgment diagnosis, Vineland Adaptive Behavior Scales, and the Mullen Scales of Early Learning. However, as noted in the results section, some data is missing due to error. In almost all cases, parents and children stayed in the same room for the evaluation, allowing both evaluators to observe the child's behavior. Following the evaluation, both clinicians completed the CARS. The reliability for 30 randomly selected pairs of raters was $r = .93$. The CARS completed by the psychologist or developmental pediatrician was used in the data analysis.

All children were diagnosed at this initial evaluation when they were between 16 and 32 months. There was high agreement on diagnostic classification among the ADOS, CARS, and clinical judgment (Ventola et al., in press). For a diagnosis of an ASD, the child's scores on the various diagnostic measures were considered, but the final diagnosis was determined by the clinical judgment of the expert clinicians, which has been found to be the "gold standard" (Klin et al., 2000; Spitzer & Siegel, 1990). The child was diagnosed with global developmental delay if he or she did *not* meet criteria for an ASD *and* if his or her scores on three or more areas of the developmental and adaptive measures (language, non-verbal problem solving, motor skills, adaptive skills) were greater than two standard deviations below the mean. The child was diagnosed with developmental language disorder if he or she did *not* meet criteria for an ASD *and* if his or her scores on either the expressive or receptive subtests of the developmental or adaptive measures were greater than two standard deviations below the mean, and if his or her scores on the non-verbal subtests were less than two standard deviations below the mean. For the ASD group, the mean ADOS reciprocal social interaction score was 9.81, the mean ADOS communication score was 5.74, and the mean CARS total score was 32.94. For the DD/DLD group, the mean ADOS reciprocal social interaction score was 3.05, the mean ADOS communication score 2.95, and the mean CARS total score was 23.20.

To date, 63 children have been seen for a follow-up evaluation. Of those children, 46 were initially given a diagnosis on the autistic spectrum, and 38 continue to

meet criteria for an ASD at follow-up 2–3 years later. Eight children no longer met criteria for an ASD at follow-up. Of these 8 children, two met criteria for developmental language disorder, one met criteria for global developmental delay, and five no longer met criteria for any DSM-IV diagnosis. All of these children received intensive intervention services. Of the 17 children who were not diagnosed with an ASD at their initial evaluation, none met criteria for an ASD at follow-up. This preliminary data demonstrates that the diagnosis given to these children at age 16–32 months seems to be stable and reliable, which is consistent with previous findings (Cox et al., 1999; Gillberg et al., 1990; Stone et al., 1999).

Results

The behavioral differences between children with autistic spectrum disorders (ASD) and children with global developmental or language delay (DD/DLD) were evaluated in several ways. Comparisons of means were performed to determine if cognitive or adaptive level, ADOS diagnostic algorithm items, or CARS items were different for the two groups of children (ASD and DD/DLD). χ^2 analyses and step-wise logistic regression were completed to determine the differences in failure rates between the groups of children for each M-CHAT item. Due to multiple comparisons, the critical P -value was set at .01. For the major analyses, the children with DD and DLD were combined into one group due to small sample sizes in the groups and because the focus of the study was to compare children with ASD to children with other types of developmental delays. However, exploratory analyses were conducted on the DD and DLD groups separately.

Given the possible confound of developmental and language level, communication impairments are critical to consider in group-matching designs (Charman, 2004). Charman (2004) suggests performing ANCOVA with covarying for language level and also to perform analyses with a sub-sample of pair-wise matched cases. The current study follows this suggestion, and the analyses on the M-CHAT, ADOS, and CARS were repeated covarying language level and then with a pair-wise matched sub-sample of cases (ASD: $n = 30$, DD/DLD: $n = 30$, and of this 30, DD $n = 12$ and DLD $n = 18$). The pair-wise matched cases were chosen based on Vineland Communication standard score (VComm). The mean VComm for the ASD group was 71.25, s.d. = 8.43, and the mean VComm for the DD/DLD group was 70.07, s.d. = 8.10. Vineland communication standard score was used for a variety of

reasons. It reflects the children's overall level of language, and as Charman (2004) discusses, a global assessment of language is critical to consider in an ASD sample. Also as Mervis and Klein-Tasman (2004) suggest, matching on standard scores is generally the best measurement choice. Additionally, the Vineland is the only measure that all of the children in the study received and therefore, eliminates the confound of matching on different measures. Furthermore, the Vineland communication standard score (VComm) had greater variability and less of a floor effect than the cognitive scores. VComm scores were more normally distributed than Mullen cognitive scores, with skewness = .88 for VComm compared with skewness = 1.91, 2.27, and 1.49 for Mullen expressive language *t*-score, receptive language *t*-score, and Early Learning Composite standard score, respectively.

