

Early Indicators of Autism Spectrum Disorders in the Second Year of Life

Amy M. Wetherby^{1,3}, Juliann Woods¹, Lori Allen¹, Julie Cleary¹, Holly Dickinson¹ and Catherine Lord²

Three groups of 18 children were selected for this study, one group with autism spectrum disorders (ASD), one group with developmental delays in which ASD was ruled out (DD), and one group with typical development (TD), from a pool of 3026 children who were screened with the Communication and Symbolic Behavior Scales Developmental Profile (CSBS DP, Wetherby & Prizant, 2002) Infant-Toddler Checklist under 24 months of age. The CSBS DP Behavior Sample was videotaped on selected children as a second-level evaluation during the second year of life. The Infant-Toddler Checklist had a sensitivity and specificity of 88.9% for this sample of children. Significant group differences were found on the Infant-Toddler Checklist and the Behavior Sample, however, these differences did not distinguish children with ASD and DD with high accuracy. The videotapes of the Behavior Sample were reanalyzed to identify red flags of ASD. Nine red flags differentiated children in the ASD group from both the DD and TD groups and four red flags differentiated children in the ASD Group from the TD group but not the DD group. These 13 red flags were found to discriminate the three groups with a correct classification rate of 94.4%.

KEY WORDS: Autism, CSBS DP, developmental delay, early identification, screening,

INTRODUCTION

There is now mounting evidence demonstrating the effectiveness of intensive early intervention with a substantial proportion of young children with autism spectrum disorders (ASD; Dawson & Osterling, 1997; National Research Council, 2001). Research indicates that intervention provided before age 3½ has a greater impact than that after age 5 (Fenske, Zalski, Krantz, & McClannahan, 1985; Harris & Handleman, 2000). In spite of the severity of the behavioral characteristics of most children with ASD, the average age for diagnosis in the United States is not until 3–4 years (Filipek *et al.*, 1999).

The findings of McGee, Morrier, and Daly (1999) suggest that intervention beginning before age 3 may have an even greater impact. Thus, there is a pressing need to improve early identification so that children with ASD can access intervention as early as possible. Although there have been significant advances in genetic and biomedical research on ASD, there is currently no biological marker for either autism or ASD, therefore screening and diagnosis must be based on behavioral features (Filipek *et al.*, 1999).

Core Social Communication Deficits in Preschool Children with ASD

Preschool children with ASD have a distinct profile of social communication that has important implications for early identification. Numerous studies have documented a deficit in joint attention skills of children with ASD, including difficulties using eye gaze to coordinate attention, following

¹ Florida State University

² University of Michigan

³ Correspondence should be addressed to Amy M. Wetherby, Department of Communication Disorders, RRC 107, Florida State University, Tallahassee, Florida, 32306-7814. E-mail: awetherb@fsu.edu

the attentional focus of another person, and drawing another's attention to an object or event of interest (Mundy, Sigman, & Kasari, 1990; Stone, Ousley, Yoder, Hogan, & Hepburn, 1997; Wetherby & Prutting, 1984; Wetherby, Prizant, & Hutchinson, 1998). Longitudinal research findings suggest that the failure to acquire gestural joint attention may be a core deficit in ASD and a critical milestone that impairs language development (Mundy *et al.*, 1990; Sigman & Ruskin, 1999).

There is much variability in the capacity to use vocal communication in young children with ASD, which likely contributes to the wide range of speech and language skills. Some children with ASD have been found to use a limited consonant inventory and less complex syllabic structure, while others show adequate complexity of vocalizations (McHale, Simeonsson, Marcus, & Olley, 1980; Stone & Caro-Martinez, 1990; Wetherby & Prutting, 1984; Wetherby *et al.*, 1998). Sheinkopf, Mundy, Oller, and Steffens (2000) found that compared to children with developmental delays (DD), preschool children with ASD used a comparable proportion of syllables containing consonants but a significantly greater proportion of syllables with atypical phonation, such as squeals, growls, and yells. The vocal atypicalities were independent of joint attention deficits in this small sample but were negatively correlated with mental age, suggesting that the joint attention and vocal deficits arise from different pathological processes.

Deficits in the capacity to use conventional and symbolic gestures have been documented in numerous studies. Children with ASD do not compensate for their lack of verbal skills with gestures, but rather, they show limited gestural use both in quantity and quality. They predominantly use primitive motoric gestures (i.e., contact gesture of leading or pulling another's hand) to communicate. They lack the use of many conventional gestures, such as showing, waving, pointing, and symbolic gestures, such as nodding head and depicting actions (Loveland & Landry, 1986; McHale, Simeonsson, Marcus, & Olley, 1980; Stone & Caro-Martinez, 1990; Stone *et al.*, 1997; Wetherby *et al.*, 1998). In lieu of conventional means of communicating, children with ASD may develop idiosyncratic, unconventional, or inappropriate behaviors to communicate, such as self-injurious behavior, aggression, or tantrums.

Children with ASD show significant deficits in symbolic or make-believe play (i.e., using pretend

actions with objects) and limited abilities to use objects functionally (Dawson & Adams, 1984; Sigman & Ungerer, 1984; Wetherby *et al.*, 1998; Wetherby & Prutting, 1984; Wing, Gould, Yeates, & Brierly, 1977). It is noteworthy that a lack of varied, spontaneous make-believe play is one of the four possible features of the impairment in communication in the DSM-IV (APA, 1994). Functional and symbolic play have been found to be significantly correlated with receptive and expressive language (Mundy, Sigman, Ungerer, & Sherman, 1987; Sigman & Ruskin, 1999), suggesting a shared symbolic deficit.

Exploring developmental profiles of social communication has contributed to distinguishing children with ASD from children with other DD. Studies by Stone *et al.* (1997) and Wetherby *et al.* (1998) examined the developmental profiles of 2- to 4-year-old children with ASD compared to children with DD who were at the same language stage, using similar strategies for gathering communication samples. The children with ASD showed comparable or higher use of communication to request and protest, but significantly less use of gaze shifts, shared positive affect, conventional gestures, coordinated gestures with vocalizations and eye gaze, and communication for joint attention. The children with ASD performed at comparable levels of constructive play but significantly poorer levels of language comprehension and symbolic play.

Research on social communication in preschool children with ASD has important implications for earlier identification of children with ASD because the deficits identified are in skills that typically develop during the first 12–18 months of life. These findings suggest that there is a constellation of prelinguistic behaviors that may be important early indicators of ASD and may distinguish children with ASD from both typically developing children and children with other DD. These prelinguistic behaviors include gaze shifts, shared positive affect, gaze/point following, communication for joint attention, conventional gestures, communicative vocalizations, language comprehension, and symbolic play.

Earliest Indicators of ASD in Infants and Toddlers

The diagnostic features of ASD should be evident in very young children since they involve abilities that typically develop in the first 2 years of life. Most children identified as having ASD are

reported by their caregivers to demonstrate symptoms within the first 2 years of life, based on retrospective accounts (Short & Schopler, 1988; Wimpory, Hobson, Williams & Nash, 2000). Furthermore, most families initially express concern to their pediatrician by the time their child is 18 months old (Howlin & Moore, 1997; Siegel, Pliner, Eschler, & Elliot, 1988).

Although most children with ASD are not diagnosed until at least 3 years of age, a clinical diagnosis of ASD at 2 years of age was found to be associated with the same diagnosis at 3 years of age or older in 90% of children studied (Lord, 1995; Stone *et al.*, 1999). These studies indicate that impairments in social interaction and impairments in communication were found to be evident by 2 years but restricted and repetitive activities and interests were not evident in some children until closer to 3 years of age. Lord (1995) found that lack of seeking another's attention and lack of response to voice were the clearest discriminators of children with ASD at 2 years based on parent interview. She also found that children with ASD differed from other children with DDs in understanding gesture, unusual use of others' bodies, seeking to share enjoyment, hand and finger mannerisms, and unusual sensory behavior.

Further support for early social and communicative deficits is provided by Wimpory *et al.* (2000), who conducted a semi-structured interview of 10 parents of preschoolers with ASD and 10 parents of matched control children with DD. They found that limitations on social engagement based on retrospective parent report, such as lack of appropriate eye gaze, showing objects, pointing to objects, following others point, waving, and preverbal turn-taking, distinguished the children with ASD from the controls.

Research on children under 24 months of age who are later diagnosed with ASD further clarifies patterns of early indicators of ASD. One research approach has been a longitudinal, prospective study conducted by Baron-Cohen and colleagues (Baron-Cohen, Allen, & Gillberg, 1992; Baron-Cohen *et al.*, 1996; Baird *et al.*, 2000) using the *Checklist for Autism in Toddlers* (CHAT). The CHAT consists of nine items reported by parents and five items observed by a health professional at the 18-month checkup. Baron-Cohen *et al.* (1996) screened over 16,000 children using the CHAT and identified 19 children at 18 months who were later diagnosed with ASD based on failure of the following three

key items: (1) protodeclarative pointing, (2) pretend play, and (3) gaze monitoring. However, in a follow-up study of children in this sample at age 7 years, 94 cases of ASD were identified (Baird *et al.*, 2000). These findings indicate that the CHAT has a specificity of 98% but a sensitivity of 38%, and missed many children at 18 months who were later diagnosed with ASD. While the validity of the CHAT is disappointing, it indicates the need for further research on young children with ASD and provides important clues to early indicators of ASD, based on the children they were able to identify early.

Robins, Fein, Barton, and Green (2001) have investigated the Modified Checklist for Autism in Toddlers (M-CHAT), which consists of 23 yes/no items reported by parents, to screen a nonselected pediatric sample of 1,122 children and high-risk sample of 171 children. Half of the sample was screened at 18 months of age and half at 24 months. For children who failed the M-CHAT, a phone follow-up was conducted to confirm checklist responses, and families of children who still failed were invited to participate in a developmental evaluation. Fifty-eight children were evaluated; 39 of the 58 received a diagnosis of ASD, and 3 of the 39 were from the unselected pediatrician sample. They have not yet completed a follow-up of the sample, and therefore, sensitivity and specificity cannot yet be determined. However, the results of a discriminant function analysis indicated that six items pertaining to social relatedness and communication distinguished the children diagnosed with ASD from the children who did not fail the M-CHAT: protodeclarative pointing, following a point, bringing objects to show parent, interest in other children, imitation, and responding to name.

Another line of research has been to study home videotapes of children who are later diagnosed with ASD. Osterling and colleagues (Osterling & Dawson, 1994, 1999; Werner, Dawson, Osterling, & Dinno, 2000) conducted a retrospective study using home videotapes of first year birthday parties. Osterling and Dawson (1994) found that children later diagnosed with ASD displayed significantly fewer social and communicative behaviors and significantly more autistic symptoms at this young age compared to typically developing controls. Lack of the following four behaviors correctly classified 10 of the 11 children with ASD: pointing, showing objects, looking at the face of another, and orienting to name. Osterling and Dawson (1999)

added a control group of children with DDs who did not have ASD and found that only the latter two of the four behaviors distinguished these children from the children with ASD. Furthermore, using videotapes collected at 8–10 months of age on these 11 children and 4 additional children later diagnosed with ASD, these same two behaviors distinguished infants with ASD and typically developing controls at 8–10 months of age (Werner *et al.*, 2000). Baranek (1999) examined home videotapes of 9–12 month olds with 11 children with ASD, 10 with developmental disabilities, and 11 typically developing children. She found that measures of sensory-motor functioning including response to social touch and motor and object stereotypies distinguished children with ASD, in addition to indicators of social responsiveness. These results suggest that impairments in these early social, communication, and sensory-motor behaviors should contribute to earlier detection of ASD and that differences should be detectable by the end of the first year of life, if not earlier.

