# **Evidence-Based Assessment of Autism Spectrum Disorders** in Children and Adolescents

# Sally Ozonoff, Beth L. Goodlin-Jones, and Marjorie Solomon

M.I.N.D. Institute; Department of Psychiatry & Behavioral Sciences, University of California, Davis

This article reviews evidence-based criteria that can guide practitioners in the selection, use, and interpretation of assessment tools for autism spectrum disorders (ASD). As Mash and Hunsley (2005) discuss in this special section, evidence-based assessment tools not only demonstrate adequate psychometric qualities, but also have relevance to the delivery of services to individuals with the disorder (see also Hayes, Nelson, & Jarrett, 1987). Thus, we use what is known about the symptoms, etiologies, developmental course, and outcome of ASD to evaluate the utility of particular assessment strategies and instruments for diagnosis, treatment planning and monitoring, and evaluation of outcome. The article begins with a review of relevant research on ASD. Next we provide an overview of the assessment process and some important issues that must be considered. We then describe the components of a core (minimum) assessment battery, followed by additional domains that might be considered in a more comprehensive assessment. Domains covered include core autism symptomatology, intelligence, language, adaptive behavior, neuropsychological functions, comorbid psychiatric illnesses, and contextual factors (e.g., parent well-being, family functioning, quality of life). We end with a discussion of how well the extant literature meets criteria for evidence-based assessments.

# Autism Spectrum Disorders (ASD): Background

The Diagnostic and Statistical Manual of Mental Disorders (4th ed., text rev. [DSM-IV-TR]; American Psychiatric Association, 2000) lists five pervasive developmental disorders!: autistic disorder, Asperger's disorder, Rett's disorder, childhood disintegrative disorder, and pervasive developmental disorder not otherwise specified (PDDNOS). Symptoms of autistic disorder fall under three domains: social relatedness, communication, and behaviors and interests, with delays or abnormal functioning in at least one of these areas prior to age 3 years. In the social domain, symptoms include impaired use of nonverbal behaviors (e.g., eye contact, facial expression, gestures) to regulate social interaction, failure to develop ageappropriate peer relationships, little seeking to share

enjoyment or interests with other people, and limited social-emotional reciprocity. Communication deficits include delay in or absence of spoken language, difficulty with conversational reciprocity, idiosyncratic or repetitive language, and imitation and pretend play deficits. In the behaviors and interests domain, there are often encompassing, unusual interests, inflexible adherence to nonfunctional routines, stereotyped body movements, and preoccupation with parts or sensory qualities of objects (American Psychiatric Association, 2000). To meet criteria for autistic disorder, an individual must demonstrate at least 6 of 12 symptoms, with at least 2 coming from the social domain and 1 each from the communication and restricted behaviors/interests categories.

Asperger's disorder (or Asperger's syndrome [AS]) shares the social disabilities and restricted, repetitive behaviors of autism, but language abilities are well developed and intellectual functioning is not impaired.<sup>2</sup> Its symptoms are identical to those listed for autistic disorder, except that there is no requirement that the child demonstrate any difficulties in the second category, communication. The main point of differentiation from autistic disorder, especially the higher functioning subtype, is that those with AS do not exhibit

Our comments in this article have been influenced by the membership of Sally Ozonoff on the National Institute of Mental Health's Workgroups on Interventions Research in Autism Spectrum Disorders (September 5–6, 2002, Rockville, MD, and May 6, 2004, Sacramento, CA). Sally Ozonoff was supported during the writing of this article by funding from the National Institutes of Health (R01–MH068398–01 and U19–HD35468–06).

Requests for reprints should be sent to Sally Ozonoff, M.I.N.D. Institute, UC Davis Medical Center, 2825 50th Street, Sacramento CA 95817. E-mail: sjozonoff@ucdavis.edu

<sup>&</sup>lt;sup>1</sup>This term is used synonymously with ASD in this article.

<sup>&</sup>lt;sup>2</sup>Generally defined as IQ scores above 69, although no operational definition exists and other thresholds, such as IQ > 84, are sometimes used.

significant delays in the onset or early course of language. Communicative use of single words must be demonstrated by age 2 and meaningful phrase speech by age 3. Autistic disorder must be ruled out before a diagnosis of AS is justified. The DSM-IV-TR mandates that the diagnosis of autism always take precedence over that of AS. Thus, if a child meets criteria for autistic disorder, the diagnosis must be autism even if he or she displays excellent language, average or better cognitive skills, and other "typical" features of AS. As discussed in some detail later in this article, consensus has not been achieved on the validity of the distinction between higher-functioning forms of autistic disorder and AS (Howlin, 2003; Macintosh & Dissanayake, 2004). Whether the two conditions are different enough to warrant separate names is of more than academic interest, because in many states resources are provided differentially to children based on the particular autism spectrum diagnosis they receive.