Correlations between VComm and other measures of language and cognitive ability were moderate: $r = .67$ for Mullen expressive language *t*-score, $r = .55$ for Mullen receptive language *t*-score, $r = .56$ for Mullen Early Learning Composite standard score, and $r = .43$ for Bayley Mental Developmental Index standard score.

Cognitive and Adaptive Level Analyses

The ASD group had significantly lower standard scores than the DD/DLD group on the Vineland subtests of communication, daily living, socialization, and motor functioning ($P < .01$) (See Table 2). Additionally, for the children who received the Bayley ($n = 65$), the ASD group had a significantly lower overall cognitive score (Bayley Mental Development Index: MDI) than

Table 2 Adaptive and cognitive data for ASD and DD/DLD groups

Scale ^{1,2}	ASD ($n = 150$)	DD/DLD ($n = 45$)	<i>t</i> ^a	<i>P</i> ^a	DD ($n = 15$)	DLD ($n = 30$)	<i>F</i> ^b	<i>P</i> ^b
VABS-Comm								
M	64.33	74.03	5.12	.001	69.15 ^{a,b}	74.73 ^b	24.03	.001
SD	6.32	10.99			11.22	9.65		
VABS-DL								
M	68.59	75.03	2.96	.005	69.00 ^a	77.30	13.15	.001
SD	8.26	11.44			5.28	12.02		
VABS-Soc								
M	67.22	77.50	5.57	.001	69.85 ^a	81.13	31.49	.001
SD	8.36	11.07			8.38	10.28		
VABS-Motor								
M	82.29	90.75	3.55	.001	77.92 ^a	91.27	10.53	.001
SD	10.82	14.17			9.53	14.66		
Bayley MDI ($n = 65$)								
M	52.02	69.45	4.35	.001	61.78 ^{a,b}	73.00 ^b	17.98	.001
SD	10.98	18.02			15.31	18.85		
Mullen-EL ($n = 111$)								
M	25.16	29.14	1.56	.119	20.00 ^a	31.93	6.38	.01
SD	8.21	8.30			0.00	7.53		
Mullen-RL ($n = 111$)								
M	22.45	35.14	2.10	.014	20.00 ^{a,b}	30.20 ^b	4.50	.02
SD	7.73	15.23			0.00	12.14		
Mullen-FM ($n = 111$)								
M	29.48	41.13	1.95	.058	24.00 ^{a,b}	31.93 ^b	5.48	.01
SD	13.04	15.25			8.00	7.53		
Mullen-VR ($n = 111$)								
M	28.10	45.13	3.34	.009	24.50 ^{a,b}	37.56 ^b	6.67	.01
SD	8.86	13.72			9.00	13.57		

¹ Mullen Scales of Early Learning yield *T*-Scores with mean of 50 and standard deviation of 10. Bayley Scales of Infant Development and Vineland Adaptive Behavior Scales yield Standard Scores with mean of 100 and standard deviation of 15

² Mullen EL = Mullen Scales of Early Learning Expressive Language Subtest *T*-Score; Mullen RL = Mullen Scales of Early Learning Receptive Language Subtest *T*-Score; Mullen FM = Mullen Scales of Early Learning Fine Motor Subtest *T*-Score; Mullen VR = Mullen Scales of Early Learning Visual Reception Subtest *T*-Score; VABS-Comm = Vineland Adaptive Behaviors Scales Communication Domain Standard Score; VABS-DL = Vineland Adaptive Behaviors Scales Daily Living Domain Standard Score; VABS-Social = Vineland Adaptive Behaviors Scales Socialization Domain Standard Score; VABS-Motor = Vineland Adaptive Behaviors Scales Motor Domain Standard Score

^a Results based on ASD and DD/DLD group comparisons

^b Results based on ASD, DD, and DLD group comparisons

the DD/DLD group ($P < .001$), and for the children who received the Mullen ($n = 111$), the ASD group had significantly lower visual reception ($P < .01$) scores than the DD/DLD group did. The ASD group had lower receptive and expressive language and fine motor skills as well, but these differences were not significant (See Table 2). When comparing the three groups separately using ANOVA and post-hoc Tukey HSD analyses, the children with DLD had significantly higher scores than the children with ASD on all of the adaptive and cognitive skills (Table 2). The children with DLD had significantly higher scores than the children with DD in daily living skills, adaptive motor skills, adaptive socialization skills, and expressive language skills on the Mullen. There were no significant differences in adaptive or cognitive skills for the children with ASD and the children with DD.