A multidisciplinary consensus panel endorsed by the *American Academy of Neurology* and *Child Neurology Society* (Filipek *et al.*, 1999) reviewed research on screening and diagnosis of ASD and made recommendations on practice parameters. The panel pointed out that fewer than 30% of primary care providers conduct regular standardized screening tests and recommended that primary care providers perform routine developmental screenings for ASD at each well-child visit with standardized tools that use parent report. Because there was no validated screening tool for ASD available, the panel suggested that failure to meet any of the following milestones is an absolute indication for immediate further evaluation: (a) no babbling by 12 months; (b) no gesturing by 12 months; (c) no single words by 16 months; (d) no 2-word spontaneous (not just echolalic) phrases by 24 months; and (e) ANY loss of ANY language or social skills at ANY age. However, these early indicators are not distinguishing of ASD, but would also be common early indicators of other DDs.

Thus, there is a constellation of social communication parameters that are important early indicators of ASD, particularly deficits in joint attention and symbolic communication. The lack of language and limitations in communication development may be among the first symptoms evident to parents and professionals. The early indicators for further evaluation identified by the consensus panel are in the

use of sounds, gestures, words, and word combinations, which may also be evident in children with DD who do not end up with ASD and in some late bloomers who catch up without intervention. Further research is needed to identify more precise “red flags” that distinguish children with ASD from other populations.

Communication and Symbolic Behavior Scales Developmental Profile

The Communication and Symbolic Behavior Scales Developmental Profile (CSBS DP, Wetherby & Prizant, 1998, 2002) is a standardized tool designed for screening and evaluation of communication and symbolic abilities of children between 12 and 24 months of age. The CSBS DP was recently developed based on the CSBS (Wetherby & Prizant, 1993), which is a more in-depth tool designed for program planning. The CSBS DP consists of three measures: (1) a one-page, 24-item Infant-Toddler Checklist for screening that can be completed quickly by a parent at a physician’s office or child-care center; (2) a four-page follow-up Caregiver Questionnaire; and (3) a Behavior Sample, which is a face-to-face evaluation of the child interacting with a parent and clinician that is videotaped for later analysis. The Checklist, Caregiver Questionnaire, and Behavior Sample were designed to measure seven prelinguistic skills organized into three composites: the Social composite, which includes Emotion and Eye Gaze, Communication (Rate and Function), and Gestures; the Speech composite, which includes Sounds and Words; and the Symbolic composite, which includes Understanding and Object Use. The CSBS DP has been nationally field-tested on 2188 children for the Checklist, 790 children for the Caregiver Questionnaire, and 337 children for the Behavior Sample (Wetherby & Prizant, 2002).

Wetherby and colleagues (Wetherby, Allen, Cleary, Kublin, & Goldstein, 2002; Wetherby, Goldstein, Cleary, Allen, & Kublin, 2003; Wetherby & Prizant, 2002) reported a series of studies to evaluate the validity and reliability of the three measures of the CSBS DP. These findings, which are summarized in the Method section, provide good evidence for reliability and validity and support the use of the Checklist as a first-level screening and the Behavior Sample as a second-level evaluation following the Checklist. It is important to point out that a small number of children with ASD may

have been included in those studies but most of the children had specific language delays or general DD and likely did not have ASD. Research is needed to investigate the effectiveness of the CSBS DP for screening and evaluation of children with ASD. The CSBS DP seems well suited for early identification of very young children with ASD because it measures prelinguistic skills that have been identified as deficits in preschoolers with ASD and should be evident in younger children.

The purpose of this study was to identify more precise early indicators of ASD during the second year of life. This study is part of an ongoing longitudinal, prospective investigation of the FIRST WORDS[®] Project to examine the relationship between prelinguistic communication measured with the CSBS DP (Wetherby & Prizant, 2002) in the second year of life and language outcomes in preschool children. Three groups of children were selected for this study, one group with communication delays who were later diagnosed with ASD, one group with communication delays in which ASD was ruled out, and one group with TD. The first objective of this study was to examine the usefulness of the CSBS DP Infant-Toddler Checklist and Behavior Sample for identifying children with ASD. The second objective of this study was to examine possible red flags for ASD from the videotapes of the Behavior Sample collected during the second year of life. We were interested in determining which red flags distinguished children with and without ASD and the accuracy of identification of ASD based on the red flags.

METHOD

Participants

Three groups of children participated in this study: one with ASD ($n = 18$), one with DD ($n = 18$), and one with TD; ($n = 18$). The participants in this study were drawn from a pool of children who are part of an ongoing longitudinal study of the FIRST WORDS[®] Project. Children were recruited to the Project by having parents complete the CSBS DP Infant-Toddler Checklist distributed by childcare and healthcare agencies and by staff in public places that serve families of young children. The target population for screening is children who have not yet been identified as having DD. This investigation is reporting on findings of participants

drawn from a pool of 3,021 children who were not previously identified (NPI) as having DD and were screened with the Infant-Toddler Checklist under 2 years of age. Additionally, five children who were developmentally delayed and had been identified (DDI) under the age of 2 years when referred to the Project, were included in this study, three who received a diagnosis of ASD from a pediatric neurologist and two children who had Down syndrome. These five DDI children were also screened with the Infant-Toddler Checklist under 2 years of age.

Figure 1 illustrates the steps involved in selecting subjects for this study from the pool of 3026 NPI and DDI children beginning with screening using the Infant-Toddler Checklist. The second step was to invite families for an evaluation, which included videotaping the CSBS DP Behavior Sample. The families of all children performing in the bottom 10th percentile (i.e., 1.25 SD below the mean) on the Checklist were invited for an evaluation, which included 377 NPI children or 12.48% of the 3021 NPI sample and the 5 DDI children. Additionally, 230 randomly selected children performing within normal limits on the Checklist were invited for an evaluation. Of the 612 families contacted, 122 could not be reached or did not agree to participate, leaving a pool of 490 NPI children. The third step was to invite the families of all children performing in the bottom 10th percentile and randomly selected children performing within normal limits on the Behavior Sample for a follow-up developmental evaluation using the *Mullen Scales of Early Learning* (MSEL, Mullen, 1995) between 2 and 3 years of age as a general measure of cognitive functioning. The MSEL was selected because it provides separate normative scores for four cognitive scales— Visual Reception, Fine Motor, Receptive Language, and Expressive Language. Of 445 families contacted, 147 children were lost from attrition, leaving a pool of 298 NPI children. At the time of the follow-up evaluation, the families completed an intervention history form inquiring about whether the child had been diagnosed with or was suspected of having DD, what type, and whether the child was receiving intervention. In addition to the 298 NPI children, all 5 DDI children completed the first three steps of screening and evaluation.

A group of children with communication delays was drawn from the pool of 303 children (i.e., 298 NPI and 5 DDI) who completed the first three steps of the Project and selected if they displayed a

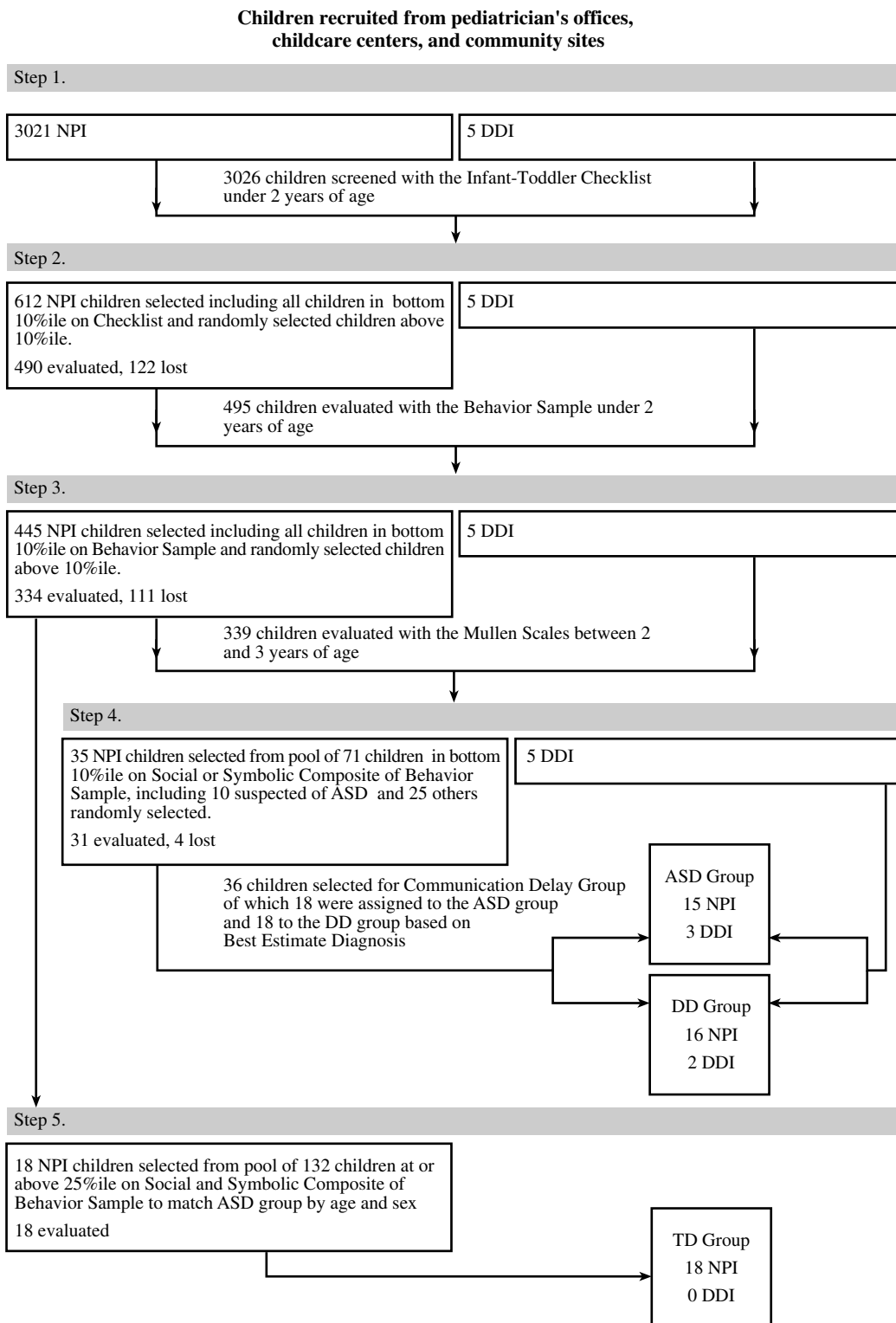


Fig. 1. Steps in the evaluation process for children not previously identified (NPI) and children with developmental delays who were previously identified (DDI) to find participants in the groups with autism spectrum disorders (ASD), developmental delays (DD), and typical development (TD).

communication delay during the second year of life based on performance in the bottom 10th percentile on either the Social or Symbolic Composite of the Behavior Sample. There were 76 children who met this criterion, 71 NPI children and the 5 DDI children. All 5 DDI children were contacted to participate in this study and all 5 agreed. Of the 71 NPI children, information provided by families on the intervention history form in the FIRST WORDS® Project archival database indicated that ASD was suspected in 10 of these children between 2 and 3 years of age and the families of these 10 children were contacted for participation in this study. All 10 agreed to participate, although one family had moved out of state. Of the remaining 61 NPI children, 25 children were randomly selected and their families were contacted to participate in this study; 4 families could not be found or refused. This resulted in 36 children with communication delays, 31 NPI and 5 DDI children, who were divided into two groups, ASD and DD, based on an evaluation to make a *best estimate diagnosis*, described below. The DD group included children with global DD and specific language delay.