PDDNOS is a label used for children who experience difficulties in at least two of the three autismrelated symptom clusters (clear difficulty relating to others, as well as either communication problems or repetitive behaviors) but who do not meet criteria for any of the other pervasive developmental disorders. The same list of 12 symptoms outlined previously is used to diagnose PDDNOS, but only one difficulty within the reciprocal social interaction domain and one symptom from either the communication deficits or repetitive, restricted behaviors domains is required. Thus, this is a very heterogeneous category (Walker et al., 2004). The diagnosis is often misused, with substantial proportions of children carrying this label either meeting full criteria for autism or not meeting criteria for any ASD (Buitelaar, van der Gaag, Klin, & Volkmar, 1999). Two other conditions also appear in the DSM-IV-TR within the pervasive developmental disorders category: Rett's disorder and childhood disintegrative disorder. Both involve a period of typical development, followed by a loss of skills and regression in development. These conditions are not discussed further in this article.

Kanner (1943), who provided the first description of autism (and coined the term), was the first to identify the much greater preponderance of affected boys. Recent meta-analysis suggests that the widely reported 4:1 ratio of boys to girls is quite consistent across studies, geographical regions, ethnicities, and time (Fombonne, 2003). Early research suggested that autism (strictly defined as children meeting full criteria for the disorder) occurred at the rate of 4 to 6 affected individuals per 10,000 (Lotter, 1966; Wing & Gould, 1979). Newer studies have given prevalence estimates of 60 to 70 per 10,000 or approximately 1 in 150 across the spectrum of autism and 1 in 500 for children with the full syndrome of autistic disorder (Bertrand et al.,

2001; Chakrabarti & Fombonne, 2001). Thus, ASDs are no longer rare conditions, and it is likely that many or most practitioners will encounter individuals with suspected ASD in their practices.

The causes of autism are not yet known, but it has become clear that genetic factors play an important role (Bailey et al., 1995; International Molecular Genetic Study of Autism Consortium, 2001) and that the brain is both structurally and functionally different from normal (Bailey et al., 1998; Courchesne et al., 2001; Schultz et al., 2000), although results are inconsistent across studies and samples and no signature "autistic anomaly" has been identified.

#### Issues in Assessment of ASD

Specific practice parameters for the assessment of ASD have been published by the American Academy of Neurology (Filipek et al., 2000), the American Academy of Child and Adolescent Psychiatry (Volkmar, Cook, Pomeroy, Realmuto, & Tanguay, 1999), and a consensus panel with representation from multiple professional societies (Filipek et al., 1999). These practice parameters describe two levels of screening and evaluation. The first, Level 1 screening, involves routine developmental surveillance by providers of general services for young children, such as pediatricians; Level 2 evaluation involves a comprehensive diagnostic assessment by experienced clinicians for children who fail the initial screening (Filipek et al., 1999, 2000; Volkmar et al., 1999). These publications have been significant milestones in the field of autism, as they lay out, for the first time, consensus guidelines for ASD assessment. We cover Level 2 evaluation in this article, focusing primarily but not exclusively on assessment of school-age children.

There are several important considerations that should inform the assessment process. First, a developmental perspective must be maintained (Burack, Iarocci, Bowler, & Mottron, 2002). Autism is a lifelong disorder. It is first diagnosed in early childhood and continues to be apparent throughout a person's life. It is characterized by unevenness in development that differs over the life span of the individual. Studying a child within a developmental framework provides a benchmark for understanding the severity or quality of delays or deviance. Delays in one developmental achievement can significantly impact the acquisition of later developmental milestones, as when early levels of joint attention predict later language acquisition (Mundy, Sigman, & Kasari, 1990) and theory of mind abilities (Baron-Cohen, 1991). Autism symptoms are usually at their worst in preschool and may substantially improve over time. Children who have very poor eye contact and make few social initiations at this age

may have quite different social symptoms when they are teenagers. They may be relatively interested in social engagement by this later stage and may have acquired some more advanced social skills. Their social difficulties may be manifested as awkwardness or inappropriateness rather than the lack of interest seen in young childhood. Thus, the form and quality of symptoms change with age. There are also characteristic patterns of delays in ASD that differ across domain and developmental level. For example, a child with autism may have meaningful expressive language, a large vocabulary, and adequate syntactic abilities but may not be able to participate in a conversation or even adequately answer questions.