Comparison of ADOS Diagnostic Algorithm Items

Group means were compared on the ADOS diagnostic algorithm items for the ASD children ($n = 79$) and the DD/DLD children ($n = 25$). In the communication domain, frequency of vocalizations directed to others, ($P < .001$), and pointing, ($P < .001$) were significantly different for the two diagnostic groups (Table 3). In the reciprocal social interaction domain, the two groups were significantly different on all of the algorithm items: unusual eye contact, ($P < .001$), facial expression directed to others, ($P < .001$), shared enjoyment in interaction, ($P < .001$), showing, ($P < .001$), spontaneous initiation of joint attention, ($P < .001$), response to joint attention, ($P < .001$), and quality of social overtures, ($P < .001$). For all of these items, the ASD group had significantly higher

scores (indicating greater degree of impairment) than the DD/DLD group (Table 3).

Analyses of covariance (ANCOVA) were then performed using language level (Vineland Communication standard score) as the covariate, and effect size was calculated as measured by η^2 . All of the items remained significant, even with language level controlled (Table 3). With the language matched sub-sample, the results were very similar, except frequency of vocalizations directed towards others was no longer significant.

Comparison of CARS Items

Group means were compared on the CARS items for the ASD children ($n = 123$) and the DD/DLD children ($n = 39$). All but three of the CARS items significantly differentiated between the groups (Table 4), and the ASD group scored significantly higher than the DD/DLD group on the significantly different items: relating to people, ($P < .01$), imitation, ($P < .01$), emotional response, ($P < .01$), body use, ($P < .01$), object use, ($P < .01$), adaptation to change, ($P < .01$), visual response, ($P < .01$), listening response, ($P < .01$), verbal communication, ($P < .01$), nonverbal communication, ($P < .01$), level and consistency of intellectual response, ($P < .01$), and general diagnostic impressions, ($P < .01$).

When controlling for language level by using ANCOVA with Vineland communication standard score as the covariate, the results were similar, and the following items were significantly different: relating to people, imitation, emotional response, body use, object use, visual response, listening response, verbal communication, nonverbal communication, level and consistency of intellectual response, and general diagnostic

Table 3 Comparison between two diagnostic groups (ASD and DD/DLDD) on ADOS Algorithm items

Item	Mean score ASD ($n = 79$)	Mean score DD/DLD ($n = 25$)	<i>F</i>	<i>P</i>	Eta ² (η^2)
Vocalization to others	1.79 ^{a,b}	.95 ^{a,b}	14.13	.001	.24
Stereotyped phrases	0	0	1.66	.206	.04
Use of other's body	.82	.38	1.82	.167	.04
Pointing	2.38 ^{a,b,c}	1.19 ^{a,b,c}	22.89	.001	.34
Gestures	1.49	.67	.28	.602	.01
Eye contact	1.33 ^{a,b,c}	.38 ^{a,b,c}	9.71	.001	.18
Facial expressions	1.0 ^{a,b,c}	.38 ^{a,b,c}	11.28	.001	.20
Shared enjoyment	1.11 ^{a,b,c}	0 ^{a,b,c}	20.57	.001	.33
Showing	1.67 ^{a,b,c}	.48 ^{a,b,c}	67.53	.001	.60
Initiation of joint attention	1.63 ^{a,b,c}	.81 ^{a,b,c}	20.94	.001	.32
Response to joint attention	1.61 ^{a,b,c}	.52 ^{a,b,c}	22.45	.001	.33
Social overtures	1.67 ^{a,b,c}	.62 ^{a,b,c}	29.34	.001	.39

^a Significant at the ($P < .01$) level on *t*-test

^b Significant at the ($P < .01$) level on ANCOVA when controlling for language level

^c Significant at the ($P < .01$) level on *t*-test with language matched sub-sample

Table 4 Comparison of CARS scores between the two diagnostic groups (ASD and DD/DLD)