Children in the TD group were drawn from the pool of 298 NPI children who completed the first three steps of the Project and selected if they displayed performance above the 25th percentile on the Social and Symbolic Composites of the CSBS DP Behavior Sample during the second year of life. Additionally, none of these children were suspected of having ASD by the clinician at the follow-up developmental evaluation or by the parent on the Information History Form. From a pool of 132 children meeting this criterion, 18 were selected to match the participants in the ASD group on sex and chronological age at the Behavior Sample. A summary of participant demographics for the ASD, DD, and TD groups are presented in Table I.

All children were 24 months of age or less when their families completed the Infant-Toddler Checklist. The age when the children were videotaped for the Behavior Sample is reported in Table I. The children in the ASD and TD groups were slightly older than the children in the DD group, however, this difference was not significant, $F(2, 51) = 2.61, p = .08$. Most of the children in all three groups were over 18 months of age at the time of the Behavior Sample. There were four children in the ASD group who were 18 months of age or younger, seven children in the DD group, and four children in the TD group. The groups were fairly

Table I. Summary of Participant Demographics

| Characteristic | ASD (n = 18) | DD (n = 18) | TD (n = 18) |
|--|-----------------|----------------|----------------|
| Child's chronological age at the behavior sample | | | |
| Mean age in months (M, SD) | 21.0 (3.5) | 18.4 (3.7) | 20.3 (3.2) |
| Range of age in months | 13.1–26.9 | 13.0–24.9 | 13.6–24.4 |
| Percentage of males | 88.9% | 83.3% | 88.9% |
| Percentage of first born | 27.8% | 33.3% | 22.2% |
| Referral source | | | |
| Public places/self referrals | 38.9% | 61.1% | 61.1% |
| Childcare providers | 0.0% | 5.6% | 11.1% |
| Healthcare providers | 61.1% | 33.3% | 27.8% |
| Race | | | |
| Caucasian | 83.3% | 77.8% | 94.4% |
| Other | 16.7% | 22.2% | 5.6% |
| Hispanic | 5.6% | 5.6% | 0.0% |
| Parent's education in years completed | | | |
| Mother (M, SD) | 14.7 (2.1) | 15.4 (2.5) | 15.5 (2.2) |
| Father (M, SD) | 15.8 (2.8) | 15.1 (2.6) | 15.4 (2.1) |
| Parent's age at child's birth in years | | | |
| Mother (M, SD) | 30.5 (5.8) | 32.7 (6.5) | 31.0 (4.9) |
| Father (M, SD) | 31.8 (5.8) | 35.9 (5.3) | 34.3 (4.9) |

similar with regard to most aspects of demographics. The ASD group had more children who were referred from healthcare providers than the other two groups, which may reflect their slightly older age and the providers' concerns about these children's development. Additionally, the TD group had a larger proportion of children who were Caucasian and none who were Hispanic.

Best Estimate Diagnosis

Families in the ASD and DD groups were contacted by phone when their child was between 30 months and 5 years of age to inquire if they were willing to bring their child back for follow-up to conduct a diagnostic evaluation for ASD. An interdisciplinary team consisting of a licensed speech-language pathologist and psychologist made a *best estimate diagnosis* based on information gathered from the following measures: (1) the most recent MSEL in the FIRST WORDS® Project archival database to determine nonverbal and verbal developmental level of functioning; (2) the *Vineland Adaptive Behavior Scales*, Interview Edition, Survey

Form (VABS, Sparrow, Balla, & Cicchetti, 1984) which yields a standard score in four domains—Communication, Daily Living, Social, and Motor, and an Adaptive Behavior Composite to index adaptive behavior/ developmental maturity; and (3) the *Autism Diagnostic Observation Schedule* (ADOS, Lord, Rutter, DiLavore, & Risi, 1999), which is a semi-structured, standardized assessment of communication, social interaction, and play or imaginative use of materials for children referred because of possible autism. The VABS was completed by phone interview if the families lived at a distance or preferred that to completing it face-to-face. The ADOS was administered in a small clinical room.

Using the information gathered from the diagnostic evaluation and provided by the parents, the team made a *best estimate diagnosis* based on the diagnostic criteria for Autistic Disorder or Pervasive Developmental Disorder- Not Otherwise Specified (PDD-NOS) defined in the DSM-IV (APA, 1994). The term *best estimate diagnosis* is used because the clinicians were allowed to make a judgment about how to put the information together to make a diagnosis. A diagnosis of Asperger syndrome was not made or ruled out due to the young age of these children. Children were assigned to the ASD group if they received a DSM IV diagnosis of Autistic Disorder or PDD-NOS and their Communication and Social Interaction Total score on the ADOS fell at or above the cut-off for autism spectrum; otherwise they were assigned to the DD group. There was only one child whose ADOS scores fell above the cut-off for autism spectrum but did not receive a diagnosis of PDD because the diagnostic team felt that his heightened scores were due to the severity of his DD. The ADOS could not be completed for two children, one unidentified and one identified child, because they lived at a distance and families could not return for the follow-up evaluation. Both of these children had received a diagnosis of ASD at over 30 months of age by a pediatric neurologist and had qualified for special education in the public schools, and therefore, were assigned to the ASD group. Of the 36 children with communication delays, 18 were assigned to the ASD group, including the 3 identified children, the 10 children suspected of ASD by community professionals and 5 children who were not suspected of having ASD prior to the research assessment. Of these 18 children, 9 were diagnosed with Autistic Disorder and 9 with PDD-NOS. The remaining 18 children were

assigned to the DD group, including the 2 identified children with Down syndrome.

Families in the TD group were mailed the *Social Communication Questionnaire*, Lifetime Version (SCQ, Rutter, Bailey, Berument, Lord, & Pickles, 2001), which is a 40-item parent report tool for diagnostic screening of ASD. The SCQ was selected because there is no validated screening tool for ASD in very young children and recent research indicates that the specificity and sensitivity of the SCQ is comparable for younger children based on 112 children between 2 and 5 years of age compared to 181 individuals over 5 years of age (Corsello, Cook & Leventhal, 2003). The SCQ yields a total score with 0 indicating no risk for ASD and 15 or higher indicating risk for ASD. If the families did not return the SCQ within three weeks, they were contacted by phone to remind them to return it and sent another copy of the questionnaire. The SCQ was returned by nine families which is half of the children in the TD group and these children received an average of 1.8 and range from 0 to 6.

The results of the MSEL for the ASD, DD and TD groups and the VABS for the ASD and DD groups are presented in Table II. Because so many of the children in this study achieved the lowest possible standard score on the MSEL (i.e., $T = 20$), a developmental quotient (DQ) was used based on age equivalent divided by chronological age multiplied by 100. A nonverbal DQ was calculated from the average of the Fine Motor and Visual Reception scales, and a verbal DQ was calculated from the average of the Receptive and Expressive Language scales. These scores indicate that the ASD and DD groups show a wide range of cognitive and adaptive functioning. On the MSEL, the ASD and DD groups did not differ on their nonverbal DQ, $t = -.34$, $p = .737$, or on their verbal DQ, $t = -1.63$, $p = .114$. Although the range of the nonverbal DQ was larger for the ASD group than the DD group, the variance and distribution of scores were similar; 9 of the 18 children in both the ASD and DD groups had a nonverbal DQ over 80. The range and variance of the verbal DQ was larger for the ASD group than the DD group and the distribution of scores was different; 7 children in the ASD group had a verbal DQ over 80 compared to 10 children in the DD group. On the VABS, the ASD and DD groups did not differ on the Communication standard score, $t = -.85$, $p = .403$, the Motor standard score, $t = .25$, $p = .803$, or the Adaptive Behavior Composite, $t = -1.94$, $p = .062$.

Table II. Summary of Follow-up Evaluation

| Characteristic | ASD (<i>n</i> = 18) | | | DD (<i>n</i> = 18) | | | TD (<i>n</i> = 18) | | |
|--|----------------------|-------------------|------------|---------------------|-------------------|-------------|---------------------|-----|------------|
| | M | SD | Range | M | SD | Range | M | SD | Range |
| Mullen scales of early learning ^a | | | | | | | | | |
| Chronological age (months) | 35.9 | 9.8 | 24.8–64.9 | 30.5 | 8.0 | 24.5–43.8 | 33.2 | 8.0 | 24.8–45.8 |
| Nonverbal developmental quotient | 78.0 | 21.2 | 37.8–125.2 | 80.3 | 20.1 | 49.4–108.01 | 104.9 | 9.8 | 87.8–116.2 |
| Verbal developmental quotient | 61.4 | 32.2 | 16.5–134.7 | 77.1 | 23.8 | 31.3–105.2 | 100.4 | 9.4 | 85.0–117.8 |
| Vineland adaptive behavior scales ^b | | | | | | | | | |
| Chronological age (months) | 47.7 | 15.4 | 30.9–70.7 | 43.4 | 11.7 | 30.0–66.8 | | | |
| Communication | 76.1 | 19.9 | 48–115 | 81.7 | 18.0 | 53–114 | | | |
| Daily living | 70.7 | 13.7 ^c | 49–93 | 84.1 | 21.8 ^c | 47–124 | | | |
| Social | 69.3 | 14.8 ^c | 50–110 | 88.2 | 19.2 ^c | 58–118 | | | |
| Motor | 76.7 | 14.5 | 52–101 | 78.6 | 25.9 | 31–116 | | | |
| Adaptive behavior composite | 68.3 | 12.9 | 51–94 | 80.6 | 21.53 | 51–123 | | | |

^aDevelopmental quotients based on age equivalent divided by chronological age multiplied by 100.

^bStandard scores based on a M of 100 and SD of 15.

^c $p < .05$ (t test, ASD and DD groups).

The ASD group had a significantly lower mean on the Daily Living standard score, $t = -2.06$, $p < .05$, and the Social standard score, $t = -3.11$, $p < .01$.

Procedure

Initial Measures from the Second Year of Life

As outlined in Fig. 1, families completed the CSBS DP Infant-Toddler Checklist at initial contact. The Checklist includes 24 questions with three to five choices about developmental milestones. It also asks if the family has any concerns about their child's development, and if so, to describe these concerns. The response to the question about parent concern was categorized for this study into concerns about language, social development and behavior, cognitive and motor development, and family history of developmental disabilities.