A second important consideration is that the evaluation of a child with ASD should include information from multiple sources and contexts, as symptoms of ASD may be dependent on characteristics of the environment. For example, high-functioning children may present as charming, precocious, and highly intelligent when provided with one-on-one attention and conversational scaffolding from a well-meaning adult professional. The same child may look much more symptomatic with peers on a playground or in a distracting classroom situation where individual adult attention is unavailable. Conversely, children with severe learning and behavioral deficits may look much more competent in a known environment, such as the classroom, than in an evaluation room without familiar, well-practiced routines. Thus, measures of parent report, teacher report, child observation across settings, cognitive and adaptive behavior assessments, and clinical judgments may all be part of the most comprehensive ASD assessment (Filipek et al., 1999).

Third, it is recommended that assessments of ASD are multidisciplinary whenever possible, including professionals from psychology, psychiatry, other medical specialties as needed (e.g., pediatrics, neurology) and speech and language. On interdisciplinary teams, it is important that one member act as the evaluation coordinator. The person in this role communicates with parents and referring professionals before the evaluation to understand the referral questions, organizes appropriate team members, plans the components of the assessment, establishes contact with the service providers in the community who will implement the recommendations from the evaluation, and perhaps monitors later treatment. This type of coordination is critical for the successful outcome of an evaluation.

We first describe the components of what we call a "core" assessment battery, one that covers the foundational elements that are both necessary and sufficient for an evaluation of suspected ASD. The specific approach will depend on the goal of the assessment (e.g., diagnosis, treatment planning, annual or other regularly scheduled assessment, evaluation of treatment

progress, program admission or discharge, eligibility for entitlements, and so on), but the domains we consider first are part of many of these evaluation contexts. After describing the core assessment battery, we discuss other domains that might be part of a more comprehensive assessment or might be necessary for a particular individual, depending on the referral question or evaluation goals.

# Core Autism Assessment: Necessary and Sufficient Domains

The first step of the core assessment process is to review with parents the child's early developmental history and current concerns. The critical aspects of this history-taking are reviews of communication, social, and behavioral development; additionally, a brief screening of potential medical and psychiatric issues, such as anxiety and depression, should be conducted at this stage to determine the need for more detailed evaluation (possibly including referral to specialists). A review of available records (e.g., medical, school, previous testing, intervention reports) rounds out the history-taking aspect of the evaluation. Combined with this review is direct observation of and interaction with the child. Whenever feasible, teachers should be consulted to provide their observations about child functioning in the less structured, socially challenging school setting.

## **Autism Diagnostic Measurement**

There is general agreement on the primary characteristics of autism in North America and Europe, as evidenced by close overlap of the diagnostic criteria laid out in the DSM-IV-TR and International Classification of Diseases-10 (Sponheim, 1996). All professional practice parameters state the necessity of interviewing the parents about early development and specific symptoms of autism as well as observing the child directly (Filipek et al., 1999, 2000; Volkmar et al., 1999), ideally using the types of standardized instruments reviewed later in this article. In the relatively short observation of the child done in most clinic settings, the full range of difficulties experienced by the child will likely not be evident, so parent report is vital. Parents, however, do not have the professional expertise and experience to recognize or interpret all difficulties, so observation and testing by informed practitioners in a controlled setting is also necessary. The information gained from these sources can then be integrated into a DSM-IV-TR diagnosis. In the following we describe first parent report and then observational tools for the diagnostic assessment of ASD.

Clinical impression, oral traditions, and subjective observations dominated the assessment process of ASD until recently (Klinger & Renner, 2000). Use of standard diagnostic criteria and recognition and interpretation of symptoms differed across settings (university clinics, private practice settings, research projects). The publication of two standardized assessment tools, the parent-interview Autism Diagnostic Interview-Revised (ADI-R; Lord, Rutter, & LeCouteur, 1994; Rutter, LeCouteur, & Lord, 2003) and the performance-based Autism Diagnostic Observation Schedule (ADOS; Lord et al., 2000), have ended many of these disparities and are currently considered the "gold standards" for diagnosis of ASD. Use of these and other tools described in this article has advanced scientific progress and improved the accuracy and reliability of diagnostic assessment (Filipek et al., 1999).