Item	Mean score ASD (<i>n</i> = 123)	Mean score DD/DLD (<i>n</i> = 39)	<i>F</i>	<i>P</i>	Eta ² (<i>η</i> ²)
Relating to people	2.55 ^{a,b,c}	1.45 ^{a,b,c}	41.42	.001	.40
Imitation	2.65 ^{a,b,c}	1.67 ^{a,b,c}	32.52	.001	.34
Emotional response	1.97 ^{a,b,c}	1.42 ^{a,b,c}	13.52	.001	.18
Body use	2.08 ^{a,b,c}	1.45 ^{a,b,c}	18.12	.001	.23
Object use	2.19 ^{a,b,c}	1.48 ^{a,b,c}	22.08	.001	.26
Adaptation to change	1.88 ^a	1.42 ^a	4.47	.013	.07
Visual response	2.08 ^{a,b,c}	1.40 ^{a,b,c}	13.86	.001	.18
Listening response	2.38 ^{a,b,c}	1.67 ^{a,b,c}	20.80	.001	.25
Taste, smell, and touch response	1.76	1.45	4.32	.040	.03
Fear	1.54	1.25	4.64	.033	.04
Verbal communication	2.96 ^{a,b,c}	2.45 ^{a,b,c}	6.00	.001	.23
Nonverbal communication	2.50 ^{a,b,c}	1.57 ^{a,b,c}	39.94	.001	.40
Activity level	1.73	1.43	3.86	.052	.03
Level of intellectual response	2.40 ^{a,b}	1.90 ^{a,b}	30.46	.001	.33
General impressions	2.50 ^{a,b,c}	1.25 ^{a,b,c}	87.53	.001	.59

^a Significant at the (*P* < .01) level on *t*-test

^b Significant at the (*P* < .01) level on ANCOVA when controlling for language level

^c Significant at the (*P* < .01) level on *t*-test with language matched sub-sample

impressions. Effect sizes were estimated using (η^2 (Table 4). With the language matched sub-sample of cases, the results were similar, but item 14 (level and consistency of intellectual responses) was no longer significant (see Table 4).

Comparison of M-CHAT Items

Failure percentages are displayed by item for each group (Table 5). χ^2 analyses were done to determine the difference in failure rates by item for each group.

Table 5 Percent failure for diagnostic groups

Item	ASD <i>n</i> = 150	DD/DLD (together) <i>n</i> = 45	DD <i>n</i> = 15	DLD <i>n</i> = 30
1 (enjoy being swung/bounced)	1.9	8.7	6.3	8.8
2 (interest in children)	55.7 ^a	32.6 ^a	12.5	38.2
3 (enjoy climbing)	6.3	8.7	6.3	11.8
4 (peek-a-boo)	17.1	17.4	6.3	26.5
5 (pretend play)	62.7	45.7	56.3	35.3
6 (point to ask)	84.2 ^{a,b}	54.3 ^{a,b}	81.3	38.2
7 (point for interest)	88.0 ^{a,b,c}	54.3 ^{a,b,c}	68.8 ^d	47.1 ^d
8 (play properly)	43.0	34.8	37.5	35.3
9 (bring to show)	51.9	41.3	50.0	35.3
10 (eye contact)	27.8	19.6	18.8	17.6
11 (oversensitive to noise)	38.6	52.2	56.3	47.1
12 (respond to smile)	19.0	8.7	18.8	2.9
13 (imitation)	65.2	54.3	68.8	41.2
14 (response to name)	54.4 ^{a,b,c}	21.7 ^{a,b,c}	6.3	26.5
15 (follow point)	75.9 ^{a,b,c}	43.5 ^{a,b,c}	43.8	41.2
16 (walk)	1.9	2.2	6.3	0.0
17 (follow gaze)	57.0 ^a	34.8 ^a	43.8	29.4
18 (finger movements)	37.3 ^a	17.4 ^a	12.5	20.6
19 (attract attention to activity)	74.7 ^a	56.5 ^a	62.5	50.0
20 (appear deaf)	47.5 ^a	26.1 ^a	25.0	23.5
21 (understand what others say)	58.9 ^a	32.6 ^a	18.8	32.4
22 (stare at nothing)	60.1	47.8	50.0	41.2
23 (check parent's reaction)	39.9 ^a	21.7 ^a	12.5	20.6

^a Significant at the (*P* < .01) level for the initial χ^2 with the ASD and DD/DLD groups

^b Significant at the (*P* < .01) level for the step-wise logistic regression controlling for language level

^c Significant at the (*P* < .01) level for the language matched sub-sample of cases

^d Significant at the (*P* < .01) level for the comparison between the DD and DLD groups

Table 5 indicates which items were significantly different by group. The ASD group failed all of the significantly different items more frequently than the DD/DLD group. An exploratory χ^2 analysis investigated the differences on M-CHAT items between the DD and DLD groups. One item: pointing for interest, ($P < .01$), was significantly different, with the DD group failing significantly more frequently than the DLD group.