Families were contacted by phone to invite them to bring their child in for a face-to-face evaluation to conduct the Behavior Sample. The Behavior Sample was collected in a small clinical room using the standard sampling procedures and the standard Behavior Sample kit of materials (Wetherby & Prizant, 2002). A child's caregiver was present during the full evaluation and was instructed to respond naturally, but not to direct the child's behavior, in order to encourage spontaneous communication and play. The evaluation session began with a warm-up of about 10 minutes and lasted 30–40 minutes. The Behavior Sample uses a standard set of systematic procedures

designed to encourage spontaneous behavior that range in degree of structure provided. The child is first presented with a series of communicative temptations to entice spontaneous communication using a windup toy, balloon, bubbles, jar with food, bag with toys and books designed for young children. The child is then presented with a feeding toy set and stuffed animal to play symbolically and blocks to play constructively. The sample includes probes of gaze/point following and comprehension of person name, body part, and object name. The Behavior Sample was videotaped and scored using the standard CSBS DP procedures by one of five trained examiners who were blind to the child's diagnostic classification.

Standard scores and percentiles were available for the CSBS DP Infant-Toddler Checklist and Behavior Sample in the archival database of the FIRST WORDS[®] Project, which included the Social composite (sum of Emotion and Eye Gaze, Communication, and Gestures clusters), the Speech composite (sum of Sounds and Words clusters), and the Symbolic composite (sum of Understanding and Object Use clusters) using the norms for the CSBS DP First Normed Edition (Wetherby & Prizant, 2002). The standard scores were scaled to a mean of 10 and SD of 3 for the composites and a mean of 100 and SD of 15 for the Total score.

Information about the reliability and validity of the CSBS DP has been reported in Wetherby *et al.* (2002, 2003), and Wetherby and Prizant (2002). Based on the normative sample, all three measures of the CSBS DP were found to have a high degree

of internal consistency (α coefficients ranging from .86 to .92) and good test–retest reliability for standard scores over a 4-month interval, with significant increases in raw scores. Construct and concurrent validity has been supported by the developmental progression of scores from 6 to 24 months of age, intercorrelations among cluster and composite scores, and correlations between the parent report measures and the Behavior Sample. Wetherby *et al.* (2003) compared the accuracy of the Checklist to standardized testing on 232 children between 12 and 24 months of age, half with language delays and half with TD. Sensitivity was 87.4% and specificity was 75.2% using the bottom 10th percentile or 1.25 standard deviations below the mean as criterion for risk, which is comparable to or better than other instruments gathered with infants and toddlers using 2 standard deviations below the mean as risk criterion (e.g., Squires, Potter, & Bricker, 1999; Glascoe, 1999). Other analyses indicated that the three composites of the Checklist and Behavior Sample were a significant predictor of receptive and expressive language outcomes at 2 and 3 years of age and that the Behavior Sample explained a significant amount of unique variance in language outcomes beyond the Checklist. Thus, the CSBS DP Checklist and Behavior Sample are appropriate screening and evaluation tools for identifying children with DD at 12 to 24 months of age.

Reanalysis of the Behavior Sample Videotapes

Videotapes of the Behavior Sample were recoded for this study using the Systematic Observation of Red Flags (SORF) for Autism Spectrum Disorders in Young Children (Wetherby & Woods, 2002). Two raters blind to the diagnostic classification of the participants coded the SORF. The SORF was developed as a tool to decide if a referral for a diagnosis of ASD is needed based on the Behavior Sample of the CSBS DP. It includes 29 items derived from the diagnostic criteria for ASD (APA, 1994) and research on very young children with ASD (Adrien *et al.*, 1992; Baranek, 1999; Lord, Rutter, DiLavore, & Risi, 1999; Osterling & Dawson, 1994; Sheinkopf *et al.*, 2000; Wetherby *et al.*, 1998) grouped into the following five composite areas: Reciprocal Social Interaction, Unconventional Gestures, Unconventional Sounds and Words, Repetitive Behaviors and Restricted Interest, and Emotional Regulation. Definitions of each of the 29 items

are presented in the Appendix. The SORF rates the presence of behaviors that are *atypical* (i.e., rarely if ever displayed by typically developing children) and the absence of behaviors that are *typical* (i.e., usually displayed by typically developing children). Each item is rated as 0, 1, or 2, based on the child's behavior during the sample. For behaviors that are atypical, a score of 0 indicates absence of the behavior and a score of 2 indicates that the behavior is displayed often and/or in at least three different activities during the sample, with 1 indicating that the behavior is displayed one or a few times. For behaviors that are typical, a score of 0 indicates that the behavior is used frequently and/or in at least three different activities and a score of 2 indicates absence of the behavior during the sample. Thus, the score can range from 0 to 58, and a higher score indicates that more red flags of ASD were observed.

Interrater Reliability

Interrater reliability for the CSBS DP Behavior Sample was calculated using generalizability (g) coefficients for pairs of five independent raters on randomly selected videotapes of the Behavior Sample for at least 20% of the samples scored by each rater. A g coefficient approaches 1 as the variance accounted for by the subjects is large in comparison with the variance accounted for by raters. G coefficients that are at least .6 are considered acceptable for demonstrating inter-rater reliability (Mitchell, 1979). The g coefficients ranged from .92 to .97 for the composites and total, which indicate that CSBS DP raters exhibited high inter-rater reliability.

Interrater reliability for the SORF was calculated using the percent agreement between two raters who independently scored the videotapes of 12 randomly selected children, which is one-fifth of the sample. Both raters had extensive clinical experience and were familiar with the CSBS DP standard scoring and the ADOS. About five hours of training was needed to learn the SORF scoring and practice on videotapes of children who were not part of this study. Percent agreement was calculated for each item across children and for each child across the 29 items of the SORF. The mean percent agreement was 97.1%, and ranged from 89.7 to 100% across children and from 83.0 to 100% across items. Cohen's κ s were also calculated for each child across the 29 items of the SORF. The mean κ was .94 and ranged from .82 to 1.00.

Table III. Agreement of Classification for the CSBS DP Infant-Toddler Checklist Compared to Developmental Outcome and Best Estimate Diagnosis

| | ASD, DD & TD groups | | | ASD & TD groups | | | |
|---------------------------|---------------------|---------------------|----|--------------------|---------------|----|----|
| | Checklist | Communication delay | | Checklist | ASD diagnosis | | |
| | | Yes | No | | Yes | No | |
| Identified at risk | 32 | 2 | 34 | Identified at risk | 17 | 2 | 19 |
| No risk | 4 | 16 | 20 | No risk | 1 | 16 | 17 |
| | 36 | 18 | 54 | | 18 | 18 | 36 |
| Sensitivity | 88.9% | | | 94.4% | | | |
| Specificity | 88.9% | | | 88.9% | | | |
| Positive predictive value | 94.1% | | | 89.5% | | | |
| Negative predictive value | 80.0% | | | 94.1% | | | |

RESULTS

To explore the utility of the Infant-Toddler Checklist for early identification of ASD, classification of children based on Checklist scores was compared to classification based on best estimate diagnoses. Using the risk criteria established by Wetherby and Prizant (2002; i.e., “fail” the Checklist), 17 of the 18 children in the ASD group or 94.4% were at risk, 15 in the DD group or 83.3%, and 2 in the TD group or 11.1%. Agreement between classifications on the Checklist and developmental outcomes are presented in Table III for the following proportions: (1) *Sensitivity* — the proportion of children identified as at risk (i.e., receiving a positive screen or evaluation result) who failed the follow-up testing, which is also called *True Positives*; (2) *Specificity* — the proportion of children identified as no risk (i.e., receiving a negative screen or evaluation result) who passed the follow-up testing, which is also called *True Negatives*; (3) *Positive Predictive Value*— the proportion of children identified as at risk who failed the follow-up testing out of the total number of children identified as at risk.; and (4) *Negative Predictive Value*— the proportion of children identified as no risk who passed the follow-up testing out of the total number of children identified as no risk. Sensitivity, specificity, positive predictive value, and negative predictive value of the Infant-Toddler Checklist were at or above 80%, which is the recommended cutoff for screening tools (Meisels, 1989). Sensitivity was 88.9% when the ASD and DD groups were combined and increased to 94.4% when only the ASD group was examined with the TD group. Specificity was 88.9%.

To further explore the utility of the CSBS DP Infant-Toddler Checklist for early identification of ASD, the standard scores were compared for the three groups. The means and standard deviations on the social, speech, and symbolic composites and the total score are reported in Table IV. A series of one-way analyses of variances (ANOVA) were conducted to evaluate group differences on the composites and total. A Bonferonni adjustment of $.05/4 = .013$ was utilized to control for Type I error. Significant group differences were found for all three composites and the total. Eta square (η^2) is also reported in Table IV as the effect size index of the differences. Eta square is interpreted as the proportion of variance of the dependent variable that is a function of the factor; η^2 values of .01, .06, and .14 are considered small, medium and large effect sizes, respectively (Cohen, 1988). Large effect sizes were found for all three composites and the total, with differences in the social composite and total being the largest. It is noteworthy that 16 of the children in the ASD group were in the bottom 10th percentile on the Social composite of the Checklist, compared to 8 in the DD group and 2 in the TD group. Post hoc tests were conducted using Sheffé’s method to evaluate pairwise differences among means. There were no significant differences between the ASD and DD groups but there were significant differences between the ASD and TD groups on all three composites and the total. To examine whether or not group membership (ASD, DD, TD) could be correctly predicted from the Checklist composite scores, a discriminant analysis was conducted. The overall Wilks’ lambda was significant, $\Lambda = .57, \chi^2(6, n = 54) = 28.45, p < .001$,

Table IV. Mean CSBS DP Infant-Toddler Checklist Standard Scores

| CSBS DP Measure | ASD (<i>n</i> = 18) | | | DD (<i>n</i> = 18) | | | TD (<i>n</i> = 18) | | | <i>F</i> value | η^2 |
|---------------------------------------|----------------------|-----------|-------|---------------------|-----------|--------|---------------------|-----------|--------|----------------|----------|
| | M | <i>SD</i> | Range | M | <i>SD</i> | Range | M | <i>SD</i> | Range | | |
| Infant-Toddler Checklist ^a | | | | | | | | | | | |
| Social Composite | 4.00 | 1.94 | 3–10 | 7.11 | 3.61 | 3–17 | 10.44 | 3.90 | 4–17 | 17.39*** | .41 |
| Speech Composite | 5.22 | 1.77 | 3–8 | 6.39 | 1.79 | 3–9 | 8.56 | 3.35 | 4–17 | 8.82** | .26 |
| Symbolic Composite | 6.11 | 2.47 | 6–11 | 7.06 | 3.67 | 3–17 | 9.94 | 3.35 | 6–17 | 6.99** | .22 |
| Total | 71.56 | 7.21 | 65–88 | 80.39 | 11.03 | 65–106 | 97.00 | 16.89 | 72–135 | 19.64*** | .44 |
| Behavior Sample ^a | | | | | | | | | | | |
| Social Composite | 4.22 | 1.11 | 3–6 | 6.00 | 2.47 | 3–12 | 11.44 | 2.12 | 8–15 | 64.50*** | .72 |
| Speech Composite | 5.28 | 1.90 | 3–8 | 6.78 | 2.60 | 3–11 | 10.17 | 1.69 | 7–14 | 25.57*** | .50 |
| Symbolic Composite | 4.83 | 2.53 | 3–10 | 5.28 | 1.78 | 3–11 | 10.83 | 2.07 | 8–14 | 43.73*** | .63 |
| Total | 69.33 | 4.79 | 65–78 | 75.17 | 10.60 | 65–100 | 102.00 | 9.40 | 85–128 | 73.32*** | .74 |

^a Standard Scores based on a M of 10 and *SD* of 3 for the Composites and a M of 100 and *SD* of 15 for the Total.
 * *p* < 0.01, ** *p* < 0.005, *** *p* < 0.001.

indicating that overall the predictors differentiated among the three groups. As shown in Table V, only 66.7% of the cases were correctly classified based on the Checklist composites. The sensitivity and specificity of the Infant-Toddler Checklist support its use as a first stage screener for children with communication disorders, including ASD. Although low social composites were found more often in the ASD group than the DD or TD group, the results of the discriminant analysis indicate that the Infant-Toddler Checklist does not distinguish children with ASD from children with DD with high accuracy.