Parent interviews and questionnaires. The ADI-R (Lord et al., 1994; Rutter, LeCouteur, et al., 2003) is a comprehensive parent interview that probes for symptoms of autism. It is administered by a trained clinician using a semistructured interview format. The research or long version of the ADI-R requires approximately 3 hr to administer and score. A short edition of the ADI-R, which includes only the items on the diagnostic algorithm, may be used for clinical assessment and takes less time, approximately 90 min (Lord et al., 1994). The use of the ADI-R for research purposes requires attending a 3-day training seminar by a certified trainer and completion of reliability testing with the developers of the instrument. Training to use the ADI-R as a clinical tool is also available; it is helpful but not required for routine use by practitioners who do not participate in research protocols.

The ADI-R elicits information from the parent on current behavior and developmental history. It is closely linked to the diagnostic criteria set forth in the DSM-IV-TR and International Classification of Diseases-10. The significant developmental time point on the ADI-R is age 4 to 5 years for most behaviors. The rationale for the focus on this age period is that it is old enough to provide an adequate range of behavior but young enough to precede major changes that may occur with age (Lord et al., 1994). The items that empirically distinguish children with autism from those with other developmental delays are summed into three algorithm scores measuring social difficulties, communication deficits, and repetitive behaviors. The algorithm scores discriminate children with autism from those with other developmental disorders, such as severe receptive language disorders (Mildenberger, Sitter, Noterdaeme, & Amorosa, 2001) and general developmental delays (Cox et al., 1999; Lord et al., 1994). There are no thresholds yet established for other ASDs (e.g., AS or PDDNOS).

The ADI-R is a very helpful tool, but it does have some limitations. It is not sensitive to differences among children with mental ages below 20 months or IQs below 20 (Cox et al., 1999; Lord, 1995) and is not advised for use with such children. In particular, its sensitivity to the milder ASDs (AS and PDDNOS) is low at age 2, but good by 3½ years. It is not designed to assess change through repeated administrations and is best suited to confirm the initial diagnosis of autism (Arnold et al., 2000). Finally, and perhaps most important, it is labor intensive and requires more administration time than many practitioners may be able to allot.

The Social Communication Questionnaire (formerly known as the Autism Screening Questionnaire; Berument, Rutter, Lord, Pickles, & Bailey, 1999; Rutter, Bailey, et al., 2003) is a parent-report questionnaire based on the ADI-R. It contains the same questions included on the ADI-R algorithm, presented in a briefer. yes/no format that parents can complete on their own. Its agreement with the more labor-intensive ADI-R on diagnostic categorization is high (Bishop & Norbury, 2002), and it is thus an efficient way to obtain diagnostic information or screen for autistic symptoms. There are two versions available—one for current behavior and one for lifetime behavior. The lifetime version is helpful for screening and diagnostic purposes, whereas the current version is more appropriate for assessment of change over time in an individual. A cutoff score of 15 differentiates between ASD and other diagnoses for children ages 4 years and older, whereas a cutoff of 22 discriminates children with autistic disorder from those with other ASDs (PDDNOS or AS). Using these cutoffs, sensitivity of .85 and specificity of .75 have been reported in a large sample of children and adults with autism and other developmental disorders (Berument et al., 1999).

The Autism Behavior Checklist (Krug, Arick, & Almond, 1988) is an informant-report questionnaire that was once widely used in both clinics and schools but is based on conceptualizations of autism that are no longer current (e.g., emphasizing sensory dysfunction and motor stereotypies). Several studies have demonstrated that the rate of both false positives and false negatives produced by the Autism Behavior Checklist is quite high and that most higher functioning children are not identified by the cutoff of 67 (Sevin, Matson, Coe, Fee, & Sevin, 1991; Sponheïm & Spurkland, 1996; Volkmar et al., 1988; Wadden, Bryson, & Rodger, 1991). Therefore, it is not recommended for use.