Step-wise logistic regression for the M-CHAT items covarying for language level was completed. z -scores were calculated as an estimate of effect size. When overall language level was controlled, four M-CHAT items remained significant: response to name, ($P < .001$), z -score = 3.1, pointing for interest, ($P < .01$), z -score = 3.5, pointing to request, ($P < .01$), z -score = 3.2, and following a point, ($P < .01$), z -score = 3.1. With the matched sub-sample of 60 cases, similar items were significant: pointing for interest, ($P < .01$), following a point, ($P < .01$), and response to name, ($P < .01$) (see Table 5).

Discussion

The purpose of the current study was to investigate the differences between children with ASD who failed a screening instrument for ASD (M-CHAT) and children with other developmental delays who also failed the M-CHAT. Although the children with DD/DLD are likely not representative of the general population of children with these delays, as all of the children in the current study failed an ASD-specific screening, the type of sample included in the current study (children who display some symptoms of ASD) is as important to study as a more representative sample of children with delays. With increasing frequency, clinicians and researchers are asked to diagnose ASD and DD/DLD in very young children who may not have a clear presentation. As the results of the current study indicate, some children with DD/DLD display characteristics of ASD, and it is these children who often present challenges for differential diagnosis. Therefore, the aim of the current study is to begin to help clarify some of the behaviors that can be used by both clinicians and researchers to aid in the difficult process of differential diagnosis of ASD in toddlers.

The differences between the children with ASD and the children with DD/DLD (who also failed the M-CHAT) were explored based on cognitive and adaptive skills, ADOS algorithm items, CARS items, and M-CHAT items. Analyses on the diagnostic instrument items were done with and without covarying

for overall language level using the Vineland Communication standard score and also with a smaller sub-set of pair-wise matched cases as suggested by Charman (2004) and Mervis and Klein-Tasman (2004).

The cognitive, language, and adaptive skills were compared between the two groups of children. The children with ASD scored lower than the children with DD/DLD in all areas (adaptive skills, expressive language, receptive language, fine motor, and visual reception skills), and the differences were significant in all but the receptive and expressive language and fine motor subtests of the Mullen. The Bayley Scales of Infant Development yields a cognitive composite score (i.e. IQ score), whereas the Mullen Scales of Early Learning yields separate scores for each domain. Therefore, analyses on the children who received the Mullen lend additional explanation as to the source of difference in the Bayley scores: the children with ASD had significantly lower visual problem solving skills but also lower receptive and expressive language and fine motor skills. These results indicate the children with ASD were, in general, more impaired than the children with DD/DLD, but the language skills were not significantly different between the two groups possibly because language delays often are the reason for referral or concern on the part of the parent. Therefore, virtually all of the children in the study presented with language delays. Fine motor skills also did not differentiate between the groups. One possible explanation for this finding is that motor skills tended to be the relative strength for most of the children in the study, especially the children with ASD. Another explanation is that since the actual differences between the groups on all of the sub-tests are fairly similar, the statistical significance of some and not all is likely due to power limitations and the relatively small sample size of the DD/DLD group. However, overall, the adaptive and cognitive results indicate that in the current sample, the children with ASD were at a younger developmental level. Therefore, as discussed above, overall level of language was a covariate in the analyses in efforts to control for the differences in developmental level.

When comparing the children with ASD to the children with other developmental delays on the ADOS algorithm items, many of the items were found to differentiate the children. When language level was controlled, all of the items from the reciprocal social interaction domain (eye contact, shared enjoyment, showing, initiation of joint attention, response to joint attention, and quality of social overtures) and one of the items from the communication domain (pointing) were found to significantly differentiate the groups.

The results were very similar with just the pair-wise language matched sub-sample. In terms of the effect sizes, the items relating to joint attention accounted for the greatest amount of variance. For example, diagnostic grouping accounted for over 50% of the variance in 'showing.' Since all of the algorithm items from the reciprocal social interaction domain significantly differentiated between the groups, even after language level was controlled, and since the effect sizes were greatest for the social items, it indicates that social deficits and especially joint attention deficits are relatively unique to ASD. Many communication items did not differentiate between the two groups and had lower effect sizes, as children in both groups scored relatively high (indicating a high degree of impairment) on these items (Refer to Table 3). This is likely reflective of the delayed communication skills seen in both children with ASD as well as children with DD/DLD. These results support the findings of previous studies discussed in the introduction and are especially similar to the findings of Lord (1995) and Trillingsgaard et al. (2005). These two studies found that social behaviors, especially joint attention behaviors are especially salient deficits in children with ASD when compared to same-aged children with other delays.