To further explore the utility of the CSBS DP for early identification of ASD, similar analyses were conducted to examine differences between groups in the pattern of scores on the Behavior Sample using the standard scoring. The means and standard deviations on the three composites and the total are reported on the bottom of Table IV. Similar to the Checklist, large effect sizes were found for all three Behavior Sample composites and the total, with differences in the social composite and total being the largest. Post hoc tests indicated that there were no significant differences between the ASD and DD groups but there were significant differences between the ASD and TD groups and the DD and TD groups on all three composites and the total. It is noteworthy that all 18 of the children in the ASD group were in the bottom 10th percentile on the Social composite of the Behavior Sample, compared to 12 in the DD group and 0 in the TD group. Furthermore, 16 of the children in the ASD group were in the bottom 5th percentile on the Social composite of the Behavior Sample, compared

to 9 in the DD group. A discriminant analysis was conducted using the seven cluster scores of the Behavior Sample and the overall Wilks' lambda was significant, $\Lambda = .19$, $\chi^2 (14, n = 54) = 80.28$, *p* < .001, indicating that overall the predictors differentiated among the three groups. As shown in Table V, 81.5% of the cases were correctly classified based on the Behavior Sample clusters. The pattern of scores on the Behavior Sample supports its use as a second-level evaluation for children with communication disorders; however, it would not lead to precise classifications of ASD and DD.

Table V. Frequency and Percentage of Subjects Classified Correctly from Discriminant Analysis of the CSBS DP Infant-Toddler Checklist and Behavior Sample

| Actual group membership | Predicted group membership | | | | | |
|---------------------------------------|----------------------------|------|----------|------|----------|------|
| | ASD | | DD | | TD | |
| | <i>n</i> | % | <i>n</i> | % | <i>n</i> | % |
| Infant-Toddler checklist ^a | | | | | | |
| ASD (<i>n</i> = 18) | 16 | 88.9 | 2 | 11.1 | 0 | 0 |
| DD (<i>n</i> = 18) | 7 | 38.9 | 8 | 44.4 | 3 | 16.7 |
| TD (<i>n</i> = 18) | 2 | 11.1 | 4 | 22.2 | 12 | 66.7 |
| Behavior sample ^b | | | | | | |
| ASD (<i>n</i> = 18) | 13 | 72.2 | 5 | 27.8 | 0 | 0 |
| DD (<i>n</i> = 18) | 5 | 27.8 | 13 | 72.2 | 0 | 0 |
| TD (<i>n</i> = 18) | 0 | 0 | 0 | 0 | 18 | 100 |

^a66.7% of original grouped cases correctly classified based on checklist composites.

^b81.5% of original grouped cases correctly classified based on behavior sample clusters.

Table VI. Parent Concern at the Time of the CSBS DP Behavior Sample

| | ASD | DD | TD |
|------------------------------------|-----|-----|-----|
| Sample size | 18 | 18 | 18 |
| Parents with concern | 94% | 72% | 61% |
| Type of concern | | | |
| Language | 83% | 56% | 50% |
| Social or behavior | 72% | 11% | 28% |
| Cognitive or motor | 33% | 22% | 6% |
| Family history | 17% | 6% | 6% |
| Concerns in more than one category | 78% | 22% | 17% |

The results of the parents' response to the question of whether they were concerned about their child's development on the CSBS DP Infant-Toddler Checklist are reported in Table VI. All but one parent of a child in the ASD group expressed concerns about their child's development, while about two-thirds of the parents in the DD group expressed concerns. However, 61% of the parents in the TD group also expressed concerns, even though their child's performance during the Behavior Sample was within normal limits. The percentage of parents who expressed concerns was significantly greater for the ASD group than the TD group ($\chi^2 = 5.79, p = .016$) but did not differ significantly for the ASD and DD groups ($\chi^2 = 3.20, p = .074$). Most of the parent concerns for all 3 groups were about language skills, however, many of the parents of children in the ASD group expressed concerns about social development (e.g., doesn't look at me, doesn't smile or laugh, not interested in interacting with people) or behavior (e.g., behavior hard to manage, difficulty paying attention). More than three-fourths of the parents of children in the ASD group reported concerns in more than one category (e.g., concerns about language and behavior), compared to less than one-fourth in the DD and TD groups. The percentage of parents who expressed concerns in more than one category was significantly greater for the ASD group than for either the DD group ($\chi^2 = 11.11, p = .001$) or the TD group ($\chi^2 = 13.49, p < .001$) and did not differ significantly for the DD and TD groups ($\chi^2 = 0.18, p = .674$).

The SORF was used to examine more precise red flags for ASD in the second year of life, and includes behaviors that are measured on the CSBS DP Behavior Sample and additional typical behaviors as well as atypical behaviors. A series of one-way ANOVA were conducted to evaluate group

differences on the SORF composites and total. A Bonferonni adjustment of $.05/6 = .008$ was utilized to control for Type I error. Significant group differences were found for Reciprocal Social Interaction, $F(2, 51) = 50.14, p < .001$, Unconventional Gestures, $F(2, 51) = 42.27, p < .001$, Unconventional Sounds and Words, $F(2, 51) = 14.58, p < .001$, Repetitive Behaviors and Restricted Interests, $F(2, 51) = 33.40, p < .001$, Emotional Regulation, $F(2, 51) = 8.78, p < .005$, and the total, $F(2, 51) = 52.66, p < .001$. Post hoc analyses using Sheffé's method indicated significant differences between the ASD and DD groups on three composites, Reciprocal Social Interaction, Unconventional Gestures, and Repetitive Behaviors and Restricted Interests, and on the total. There were significant differences between the ASD and TD groups on all five composites and the total.

The next step in analyses was to examine differences between the groups on individual items of the SORF. The means, standard deviations, and percentage of participants in each group who had a SORF score greater than 0 (i.e., 1 or 2) on each item are presented in Table VII. A series of ANOVA were conducted to evaluate the differences between the three groups on the SORF items and the results are presented in Table VII. A Bonferonni adjustment of $.05/29 = .002$ was utilized to control for Type I error. Significant group differences (i.e., $p < .001$) were found on 13 items and marginal or nonsignificant trends ($p < .005, .01$ or $.05$) on an additional 6 items.

Post hoc tests were conducted to evaluate pairwise differences among means on the 13 SORF items with significant group differences using Sheffé's method. There were significant differences between the ASD and DD groups and the ASD and TD groups on the following nine items: (1) lack of appropriate gaze; (2) lack of warm, joyful expressions with gaze; (3) lack of sharing enjoyment or interest; (4) lack of response to name; (5) lack of coordination of gaze, facial expression, gesture, and sound; (6) lack of showing; (7) unusual prosody; (8) repetitive movements or posturing of body, arms, hands, or fingers; and (9) repetitive movements with objects. There were significant differences between the ASD and TD groups but not the ASD and DD groups on the following four items: (1) lack of response to contextual cues; (2) lack of pointing; (3) lack of vocalizations with consonants; and (4) lack of playing with a variety of toys conventionally.

To address the question of whether or not group membership (ASD, DD, TD) could be correctly pre-

Table VII. Group Comparison of Items on the SORF

| SORF items | ASD (<i>n</i> = 18) | | | DD (<i>n</i> = 18) | | | TD (<i>n</i> = 18) | | | <i>F</i> value | η^2 |
|--|----------------------|-----------|------|---------------------|-----------|-----|---------------------|-----------|-----|-------------------|----------|
| | <i>M</i> | <i>SD</i> | %>0 | <i>M</i> | <i>SD</i> | %>0 | <i>M</i> | <i>SD</i> | %>0 | | |
| <i>Difficulty with Reciprocal Social Interaction</i> | | | | | | | | | | | |
| 1. Aversion to social touch or proximity | 0.00 | 0.00 | 0% | 0.00 | 0.00 | 0% | 0.00 | 0.00 | 0% | 0.00 | .00 |
| 2. Lack of appropriate gaze | 1.33 | 0.91 | 72% | 0.28 | 0.57 | 22% | 0.01 | 0.24 | 6% | 20.81*** | .45 |
| 3. Lack of warm, joyful expressions with gaze | 1.33 | 0.77 | 83% | 0.50 | 0.79 | 33% | 0.01 | 0.24 | 6% | 18.01*** | .41 |
| 4. Lack of sharing enjoyment or interest | 1.78 | 0.55 | 94% | 0.89 | 0.90 | 56% | 0.01 | 0.24 | 6% | 34.33*** | .57 |
| 5. Lack of anticipatory posture or movement | 0.61 | 0.70 | 50% | 0.22 | 0.43 | 22% | 0.01 | 0.24 | 6% | 6.05** | .19 |
| 6. Lack of response to contextual cues | 1.39 | 0.70 | 89% | 0.94 | 0.80 | 67% | 0.01 | 0.24 | 6% | 20.98*** | .45 |
| 7. Lack of response to name when called | 1.94 | 0.24 | 100% | 1.50 | 0.62 | 94% | 0.72 | 0.75 | 56% | 20.60*** | .45 |
| 8. Lack of coordination of gaze, facial expression, gesture, and sound | 1.83 | 0.38 | 100% | 0.89 | 0.83 | 61% | 0.17 | 0.38 | 17% | 38.22*** | .60 |
| <i>Unconventional Gestures</i> | | | | | | | | | | | |
| 1. Using person's hand as a tool without gaze | 0.11 | 0.32 | 11% | 0.11 | 0.32 | 11% | 0.00 | 0.00 | 0% | 1.06 | .04 |
| 2. Lack of pointing | 1.78 | 0.55 | 94% | 1.33 | 0.84 | 78% | 0.17 | 0.38 | 17% | 32.41*** | .56 |
| 3. Lack of showing | 1.83 | 0.38 | 100% | 0.89 | 0.83 | 61% | 0.11 | 0.32 | 11% | 42.53*** | .63 |
| <i>Unconventional Sounds and Words</i> | | | | | | | | | | | |
| 1. Atypical vocalizations | 0.17 | 0.51 | 11% | 0.11 | 0.32 | 11% | 0.01 | 0.24 | 6% | 0.39 | .02 |
| 2. Unusual syllable strings | 0.33 | 0.69 | 22% | 0.00 | 0.00 | 0% | 0.00 | 0.00 | 0% | 4.25 [†] | .14 |
| 3. Unusual prosody | 0.89 | 0.96 | 50% | 0.00 | 0.00 | 0% | 0.00 | 0.00 | 0% | 15.32*** | .38 |
| 4. Immediate echolalia | 0.00 | 0.00 | 0% | 0.00 | 0.00 | 0% | 0.00 | 0.00 | 0% | 0.00 | .00 |
| 5. Idiosyncratic or repetitive use of words | 0.01 | 0.24 | 6% | 0.00 | 0.00 | 0% | 0.00 | 0.00 | 0% | 1.00 | .04 |
| 6. Lack of vocalizations with consonants | 1.33 | 0.77 | 83% | 1.17 | 0.86 | 72% | 0.17 | 0.38 | 17% | 14.62*** | .36 |
| <i>Repetitive Behaviors and Restricted Interests</i> | | | | | | | | | | | |
| 1. Repetitive movements or posturing of body, arms, hands, or fingers | 0.78 | 0.88 | 50% | 0.17 | 0.38 | 17% | 0.00 | 0.00 | 0% | 9.86*** | .28 |
| 2. Repetitive movements with objects | 1.06 | 0.80 | 72% | 0.11 | 0.32 | 11% | 0.00 | 0.00 | 0% | 24.28*** | .49 |
| 3. Unusual sensory exploration of objects | 0.39 | 0.70 | 28% | 0.01 | 0.24 | 6% | 0.00 | 0.00 | 0% | 4.40 [†] | .15 |
| 4. Excessive interest in particular toys | 0.39 | 0.70 | 28% | 0.01 | 0.24 | 6% | 0.00 | 0.00 | 0% | 4.40 [†] | .15 |
| 5. Lack of playing with a variety of toys | 1.39 | 0.78 | 83% | 0.94 | 0.64 | 78% | 0.00 | 0.00 | 0% | 26.82*** | .51 |
| <i>Emotional Regulation</i> | | | | | | | | | | | |
| 1. Fear or distress about particular objects | 0.33 | 0.69 | 22% | 0.11 | 0.32 | 11% | 0.11 | 0.32 | 11% | 1.31 | .05 |
| 2. Distress over removing particular objects or ending an activity | 0.33 | 0.69 | 22% | 0.01 | 0.24 | 6% | 0.00 | 0.00 | 0% | 3.27 [†] | .11 |
| 3. Difficulty calming when distressed | 0.50 | 0.79 | 33% | 0.01 | 0.24 | 6% | 0.00 | 0.00 | 0% | 6.02** | .19 |
| 4. Abrupt shifts in emotional or behavioral state | 0.33 | 0.69 | 22% | 0.11 | 0.32 | 11% | 0.00 | 0.00 | 0% | 2.71 | .10 |
| 5. Heightened alertness and response to stimuli or situations | 0.39 | 0.78 | 22% | 0.28 | 0.67 | 17% | 0.00 | 0.00 | 0% | 2.06 | .08 |
| 6. Flat affect or unresponsive to interactions | 0.22 | 0.55 | 17% | 0.01 | 0.24 | 6% | 0.00 | 0.00 | 0% | 2.03 | .07 |
| 7. Challenging behavior | 0.01 | 0.24 | 6% | 0.00 | 0.00 | 0% | 0.00 | 0.00 | 0% | 1.00 | .04 |