The Gilliam Autism Rating Scale (Gilliam, 1995) is a recently developed instrument that has rapidly come into wide use in schools and diagnostic clinics. It is typically completed by parents and is appropriate to rate the behavior of children and young adults ages 3 to 22 years. It consists of four scales: Social Interaction, Communication, Stereotyped Behaviors, and Developmental Disturbances. Ratings are made on a 4-point scale, summed, and converted to standard scores based on the reference sample (but not broken down by age or gender). The primary score of interest is the Autism Quotient, which is intended to measure "the likelihood that a child has autism" (Gilliam, 1995). Reference data come from more than 1,000 North American children with informant-reported (but not verified) diagnoses of autism. Enthusiasm for the Gilliam Autism Rating Scale stems from its ease of use, its recent norms, and its explicit relationship to DSM-IV-TR symptoms. However, there is only one empirical report of the psychometric properties of the Gilliam Autism Rating Scale, and it raises significant questions about its utility. In a sample of children with autism, verified by ADI-R, ADOS, and expert clinical consensus, more than half were rated as having below average or very low likelihood of autism by the Gilliam Autism Rating Scale (sensitivity of .48; South et al., 2002). This high false negative rate is seriously troubling, as it may result in many missed diagnoses when used by practitioners with little ASD expertise.

The Parent Interview for Autism (Stone, Coonrod, Pozdol, & Turner, 2003) is a new instrument developed specifically for the purpose of measuring change in autistic symptomatology over time. It is appropriate for preschool children ages 2 to 6. It has good internal consistency and can differentiate autism from nonautistic developmental delays (Stone et al., 2003). Change in Parent Interview for Autism scores after 2 years of intervention correlated highly with clinical ratings of behavioral and diagnostic improvement (Stone et al., 2003). Another instrument developed to measure behavioral change in response to treatment is the PDD Behavior Inventory (Cohen, Schmidt-Lackner, Romanczyk, & Sudhalter, 2003). Norms exist for children ages 2 to 17. The questionnaire covers both autistic symptoms and adaptive and maladaptive behaviors that might be altered by treatment. It demonstrates a high degree of internal consistency, provides adequate test-retest reliability (Cohen et al., 2003), and correlates highly with both the ADI-R and the Childhood Autism Rating Scale (Cohen, 2003). These measures may prove useful for practitioners wishing to track the progress of patients enrolled in treatment programs.

AS diagnostic tools. Within the ASDs, the differential diagnosis of high-functioning autism and AS is both difficult and of questionable nosological validity (Miller & Ozonoff, 2000; Prior, 2000). Although at one time it was proposed that individuals with AS differed from those with autism in several meaningful ways, including cognitive profile (Klin, Volkmar, Sparrow, Cicchetti, & Rourke, 1995; Ozonoff, Rogers, & Pennington, 1991), research has largely failed to confirm this, and most studies conclude that the two are more similar than different (Howlin, 2003). Differ-

ences, when present, are most distinct in early child-hood (Ozonoff, South, & Miller, 2000), and the two conditions appear to converge phenomenologically at older ages (Howlin, 2003; Starr, Szatmari, Bryson, & Zwaigenbaum, 2003). This conclusion has not necessarily been incorporated into clinical practice, however, and there remains a conviction among clinicians that the two are distinct conditions. In recent years, several parent-report measures have been developed to assist with the AS diagnosis.

The Autism Spectrum Screening Questionnaire (Ehlers, Gillberg, & Wing, 1999) is a 29-item checklist standardized for completion by lay informants. It assesses symptoms of both AS and high-functioning autism and does not purport to provide a differential diagnosis between the two. The Autism Spectrum Screening Questionnaire has high internal consistency and good validity (Ehlers & Gillberg, 1993). The Gilliam Asperger's Disorder Scale (Gilliam, 2001) is based on DSM-IV-TR criteria for AS. It was standardized on a multicultural sample of 371 participants with unverified diagnoses of AS. There are no published psychometric studies of the Gilliam Asperger's Disorder Scale, but it is widely used in some settings, such as schools. The Asperger Syndrome Diagnostic Scale (Myles, Bock, & Simpson, 2001) is appropriate for children and adolescents ages 5 through 18. There are no published studies of its psychometric qualities.

None of these tools provides an adequate differential diagnosis between AS and other high-functioning ASDs (e.g., high-functioning autism or PDDNOS). Most were not standardized on well-characterized groups of participants with confirmed diagnoses of AS, and none can reliably rule out other ASD diagnoses (Goldstein, 2002). Although all scales are based on DSM-IV-TR characteristics, all include other symptoms that are not part of the diagnostic criteria and are controversial aspects of the phenotype (e.g., clumsiness). These measures may have some utility for broadly identifying any high-functioning autism spectrum disorder. However, given the paucity of studies clearly demonstrating differences between AS and high-functioning autism, as well as the practical concerns raised by differential insurance reimbursement and eligibility for state entitlement programs, the instruments reviewed in this section are not recommended for differential diagnosis.