When comparing the two groups of children on the CARS, all of the social and communication items as well as some of the atypical sensory-related items differentiated the ASD group from the group with other delays. For all items, the ASD group scored higher (more significant degree of impairment) than the DD/DLD group. When language level was controlled, the results were very similar, indicating that the CARS is relatively robust to developmental level and is measuring autistic symptomology, as opposed to developmental level. As with the ADOS, the effect sizes were greatest for the social relatedness items, especially 'relating to people' and the also 'non-verbal communication' item. One of the sensory items (taste, smell, touch responses) and some of the general behavior items (e.g. activity level and adaptation to change) did not differentiate between the groups, as both groups showed relatively mild impairments on these items. Therefore, as with the ADOS, many of the items that differentiated between the two groups and had the highest effect sizes related especially to social interactions and also non-verbal communication skills, supporting findings from numerous other studies referenced in the introduction. Some of the CARS items relating to sensory responses did differentiate the groups, whereas this was not seen with the ADOS, as these behaviors are not measured by the ADOS algorithm items. It should be noted that the CARS in the

present study was being used by clinician/researchers who were quite experienced in diagnosing ASD in toddlers and clinical judgment played a significant role in completing the CARS ratings.

On the M-CHAT, the children with ASD failed most items significantly more frequently than the children with DD/DLD did. However, once the overall language level of the children was controlled, only four items remained significantly different, and the *z*-scores, which served as an estimate of effect size, were high (all over 3.0) for these four items as well. These items (response to name, pointing for interest, pointing to request, and ability to follow a point) relate to joint attention skills and social responsiveness. Since controlling for language level eliminated some of the differences between the two groups, it indicates that some of the differences were due not to autistic symptomology, but to differences in developmental, and particularly, language level. The children with global developmental delay (DD) were also compared to the children with developmental language disorder (DLD), and the children with DD failed one item 'point for interest' significantly more than the DLD group did. However, given the small sample sizes for the DD and DLD groups, these results should be interpreted with caution (Refer to Table 5 for failure frequencies of these groups).

Robins et al. (2001) found that 21 of the 23 M-CHAT items significantly differentiated children with ASD from typically developing children, and the current study found that, after controlling for language level, 4 of the 23 M-CHAT items significantly differentiated the children with ASD and the children with DD/DLD. However, all of the items that differentiated the children with ASD from the children with DD/DLD were also found to differentiate typically developing children from children with ASD (Robins et al., 2001). Therefore, fewer, yet overlapping, items differentiated children with ASD from children with DD/DLD than differentiated typically developing children from children with ASD, which indicates, as current research (Baron-Cohen et al., 1996; Lord et al., 1993; Wainwright & Fein, 1996) and the results of the ADOS and CARS analyses from the current study suggest, that children with ASD and children with DD/DLD have many common characteristics that are not necessarily seen in typically developing children. Nonetheless, there are behavioral markers, especially joint attention deficits that are seen in children with ASD significantly more frequently than in typically developing children or in children with DD/DLD, indicating that these symptoms may be very central to ASD.

Overall, the results from the current study indicate that children with ASD are significantly different than children with DD/DLD on a variety of behaviors as measured by both diagnostic instruments and a screening measure. Most striking is the significantly more prominent and consistent impairment in social interaction skills, especially joint attention skills, in children with ASD. This finding supports previous research on differential diagnosis (Lord, 1995; Trillingsgaard et al., 2005) that impairments in socialization, most notably, joint attention skills are one of the more unique and central deficits seen in very young children with ASD. The ASD group was also more impaired on other behaviors, such as imitation, facial expressions, eye contact, and sensory responses, on at least one of the diagnostic measures. Future studies comparing children's scores on different diagnostic and screening instruments specifically looking at the effects of the differences between sources of information, rater differences, and the different scoring procedures of the instruments, will greatly help to clarify and provide more information about the differential diagnosis of ASD in very young children.

It is important to note that although the M-CHAT detected the relative deficit in socialization in the children with ASD, the diagnostic measures detected more differences between the groups. This is probably reflective of the selection criteria of the current study. As discussed, the children were selected based on failing items on the M-CHAT. Therefore, since both groups needed to fail at least three items in order to be included, the differences between the groups on M-CHAT scores would likely be reduced. In addition, the M-CHAT is scored by yes/no responses, whereas the other measures allow for continuous scores from either 0–3 or 1–4. Therefore, the diagnostic measures are more reflective of subtle differences in behavior and allow for a more detailed consideration of the behavior in question. In sum, although the results from M-CHAT reflected the greater impairment in joint attention in children with ASD, it is a screening instrument and can not be used in place of the standardized diagnostic instruments.