* $p < 0.01$; ** $p < 0.005$, *** $p < 0.001$; [†]nonsignificant trend $p < 0.05$.

dicted from the SORF scores, a discriminant analysis was conducted using the 13 items that the ASD group was significantly different from the DD or TD groups. The discriminant analysis demonstrated significant results. Two functions emerged; the first function accounted for 80% of the variance and the second function accounted for the remaining 20% of the variance. The overall Wilks' lambda was significant, $\Lambda = .07$, $\chi^2(30, n = 54) = 119.04$, $p < .001$,

indicating that overall the predictors differentiated among the three groups. In addition, the residual Wilks' lambda was significant, $\Lambda = .42$, $\chi^2(14, n = 54) = 37.82$, $p < .01$, indicating that the predictors differentiated significantly among the three groups after partialling out the effects of the first discriminant function. The correlation coefficients of each predictor with each discriminant function are shown in Table VIII. The predictors most strongly

Table VIII. Correlation Coefficients of Predictor Variables with the Two Discriminant Functions

| SORF items | Correlation coefficients with discriminant functions | |
|---|--|------------|
| | Function 1 | Function 2 |
| Lack of showing | 0.55 | 0.24 |
| Lack of coordination of gaze, facial expression, gesture, and sound | 0.52 | 0.20 |
| Lack of sharing enjoyment or interest | 0.49 | 0.25 |
| Repetitive movements with objects | 0.42 | -0.13 |
| Lack of appropriate gaze | 0.39 | -0.67 |
| Lack of warm, joyful expressions with gaze | 0.36 | 0.07 |
| Lack of response to name when called | 0.35 | 0.32 |
| Lack of response to contextual cues | 0.35 | 0.35 |
| Unusual prosody | 0.33 | -0.17 |
| Repetitive movements or posturing of body, arms, hands, or fingers | 0.27 | -0.02 |
| Lack of anticipatory posture or movement | 0.21 | 0.02 |
| Difficulty calming when distressed | 0.21 | -0.06 |
| Lack of pointing | 0.42 | 0.49 |
| Lack of playing with a variety of toys | 0.39 | 0.40 |
| Lack of vocalizations with consonants | 0.26 | 0.40 |

correlated with the first discriminant function appear to discriminate the children with ASD from the other two groups while those most strongly correlated with the second function discriminate the children with ASD and DD from the TD group. As shown in Table IX, when group classification was predicted based on the 13 red flags, 94.4% of the cases were correctly classified. Of particular importance, 100% of the chil-

Table IX. Frequency and Percentage of Subjects Classified Correctly from Discriminant Analysis based on 13 SORF items^a

| Actual group membership | Predicted group membership | | | | | |
|-------------------------|----------------------------|-----|----------|------|----------|------|
| | ASD | | DD | | TD | |
| | <i>n</i> | % | <i>n</i> | % | <i>n</i> | % |
| ASD (<i>n</i> = 18) | 18 | 100 | 0 | 0 | 0 | 0 |
| DD (<i>n</i> = 18) | 0 | 0 | 15 | 83.3 | 3 | 16.7 |
| TD (<i>n</i> = 18) | 0 | 0 | 0 | 0 | 18 | 100 |

^a94.4% of original grouped cases correctly classified.

dren in the ASD group, 83% in the DD group, and 100% in the TD group were correctly predicted. Three children from the DD group were classified as TD. Although discriminant analysis on samples this small should be interpreted with caution, the high percentage of agreement in reclassification suggests that these groups differed substantially on these 13 red flags and that these SORF items distinguished these children with ASD and DD with a high level of accuracy. These findings should also be interpreted with caution due to the lack of an independent validation sample.

DISCUSSION

The CSBS DP Infant-Toddler Checklist was designed for routine screening for communication delays and consists of 24 questions about typical communication milestones and a question about parent concern. The Checklist is brief and can be given to a parent of any child, whether the parent has a concern about the child’s development or about possible autism or has no concerns. The findings of this study demonstrate that the Infant-Toddler Checklist was an effective tool for identifying children with ASD as having a communication delay in a first-level screening. The Checklist identified 17 of the 18 children in the ASD group or 94.4% and 15 in the DD group or 83.3%. Although we cannot calculate the true sensitivity of the Checklist, the sensitivity in this sample for identifying these children with communication delays with or without ASD was 89%, which is similar to the findings of Wetherby *et al.* (2003) with a larger and more heterogeneous sample of children with DDs. The sensitivity for identifying children with ASD was higher (94%) than that for identifying children with communication delays without ASD (83%) indicating that the Checklist does an even better job sorting out children with ASD from typical children than children with communication delays without ASD. Only 2 of the 18 children in the TD group or 11.1% were identified as at risk on the Checklist, indicating a specificity of 89%, which was higher than that found in the larger sample (75%) studied by Wetherby *et al.* (2003). The larger sample included many children with milder language delays than the children in this sample, which likely contributed to the lower specificity and corresponding higher false positive rate. More children in the ASD group performed in the bottom 10th percen-

tile on the Checklist Social composite than in the DD group, however, the Checklist did not distinguish the ASD and DD children with high accuracy. The Checklist results from the three groups were consistent with the pattern of scores obtained on the Behavior Sample and indicate that parents of children with ASD are accurate reporters of early social communication milestones on the Infant-Toddler Checklist.

Healthcare providers are in a pivotal role to detect communication problems earlier in young children by conducting developmental surveillance ideally on all children. These findings add to the growing body of research documenting the effectiveness of parent report as a screening tool for young children. There is wide variation in the age that children begin talking and the rate that children learn to talk. This makes it difficult to decide when to be concerned if a child is not talking. The Checklist includes questions about prelinguistic predictors of language, including the use of eye gaze, gestures and sounds to communicate and the ability to understand words and to play with objects, which provide important clues about the development of language. Using a parent report tool, such as the Checklist, minimizes the time required of healthcare providers, maximizes the role of the family, and provides reasonably accurate information about whether to refer a child for a developmental evaluation. Unlike the CHAT or M-CHAT, the Infant-Toddler Checklist is not designed to screen specifically for ASD, but rather, is designed as a first-level screen for children with a broad array of communication delays. The findings of this study do suggest that children with ASD are likely to have low scores on the Social composite of the Checklist and this pattern could be used to indicate the need to conduct an autism-specific screen next, such as the CHAT or M-CHAT. However, there are not yet sufficient validity data on the CHAT, M-CHAT, or any other parent report tool to support their use as a second-level screen for ASD in the second year of life, and therefore, further research is needed.

One limitation of this study is that we do not have an accurate estimate of how many of the 3021 children screened with the Infant-Toddler Checklist who were NPI actually have ASD. Of these 3021 children, we were able to find 15 children who received a best estimate diagnosis of ASD in this study. Of those 15, 14 failed the Checklist. We do not know how many children we missed who may have ASD. Current prevalence estimates for ASD

range from 3.4 (Yeargin-Allsopp *et al.*, 2003) to 6 (Chakrabarti & Fombonne, 2001) per thousand. Of the sample of 3021 children screened who were NPI, 12.5% were identified as at-risk on the Checklist, using a criterion of the bottom 10th percentile. Therefore, this sample included a slightly higher proportion of children who failed than would be expected in the general population, which may reflect more children being screened whose doctor or parent had concerns about the child. With this in mind, 5 children per thousand screened received a *best estimate diagnosis* of ASD in this study, which is in the middle of the range of prevalence estimates. Future research will be directed at following up with the larger sample into school age to address the question of how many of the children screened end up with ASD or another developmental disability.

Families are often the first to raise concerns about their child's development. Previous research on parent concerns of children with various types of developmental disabilities has suggested that concerns raised by the majority of families are warranted (Glascoe, 1999). However, there is limited research on parent concerns of children in the first or second year of life. The majority of parents in all three groups in this study had concerns about their child's communication development, at least when asked in the context of a checklist about communication milestones. A larger percentage of the parents of children with ASD had concerns, and most had concerns about more than one area of development, usually language and social or behavior, similar to findings in other studies (Howlin & Moore, 1997; Siegel *et al.*, 1988). It is important to conduct a developmental screening for any child whose family has any concern about his/her development. For families who have concerns but their child is developing typically, it is important to reassure them and provide information about developmental milestones and their child's development. For children who are delayed but families are not yet concerned, it can be difficult for parents to learn that their child is not developing as expected, and particularly that their child might have ASD. Healthcare professionals can play a critical role in early identification by listening to families concerns, conducting a developmental screening, and making referrals for a developmental evaluation so that families access intervention earlier.