Diagnostic observation instruments. The ADOS (Lord et al., 2000; Lord, Rutter, DiLavore, & Risi, 2002) is a semistructured interactive assessment of ASD symptoms. There are four different modules, graded according to language and developmental level, making it possible to administer the ADOS to a wide range of patients, from very young children with no language to verbal, high-functioning adults. Most diagnostic observation instruments are hampered by the

short time period of assessment. One cannot always be sure that a behavior is deficient after only an hour of observation, but this is often all the time a professional has with a patient. The ADOS minimizes this problem by including multiple opportunities or "presses" for social interaction and communication that elicit spontaneous behaviors in standardized contexts. There are, for example, a number of different activities and situations during the ADOS that, in a typical child, elicit eye contact or a question. Once several chances to display these typical social behaviors are missed, a clinician can be reasonably certain that the behavior in question is difficult for the child being assessed. The algorithm for the ADOS includes only social and communication symptoms, as there are no presses for repetitive and stereotyped behaviors and thus their presence or absence cannot be reliably assessed. Two empirically-defined cutoff scores, one for autistic disorder and the other for broader ASD (e.g., PDDNOS or AS) are provided.

For children with younger mental and chronological ages, items from Modules 1 and 2 of the ADOS assess social interest, joint attention, communicative behaviors, symbolic play, and atypical behaviors (e.g., excessive sensory interest, hand mannerisms). For older and more capable individuals, Modules 3 and 4 of the ADOS focus on conversational reciprocity, empathy, insight into social relationships, and special interests. As with the ADI–R, use of the ADOS for research purposes requires attending a training workshop and establishing reliability with a certified trainer. There are shorter clinical trainings for clinicians not involved in research that are, such as those for the ADI–R, very helpful but not required for routine clinical use of the instrument.

Lord et al. (2000) published a study of the psychometric properties of the four modules of the ADOS. Excellent interrater reliability, internal consistency, and test–retest reliability were reported for each module. Diagnostic validity (sensitivity and specificity) for autism versus nonspectrum disorders was also excellent. Currently, several studies are underway that explore the discriminant validity of the ADOS in children with fragile X and William's syndromes. The ADOS is widely used in empirical studies of autism and has been used as an outcome measure in several treatment studies (e.g., Owley et al., 2001).

The Childhood Autism Rating Scale (Schopler, Reichler, & Renner, 1988) is a 15-item structured observation instrument that is appropriate for children over 24 months of age. Items are scored on a 7-point scale (from typical to severely deviant) and summed into a composite score that ranges from 0 to 60. Scores above 30 are consistent with a diagnosis of autism, although lower cutoffs have been recommended for adolescents (Garfin, McCallon, & Cox, 1988). Several studies report high internal consistency, interrater and

test-retest reliability, and criterion-related validity (DiLalla & Rogers, 1994; Eaves & Milner, 1993; Sevin et al., 1991), even when used by raters with little training on the measure or sophistication about ASD (Schopler et al., 1988). The Childhood Autism Rating Scale total score correlates highly with the ADI-R (r =.81; Saemundsen, Magnusson, Smari, & Sigurdardottir, 2003) but overidentifies autism relative to the ADI-R, occasionally classifying children with mental retardation as having autism (Lord, 1997; Saemundsen et al., 2003). It was developed as a tool to rate behavior observed during developmental evaluation but has also been adapted for use as a parent questionnaire (Tobing & Glenwick, 2002). The Childhood Autism Rating Scale is a frequently used measure (Luiselli et al., 2001), but it is based on pre-DSM-IV-TR conceptualizations of autism (Van Bourgondien, Marcus, & Schopler, 1992) and does not measure some constructs now considered important to autism diagnosis and of prognostic significance (e.g., joint attention).