Some limitations of the current study included the relatively small sample size in the DD/DLD group. The small sample size in this group precluded complete comparative analyses on the DD and DLD separately. Another limitation was that all of the children presented for evaluation after failing a screening measure, which indicated some degree of risk for developmental disability, so there was no normal control group. In addition, as discussed above, all of the children in the study failed the M-CHAT, which is designed to detect

autistic symptoms. Therefore, the sample of DD and DLD children is likely not representative of all children with these delays. Another limitation is that diagnosis was made partly on the basis of ADOS and CARS scores. Therefore, the differences between the groups on items from these measures may be affected by their influence on final diagnosis.

Acknowledgements This research was supported by the University of Connecticut's Research Foundation Faculty Grant, National Alliance of Autism Research, National Institute of Child Health and Human Development (5 R01 HD039961), Maternal and Child Health Bureau (R40 MC 00270). We are very grateful to the children and families who participated in this research, as well as the doctors and early intervention sites that screened the children. We would also like to acknowledge the invaluable assistance of the undergraduate research assistants.

References

- Adrien, J. L., Deletang, N., Martineau, J., Couturier, G., & Barthelemy, C. (2001). Regulation of cognitive activity and early communication development in young autistic, mentally retarded, and young normal children. *Developmental Psychobiology*, *39*, 124–136.
- American Psychiatric Association (1994). *Diagnostic and statistics manual of mental disorders* (4th ed., DSM-IV). Washington, DC: Author.
- Baron Cohen, S., Cox, A., Baird, G., Swettenham, J., & Nighingale, N. (1996). Psychological markers in the detection of autism in infancy in a large population. *British Journal of Psychiatry*, *168*(2), 158–163.
- Bartak, L., Rutter, M., & Cox, A. (1977). A comparative study of infantile autism and specific developmental receptive language disorders: III. *Journal of Autism and Childhood Schizophrenia*, *7*, 383–396.
- Bayley, N. (1993). *The Bayley scales of infant development* (2nd ed.). San Antonio, TX: Hartcourt Brace.
- Charman, T., Swettenham, J., Baron Cohen, S., Cox, A., Baird, G., & Drew, A. (1998). An experimental investigation of social-cognitive abilities in infants with autism: Clinical implications. *Infant Mental Health Journal*, *19*(2), 260–275.
- Charman, T. (2004). Matching preschool children with autism spectrum disorders and comparison children for language ability: Methodological challenges. *Journal of Autism and Developmental Disorders*, *34*(1), 59–64.
- Cox, A., Klein, K., Charman, T., Baird, G., Baron Cohen, S., Swettenham, J. Drew, A., & Wheelwright, S. (1999). Autism spectrum disorders at 20 and 42 months of age: Stability of clinical and ADI-R diagnosis. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *40*(5), 719–732.
- Gillberg, C., Ehlers, S., Schaumann, H., Jacobsen, G., Dahlgren, S., Lindblom, R., Bagenholm, A., Tjuus, T., & Blinder, B. (1990). Autism under age 3 years: A clinical study of 28 cases referred for autistic symptoms in infancy. *Journal of Child Psychology and Psychiatry*, *6*, 921–934.
- Klin, A., Lang, J., Cicchetti, D., & Volkmar, F. (2000). Brief report: Interrater reliability of clinical diagnosis and DSM-IV criteria for autistic disorder: Results of the DSM-IV autism field trial. *Journal of Autism and Developmental Disorders*, *30*(2), 163–167.