The CSBS DP Behavior Sample was designed as a second-level evaluation to determine whether a child has a communication delay. In this study, the

standard scoring of the Behavior Sample was more accurate in detecting communication problems than the Infant-Toddler Checklist. This finding is expected because the participants were selected based on their performance on the Behavior Sample and because face-to-face evaluations have been found to be more accurate than parent report (Wetherby *et al.*, 2002, 2003). Previous research suggests that the Behavior Sample offers sufficient improvement in sensitivity and specificity to warrant use by clinicians as a follow-up evaluation to the Infant-Toddler Checklist (Wetherby *et al.*, 2002, 2003). However, although more accurate than the Checklist, the standard scoring on the Behavior Sample does not distinguish children with ASD and DD with a high level of accuracy.

This longitudinal study offered a unique opportunity to examine possible red flags in the second year of life using the SORF scoring from the videotapes of the Behavior Samples of children who were later diagnosed with ASD, compared to the children in the DD and TD groups. The SORF scoring provides more detailed information than the standard scoring of the Behavior Sample by measuring both the lack of typical behaviors and the presence of atypical behaviors that have been associated with ASD in previous research on older and younger children with ASD. The SORF results indicated that the following nine red flags differentiated children in the ASD group from children in both the DD and TD groups: (1) lack of appropriate gaze; (2) lack of warm, joyful expressions with gaze; (3) lack of sharing enjoyment or interest; (4) lack of response to name; (5) lack of coordination of gaze, facial expression, gesture, and sound; (6) lack of showing; (7) unusual prosody; (8) repetitive movements or posturing of body, arms, hands, or fingers; and (9) repetitive movements with objects. The first six of these nine red flags involve a lack of typical behaviors and the last three involve the presence of atypical behaviors. The lack of each of these six typical behaviors was evident in a large majority of these children with ASD, ranging from 72 to 100%. Unusual prosody and repetitive movements or posturing of body, arms, hands, or fingers were displayed by half of the children in the ASD group and repetitive movements with objects was displayed by about three-fourths of the children with ASD. In contrast, only a small proportion or no children in the DD group displayed these atypical behaviors. These findings on children with ASD in the second year of life are consistent with previous

research on 2-year-olds which has shown that the social communication impairments are prominent at young ages and that repetitive behaviors are more variable but are evident in observations of many young children with ASD (DiLavore, Lord & Rutter, 1995; Stone *et al.*, 1999).

The SORF results indicated that the following 4 red flags differentiated children in the ASD group from children in the TD group but not from children in the DD group: (1) lack of response to contextual cues; (2) lack of pointing; (3) lack of vocalizations with consonants; and (4) lack of playing with a variety of toys conventionally. In other words, children in both the ASD and DD groups displayed these red flags and children in the TD group did not. These four red flags involve a lack of development of typical behaviors and likely reflect the severity of a child's cognitive delay, whether the child has ASD or not.

The findings from this study give us a glimpse of the warning signs of ASD in the second year of life. Some of these warning signs are also seen in other children with DD. Young children with ASD are likely to be delayed in using words and their vocalizations are likely to lack consonants and to have unusual prosody. They are likely not to respond to their name or to instructions even with contextual cues. They are likely to be delayed in using objects conventionally in play and also are likely to display repetitive movements with their body or objects. Furthermore, young children with ASD are very likely to be delayed in sharing attention with eye gaze, sharing affect, and drawing others attention to objects or events of interest. They are very likely to lack the indicating gestures of pointing and showing and lack the coordination of gestures with eye gaze, facial expression, or vocalizations. We do not yet have a sufficient sample size to determine cutoff scores to use the SORF for evaluation decisions, however, the results of this study suggest that children showing most of these 13 red flags should be suspected of having ASD and should be referred for a diagnostic evaluation. Deciding to refer a young child for a diagnostic evaluation using an autism-specific tool, such as the ADOS and informing the parent of this need are very difficult and delicate clinical responsibilities, particularly for children under 2 years of age. The Behavior Sample offers a valuable evaluation context to examine a young child with the parent present, inform the parent about the child's language or social communication delay, and then

decide about the referral for a diagnostic evaluation. If the Behavior Sample is videotaped, a clinician can decide after the sample is collected to use the SORF scoring. The results of this study suggest that children who score in the bottom 5th percentile on the Social composite using the standard scoring should be scored with the SORF and that children who score between the 6th and 9th percentiles may be considered for the SORF scoring. Clinicians who are not familiar with the ADOS will likely require more training than the raters in this study to learn to score the SORF reliably, which may limit clinical feasibility.

The results of this study extend findings from other studies on samples of very young children with ASD. They are fairly consistent with the findings from home videotapes at 1 year (Osterling and Dawson, 1994, 1999) but provide more detail. One difference was that we found that lack of showing and repetitive movements did differentiate the children in the ASD and DD group. Consistent with Osterling and Dawson (1999), we found that lack of pointing did not differentiate children in the ASD and DD groups, but did differentiate both with children in the TD group. These findings are very consistent with the findings of Lord (1995) on 24 month olds with ASD based on parent report and DiLavore *et al.* (1995) based on sampling and observation. Further research is needed to study the ontogeny of red flags for ASD over the first three years of life. Many of the red flags that were not evident in any or very many children in the second year of life, are likely to be prominent in the third year of life (e.g., using person's hand as a tool, immediate echolalia). The findings of this study provide information about items that would be important to include on an autism-specific second level screening. However, it is not known whether the red flags identified in the CSBS DP Behavior Sample, which uses a naturalistic but structured sampling procedure, would also be evident in unstructured observations at home, in childcare settings, or in pediatrician's offices during well child visits. Further research is needed to document second-level screening and evaluation tools for ASD in very young children.

REFERENCES

- Adrien, J.L., Barthelemy, C., Perrot, A., Roux, S., Lenoir, P., & Hameury, L., *et al.* (1992). Validity and reliability of the Infant Behavioral Summarized Evaluation (IBSE): A rating scale for the assessment of young children with autism and developmental disorders. *Journal of Autism and Developmental Disorders*, *22*, 375–394.
- American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.), Washington, DC: APA.
- Baird, G., Charman, T., Baron-Cohen, S., Cox, A., Swettenham, J., & Wheelwright, S., *et al.* (2000). A screening instrument for autism at 18 months of age: A 6-year follow-up study. *Journal of the American Academy of Child Adolescent Psychiatry*, *39*, 694–702.
- Baranek, G. T. (1999). Autism during infancy: A retrospective video analysis of sensory-motor and social behaviors at 9–12 months. *Journal of Autism and Developmental Disorders*, *29*, 213–224.
- Baron-Cohen, S., Allen, J., & Gillberg, C. (1992). Can autism be detected at 18 months? The needle, the haystack, and the CHAT. *British Journal of Psychiatry*, *161*, 839–843.
- Baron-Cohen, S., Cox, A., Baird, G., Swettenham, J., Nightingale, N., & Morgan, K., *et al.* (1996). Psychological markers in the detection of autism in infancy in a large population. *British Journal of Psychiatry*, *168*, 158–163.
- Chakrabarti, S., & Fombonne, E. (2001). Pervasive developmental disorders in preschool children. *Journal of the American Medical Association*, *285*, 3093–3099.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.
- Corsello, C., Cook, E., & Leventhal, B. (2003). A screening instrument for autistic spectrum disorders. Poster presentation at the Society for Research in Child Development Biennial Conference, Tampa, Florida, April, 2003.
- Dawson, G., & Adams, A. (1984). Imitation and social responsiveness in autistic children. *Journal of Abnormal Child Psychology*, *12*, 209–226.
- Dawson, G., & Osterling, J. (1997). Early intervention in autism. In M. Guralnick (Ed.), *The effectiveness of early intervention*. Baltimore, MD: Paul H. Brookes.
- DiLavore, P., Lord, C., & Rutter, M. (1995). The Pre-Linguistic Autism Diagnostic Observation Schedule. *Journal of Autism and Developmental Disorders*, *25*, 355–379.
- Fenske, E., Zalski, S., Krantz, P., & McClannahan, L. (1985). Age at intervention and treatment outcome for autistic children in a comprehensive intervention program. *Analysis and Intervention for Developmental Disabilities*, *5*, 49–58.
- Filipek, P., Accardo, P., Baranek, G., Cook, E., Dawson, G., Gordon, B., Gravel, J., Johnson, C., Kallen, R., Levy, S., Minshew, N., Prizant, B., Rapin, I., Rogers, S., Stone, W., Teplin, S., Tuchman, R., & Volkmar, F. (1999). The screening and diagnosis of autistic spectrum disorders. *Journal of Autism and Developmental Disorders*, *29*, 439–484.
- Glascoe, F. P. (1999). The value of parents' concerns to detect and address developmental and behavioural problems. *Journal of Paediatric Child Health*, *35*, 1–8.
- Harris, S., & Handleman, J. (2000). Age and IQ at intake as predictors of placement for young children with autism: A four- to six-year follow-up. *Journal of Autism and Developmental Disorders*, *30*, 137–142.
- Howlin, P., & Moore, A. (1997). Diagnosis of autism: A survey of over 1200 patients in the UK. *Autism*, *1*, 135–162.
- Lord, C. (1995). Follow-up of two year-olds referred for possible autism. *Journal of Child Psychology and Psychiatry*, *36*, 1365–1382.
- Lord, C., Rutter, M., DiLavore, P., & Risi, S. (1999). *Autism diagnostic observation schedule- generic*. Los Angeles, CA: Western Psychological Services.
- Loveland, K., & Landry, S. (1986). Joint attention and language in autism and developmental language delay. *Journal of Autism and Developmental Disorders*, *16*, 335–349.