Summary. Several measures are available for collecting information from parents and direct observation of children suspected of ASD, each with strengths and weaknesses. There are few studies comparing these instruments and thus little empirical data to guide clinicians choosing among them. In many cases, practical constraints will dictate choices. Table 1 lists all measures recommended for use, with information on dimensions such as format, administration time, training requirements, and applicable age ranges to assist examiners in choosing among them. One limitation of all diagnostic observational measures for autism is their reliance on current behavior. Deviances and delays typical of autism are most apparent in early childhood and occasionally may be missed or not recognized at an older age (Boelte & Poustka, 2000). In addition, some characteristics of ASD are low baserate behaviors that are not always apparent during an observation or structured interaction with a practitioner. Thus, it is critical for diagnosis to both directly observe the child and obtain information from parents, and we recommend choosing one measure of each type from the list in Table 1. On occasion, these measures provide discordant information (de Bildt et al., 2004; Mildenberger et al., 2001). When this happens, we recommend that further data be collected from teachers (see the section on school context) and other informants in an attempt to resolve the discrepancy.

### **Intellectual Assessment**

A second important domain that must be part of the assessment is intellectual functioning. Intellectual assessment helps frame the interpretation of many observations about the child. Level of intellectual functioning is associated with severity of autistic symptoms, ability

Table 1. Recommended Measures of a Core Assessment Battery for Autism Spectrum Disorders

Measure	Format	Age Range <sup>a</sup>	Administration/ Completion Time	Training Needs <sup>b</sup>
Autism Diagnosis: Parent Report				·
ADI–R	Interview	18 months to adult	1 to 2.5 hr	Intensive
SCQ	Questionnaire	4 years to adult	10 min	Minimal
PIA	Questionnaire	2 to 6 years	20 to 30 min	Minimal
PDDBI	Questionnaire	1 to 17 years	10 to 15 min	Minimal
Autism Diagnosis: Direct Observation				
ADOS	Direct Testing	2 years to adult	30 to 50 min	Intensive
CARS	Observation	2 years to adult	5 to 10 min	Moderate
Intelligence				
Mullen	Direct Testing	Birth to 68 months	15 to 60 min	Moderate
DAS	Direct Testing	2.5 to 17 years	25 to 65 min	Moderate
WISC-IV	Direct Testing	6 to 16 years	50 to 70 min	Moderate
Stanford-Binet 5	Direct Testing	2 to 85 years	45 to 75 min	Moderate
Leiter-Revised	Direct Testing	2 to 20 years	25 to 90 min	Moderate
Language				
CELF	Direct Testing	3 to 21 years	30 to 45 min	Moderate
PPVT	Direct Testing	2.5 to 90+ years	10 to 15 min	Moderate
EOWPVT	Direct Testing	2 to 18 years	10 to 15 min	Moderate
TLC	Direct Testing	5 to 18 years	< 60 min	Moderate
CCC	Questionnaire	5 to 17 years	10 to 15 min	Minimal
Adaptive Behavior				
Vineland	Interview	Birth to 18 years	20 to 60 min	Moderate

Note: ADI-R = Autism Diagnostic Interview-Revised; ADOS = Autism Diagnostic Observation Schedule; CARS = Childhood Autism Rating Scale; CCC = Children's Communication Checklist; CELF = Clinical Evaluation of Language Fundamentals; DAS = Differential Abilities Scale; EOWPVT = Expressive One Word Picture Vocabulary Test; PDDBI = Pervasive Developmental Disorders Behavior Inventory; PIA = Parent Interview for Autism; PPVT = Peabody Picture Vocabulary Test; SCQ = Social Communication Questionnaire; TLC = Test of Language Competence; WISC-IV = Wechsler Intelligence Scale for Children (4th ed.).

<sup>a</sup>Inclusive (e.g., 2 to 6 years = from 2 years 0 months through 6 years 11 months). <sup>b</sup>Minimal = little to no training required, but presumes familiarity with instrument; Moderate = presumes prior basic interviewing/cognitive assessment training; Intensive = additional specialized training, such as workshop attendance, suggested.

to acquire skills, and level of adaptive function and is one of the best predictors of outcome (Harris & Handleman, 2000; Lotter, 1974; Rutter, 1984; Stevens et al., 2000; Venter, Lord, & Schopler, 1992). Major goals of intellectual assessment include generating a profile of the child's cognitive strengths and weaknesses, facilitating educational planning, determining eligibility for certain IQ-related services (e.g., statefunded developmental disability services), and suggesting prognosis. Measured IQ is more stable and predictive the older the age at assessment (Lord & Schopler, 1989). Scores can and do change with development and intervention (Freeman et al., 1991; Mayes & Calhoun, 2003a) and may also change as a function of the assessment instrument chosen (Magiati & Howlin, 2001).