- Landry, S. H., & Loveland, K. A. (1988). Communication behaviors in autism and developmental language delay. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 29(5), 621–634.
- Lord, C., & Schopler, E. (1989). The role of age at assessment, developmental level, and test in the stability of intelligence scores in young autistic children. *Journal of Autism and Developmental Disorders*, 19, 483–499.
- Lord, C., Storoschuk, S., Rutter, M., & Pickles, A. (1993). Using the ADI-R to diagnose autism in preschool children. *Infant Mental Health Journal*, 14(3), 234–252.
- Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism diagnostic interview—revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, 24(5), 659–685.
- Lord, C. (1995). Follow-up of two-year-olds referred for possibly autism. *Journal of Child Psychology and Psychiatry*, 36, 1–18.
- Lord, C., & Pickles, A. (1996). Language level and nonverbal social-communicative behaviors in autistic and language-delayed children. *Journal of the American Academy of Child and Adolescent Psychiatry*, 35(11), 1542–1550.
- Lord, C., Rutter, M., DiLavore, P., & Risi, S. (1999). *Autism diagnostic observation schedule—WPS edition*. Los Angeles, California: Western Psychological Services.
- Loveland, K. A., Landry, S. H., Hughes, S. O., Hall, S. K., & McEvoy, R. E. (1988). Speech acts and the pragmatic deficits of autism. *Journal of Speech and Hearing Research*, 31(4), 593–604.
- Mayes, L., Volkmar, F. R., Hooks, M., & Cicchetti, D. V. (1993). Differentiating pervasive developmental disorder not otherwise specified from autism and language disorders. *Journal of Autism and Developmental Disorders*, 23(1), 79–90.
- Mervis, C. B., & Klein-Tasman, B. P. (2004). Methodological issues in group-matching designs: A levels for control variable comparisons and measurement characteristics of control and target variables. *Journal of Autism and Developmental Disorders*, 34(1), 7–17.
- Mildenberger, K., Sitter, S., Noterdaeme, M., & Amorosa, H. (2001). The use of the ADI-R as a diagnostic tool in the differential diagnosis of children with infantile autism and children with a receptive language disorder. *European Journal of Child and Adolescent Psychiatry*, 10(5), 248–255.
- Mullen, E. (1989). *Mullen scales of early learning*. Cranston, RI: T.O.T.A.L. Child, Inc.
- Noterdaeme, M., Sitter, S., Mildenberger, K., & Amorosa, H. (2000). Diagnostic assessment of communicative and interactive behaviours in children with autism and receptive language disorder. *European Child and Adolescent Psychiatry*, 9(4), 295–300.
- Robins, D., Fein, D., Barton, M., & Green, J. (2001). The modified checklist for autism in toddlers: An initial study investigating the early detection of autism and pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, 31(2), 131–144.
- Rogers, S. J., Hepburn, S., Stackhouse, T., & Wehner, E. (2003). Parent reports of sensory symptoms in toddlers with autism and those with other developmental disorders. *Journal of Autism and Developmental Disorders*, 33(6), 631–642.
- Saemundsen, E., Magnusson, P., Smari, J., & Sigurdardottir, S. (2003). Autism diagnostic interview-revised and the childhood autism rating scale: Convergence and discrepancy in diagnosing autism. *Journal of Autism and Developmental Disorders*, 33(3), 319–328.
- Schopler, E., Reichler, R. J., & Renner, B. R. (1988). *The Childhood Autism Rating Scale*. Los Angeles: Western Psychological Services.
- Sparrow, S., Bala, D., & Cicchetti, D. (1984). *The vineland adaptive behavior scales*. Circle Pines, MN: American Guidance Services.
- Spitzer, R., & Siegel, B. (1990) The DSM-III-R field trial of pervasive developmental disorders. *Journal of the American Academy of Child and Adolescent Psychiatry*, 29(6), 855–862.
- Stone, W. L., Lee, E. B., Ashford, L., Brissie, J., Hepburn, S. L., Coonrod, E. E. et al. (1999). Can autism be diagnosed accurately in children under 3 years? *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 40(2), 219–226.
- Trillingsgaard, A., Sorensen, E. U., Nemeč, G., & Jorgensen, M. (2005). What distinguishes autism spectrum disorders from other developmental disorders before the age of four years? *European Journal of Child and Adolescent Psychiatry*, 14, 65–72.
- Ventola, P., Kleinman, J., Pandey, P., Barton, M., Allen, S., Green, J., Robins, D., & Fein, D. Agreement among four diagnostic instruments for autism spectrum disorders in toddlers. *Journal of Autism and Developmental Disorders* (in press).
- Vig, S., & Jedrysek, E. (1999). Autistic features in young children with significant cognitive impairment: Autism or mental retardation? *Journal of Autism and Developmental Disorders*, 29(3), 235–248.
- Wainwright, L., & Fein, D. (1996). Results of play analyses. In Rapin, I. (Editor for the Autism and Language Disorders Collaborative Project: Preschool Study Group). *Preschool children with inadequate communication: Developmental language disorder, autism, low IQ. Clinics in developmental medicine, No. 139*. London: Mac Keith Press, pp. 173–189.