- McGee, G., Morrier, M., & Daly, T. (1999). An incidental teaching approach to early intervention for toddlers with autism. *Journal of the Association for Persons with Severe Handicaps, 24*, 133–146.
- McHale, S., Simeonsson, R., Marcus, L., & Olley, J. (1980). The social and symbolic quality of autistic children's communication. *Journal of Autism and Developmental Disorders, 10*, 299–310.
- Meisels, S. J. (1989). Can developmental screening tests identify children who are developmentally at risk? *Pediatrics, 83*, 578–585.
- Mitchell, S. K. (1979). Interobserver agreement, reliability, and generalizability of data collected in observational studies. *Psychological Bulletin, 86*, 376–390.
- Mullen, E. (1995). *The Mullen scales of early learning*. Circle Pines, MN: American Guidance.
- Mundy, P., Sigman, M., & Kasari, C. (1990). A longitudinal study of joint attention and language development in autistic children. *Journal of Autism and Developmental Disorders, 20*, 115–128.
- Mundy, P., Sigman, M., Ungerer, J., & Sherman, T. (1987). Nonverbal communication and play correlates of language development in autistic children. *Journal of Autism and Developmental Disorders, 17*, 349–364.
- National Research Council (2001). *Educating children with autism*. Committee on Educational Interventions for Children with Autism. Washington, DC: National Academy Press.
- Osterling, J., & Dawson, G. (1994). Early recognition of children with autism: A study of first birthday home videotapes. *Journal of Autism and Developmental Disorders, 24*, 247–257.
- Osterling, J., & Dawson, G. (1999). Early identification of one-year-old with autism versus mental retardation. Poster presented at the 1999 meeting of the Society of Research in Child Development, Albuquerque, NM.
- Robins, D. L., Fein, D., Barton, M., & Green, J. A. (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*, 131–151.
- Rutter, M., Bailey, A., Berument, S., Lord, C., & Pickles, A. (2001). *Social communication questionnaire research edition*. Los Angeles, CA: Western Psychological Corporation.
- Sheinkopf, S., Mundy, P., Oller, D.K., & Steffens, M. (2000). Vocal atypicalities of preverbal autistic children. *Journal of Autism and Developmental Disorders, 30*, 345–354.
- Short, A., & Schopler, E. (1988). Factors relating to age of onset in autism. *Journal of Autism and Developmental Disorders, 18*, 207–216.
- Siegel, B., Pliner, C., Eschler, J., & Elliot, G. (1988). How children with autism are diagnosed: Difficulties in identification of children with multiple Developmental delays. *Developmental and Behavioral Pediatrics, 9*, 199–204.
- Sigman, M., & Ruskin, E. (1999). Continuity and change in the social competence of children with autism, Down syndrome, and Developmental delays. *Monographs of the Society for Research in Child Development, 64*.
- Sigman, M., & Ungerer, J. (1984). Cognitive and language skills in autistic, mentally retarded and normal children. *Developmental Psychology, 20*, 293–302.
- Sparrow, S. S., Balla, D. A., & Cicchetti, D. V. (1984). *Vineland adaptive behavior scales*. Circle Pines, MN: American Guidance.
- Squires, J., Potter, L., & Bricker, D. (1999). *The ASQ user's guide for the ages & stages questionnaires: A parent-completed, child-monitoring system* (2nd ed.). Baltimore, MD: Paul H. Brookes.
- Stone, W. L., & Caro-Martinez, L. M. (1990). Naturalistic observations of spontaneous communication in autistic children. *Journal of Autism and Developmental Disorders, 20*, 437–453.
- Stone, W., Lee, E., Ashford, L., Brissie, J., Hepburn, S., Coonrod, E., & Weiss, B. (1999). Can autism be diagnosed accurately in children under three years? *Journal of Child Psychology and Psychiatry, 40*, 219–226.
- Stone, W., Ousley, O., Yoder, P., Hogan, K., & Hepburn, S. (1997). Nonverbal communication in 2- and 3-year old children with autism. *Journal of Autism and Developmental Disorders, 27*, 677–696.
- Werner, E., Dawson, G., Osterling, J., & Dinno, N. (2000). Brief report: Recognition of autism spectrum disorder before one year of age: A retrospective study based on home videotapes. *Journal of Autism and Developmental Disorders, 30*, 157–162.
- Wetherby, A., Allen, L., Cleary, J., Kublin, K., & Goldstein, H. (2002). Validity and reliability of the Communication and Symbolic Behavior Scales Developmental Profile with very young children. *Journal of Speech, Language, & Hearing Research, 45*, 1202–1218.
- Wetherby, A., Goldstein, H., Cleary, J., Allen, L., & Kublin, K. (2003). Early identification of children with communication disorders: Concurrent and predictive validity of the CSBS Developmental Profile. *Infants and Young Children, 16* (2), 161–174.
- Wetherby, A., & Prizant, B. (1993). *Communication and symbolic behavior scales – normed edition*. Baltimore, MD: Paul H. Brookes.
- Wetherby, A., & Prizant, B. (1998). *Communication and symbolic behavior scales developmental profile – research edition*. Baltimore, MD: Paul H. Brookes.
- Wetherby, A., & Prizant, B. (2002). *Communication and symbolic behavior scales developmental profile – first normed edition*. Baltimore, MD: Paul H. Brookes.
- Wetherby, A. M., Prizant, B.M., & Hutchinson, T. (1998). Communicative, social-affective, and symbolic profiles of young children with autism and pervasive developmental disorder. *American Journal of Speech-Language Pathology, 7*, 79–91.
- Wetherby, A., & Prutting, C. (1984). Profiles of communicative and cognitive-social abilities in autistic children. *Journal of Speech and Hearing Research, 27*, 364–377.
- Wetherby, A., & Woods, J. (2002). *Systematic observation of red flags for autism spectrum disorders in young children*. Unpublished manual, Florida State University, Tallahassee, FL.
- Wimpory, D. C., Hobson, R. P., Williams, J. M. G., & Nash, S. (2000). Are infants with autism socially engaged? A study of recent retrospective parental reports. *Journal of Autism and Developmental Disorders, 30*, 525–536.
- Wing, L., Gould, J., Yeates, R.R., & Brierley, L.M. (1977). Symbolic play in severely mentally retarded and autistic children. *Journal of Child Psychology and Psychiatry, 18*, 167–178.
- Yeargin-Allsopp, M., Rice, C., Karapurkar, T., Doernberg, N., Boyle, C., & Murphy, C. (2003). Prevalence of autism in a US metropolitan area. *Journal of the American Medical Association, 289*, 49–55.

Appendix A. Rating definitions for the *Systematic Observation of Red Flags for Autism Spectrum Disorders in Young Children (SORF)*; Wetherby & Woods, 2002)

Difficulty with Reciprocal Social Interaction

1. *Aversion to social touch or proximity*: avoids touch or close proximity of another person by moving out of the way, turning away, or pushing away
2. *Lack of appropriate gaze*: lacks or avoids eye gaze to people during face-to-face interaction, displays frequent fleeting gaze, or looks at objects with peripheral gaze, rather than centrally
3. *Lack of warm, joyful expressions with directed gaze*: smiles or laughs; must be directed to another person by orienting body to, looking at, or touching that person
4. *Lack of sharing interest or enjoyment*: shows, brings to, points out, or draws attention to objects of interest or fun activities to share enjoyment
5. *Lack of anticipatory posture or movement in response to interaction*: responds to the social overture of another person with an anticipatory posture (orienting body, leaning, stiffening body to be picked up, molding to body when hugged, assisting with positioning feet when put in Sassy seat or high chair) or movement (reaching out to take an object offered or to be picked up)
6. *Lack of response to contextual cues*: responds to verbal stimuli presented with contextual cues, such as gestural cues (saying give me and holding hand out with palm up; saying bye-bye and waving when removing toy) or situational cues (saying put in after putting several objects in bag to clean up)
7. *Lack of response to name when called*: responds to name by turning toward or looking at the person immediately (within a few seconds) after the child's name is called
8. *Lack of coordination of gaze, facial expression, gestures, and sounds*: communicates with at least three of the following simultaneously: gaze, facial expression, gesture, and sound

Unconventional Gestures

1. *Using person's hand as a tool without directed gaze*: takes, moves, or pulls another person's hand as if it was a tool; either places a hand on an object or manipulates a hand that is holding an object (e.g., pulling hand to lid of jar, pushing hand holding balloon up to the person's mouth)

2. *Lack of pointing*: uses the index finger to reference another's attention to an object, picture, or event; must be with eye gaze directed to another person before, during, or after pointing; may be contact or distal point
3. *Lack of showing*: holds an object out toward another person without giving the object to reference another's attention to the object

Unconventional Sounds and Words

1. *Atypical vocalizations*: produces syllables that have abnormal phonation, such as squeals (very high pitch vocalization produced with high tension that is isolated from other vocalizations), growls (low-pitch, often creaky-voice vocalizations), and yells (high amplitude nondistress vocalizations)
2. *Unusual syllable strings*: produces strings of syllables that have odd combinations of sounds not usually combined in speech (e.g., gudagudagudaguda)
3. *Unusual prosody*: little variation in pitch (flat and mechanical sounding), odd intonation (monotonous, stiff, or stress not on high content words), irregular rhythm (stiff or jerky), or unusual voice quality (difficult to hear, whines, whimpers, nasal)
4. *Immediate echolalia*: repeats phrases (at least two different words combined) produced by another person immediately after
5. *Idiosyncratic or repetitive use of words or phrases*: uses and/or repeats words or phrases that have private meanings not easily understood by someone unfamiliar to the child; may include unintelligible utterances or jargon
6. *Lack of communicative vocalizations with consonants*: uses vocalization that are communicative (i.e., directed by eye gaze, gesture, or movement toward a person) and consist of syllables containing at least one consonant

Repetitive Behaviors and Restricted Interests

1. *Repetitive movements or posturing of body, arms, hands, or fingers*: stereotyped or repetitive movements or posturing of child's body, arms, hands, or fingers (rocks body, turns body in circles, flaps arms, flicks fingers)
2. *Repetitive movements with objects*: stereotyped or repetitive (at least three consecu-

- tive) movements with objects (taps, spins, bangs, lines up, rubs, twirls, rolls, collects objects)
3. *Unusual sensory interest or exploration with objects*: unusual, detailed, or prolonged tactile or visual examination of objects (licks, smells, rubs to feel texture, stares or fixates)
 4. *Excessive interest in or focus on particular toys*: intense interest in a particular toy; focuses on and becomes absorbed with the toy for a much longer time than with other toys
 5. *Lack of playing with a variety of toys and objects*: shows interest in and plays with a variety of toys using conventional actions with objects appropriate for child's age; imitation of actions counts if child repeats that action (appropriate play with the feeding set: mouths, bangs, drops by 8 months; uses objects functionally by 12 months; uses objects toward other by 16 months; sequences actions with objects by 20 months)

Emotional Regulation

1. *Fear or distress about particular objects*: shows signs of being afraid or wary or distressed in response to particular objects (e.g., Koosh animal in Bag of Toys; Big Bird stuffed animal)
2. *Distress over removing particular objects*: displays distress when a particular object is removed (e.g., bubbles, balloon) or an activity is ended (e.g., putting the lid on the bubbles); distress must be strong in intensity and last at least 5 seconds in duration
3. *Difficulty calming when distressed*: when distressed, difficulty calming within a brief time (5 seconds) either by self or with comfort

offered by another person; only rate if child displays distress

4. *Abrupt shifts in emotional or behavioral state*: rapid changes from neutral to positive or to negative emotional states or from relaxed tone to very tense
5. *Heightened alertness and response to stimuli or situations*: easily excited by or negative response to auditory, visual, or tactile stimuli
6. *Flat affect or unresponsive to interactions*: minimal change in facial expression; very passive; lacks interest in interacting
7. *Challenging behavior*: displays challenging behavior (hits, pinches, bites, bangs, pulls hair, scratches) that is self-injurious (hurting self) or aggressive (hurting another person)

AUTHOR NOTE

Amy M. Wetherby, Juliann Woods, Lori Allen and Holly Dickinson are associated with the Department of Communication Disorders, Florida State University. Julie Cleary is now associated with REACH Research, Kennedy Krieger Institute. Cathy Lord is associated with the Autism and Communication Disorders Center, University of Michigan. This research was supported in part by two grants from the U.S. Department of Education, Office of Special Education and Rehabilitation Services (H324M980173 and H324M010071) and a grant from the U.S. Department of Education, Institute of Education Sciences (R305T010262). The authors would like to thank the families who gave their time to participate in this project.