The child with suspected ASD often presents an assessment challenge due to social difficulties, unusual use of language, frequent off-task behaviors, high distractibility, and variable motivation. Motivation can have a tremendous influence on test results, and assessments that incorporate reinforcement procedures can result in very different test scores (Koegel, Koegel, & Smith, 1997). It is important to enhance motivation as much as possible without altering the standard administration of the instrument and consider the motivational element when interpreting scores. More frequent

breaks may be needed, and testing may need to be conducted over multiple shorter sessions. When experienced clinicians evaluate children with autism, few should be "untestable." Untestability reflects primarily lack of availability of appropriate tests or clinician inexperience. There are special concerns about the validity of testing younger, lower functioning, and nonverbal children, and care must be taken in choosing appropriate tests. It is important that the test chosen (a) is appropriate for both the chronological and the mental age of the child, (b) provides a full range (in the lower direction) of standard scores, and (c) measures verbal and nonverbal skills separately (Filipek et al., 1999).

There are several commonly used tests for children with lower mental ages (e.g., those who are younger, nonverbal, or have moderate to severe mental retardation). The Leiter International Performance Scales—Revised (Roid & Miller, 1997; Tsatsanis et al., 2003) is appropriate for individuals with a mental age of 2 years or higher and requires no expressive or receptive language skills. The Differential Abilities Scales (Elliott, 1990) assess both intellectual and academic skills. It is growing in popularity and use because it can be administered to children across a wide chronological and mental age range (2½ through 17 years), making it ideal for repeat administrations, to track progress, and

for research projects in which the developmental range of participants may vary considerably. Especially helpful for the ASD population is the option of out-ofrange testing (i.e., administration of tests usually given to children of a different age): Norms for school-age children are available for the preschool battery, permitting use of the test with older children with significant intellectual limitations. For younger children (less than 5) or those with skills that fall below the entry levels of the tests just described, there are a few additional choices for assessment of intellectual functioning, including the Bayley Scales of Infant Development-II (for ages 1 to 42 months; Bayley, 1993) and the Mullen Scales of Early Learning (for ages 1 to 68 months; Mullen, 1995). For children suspected of ASD, the Mullen Scales of Early Learning is often chosen over the Bayley Scales due to its wider age range and five distinct scales that allow separate assessment of verbal and nonverbal abilities. The Bayley Scales have a longer research tradition than the Mullen Scales of Early Learning but yield less detailed information, with one score averaging memory, problem solving, communication, and other abilities. These instruments provide both standard scores and developmental age equivalents. Thus, they can be used to evaluate children who are older than the test norms but whose developmental skills are not high enough to administer more age-appropriate instruments.

For children with spoken language, the Wechsler Intelligence Scales are the most widely used intellectual instruments. There are not yet any published studies using the most recent revision (4th ed.; Wechsler, 2003), but studies using earlier editions (i.e., the revised edition and the third edition) find that individuals with ASD often exhibit uneven subtest profiles. Performance IQ (PIQ) is often higher than verbal IQ (VIQ; Lincoln, Allen, & Kilman, 1995), but the verbal-performance discrepancy is severity dependent, and the majority of individuals with ASD do not show a significant split (> 12 points; Siegel, Minshew, & Goldstein, 1996). When present, a PIQ > VIQ pattern can have important implications for how the child learns best and what activities may be most and least enjoyable. A recent study suggests that children with significantly uneven intellectual development (in favor of nonverbal skills) are more socially impaired than those with similar overall intelligence but smaller or reversed nonverbal-verbal discrepancies (Joseph, Tager-Flusberg, & Lord, 2002). They also, as a group, demonstrate larger head circumference and brain volume than children without major nonverbal-verbal discrepancies (Tager-Flusberg, & Joseph, 2003), suggesting that such children may form an etiologically distinct subtype of autism.

Children with AS may exhibit the opposite intellectual test profile, with VIQ significantly higher than PIQ (Klin et al., 1995), but this is by no means universal and has not been replicated in all studies (see

Ozonoff & Griffith, 2000, for a review). Thus, intellectual test profiles should never be used for diagnostic confirmation or differential diagnosis of ASD subtypes (e.g., AS from high-functioning autism). However, when a VIQ > PIQ profile is evident, the child may prefer verbally based leisure activities, may benefit from verbal explanations, and may excel in subjects that require good verbal processing (Klin et al., 1995), unlike a child with the opposite (PIQ > VIQ) intellectual profile.

There are fewer published studies of the Stanford-Binet Intelligence Scale (4th ed.) with children with ASD, but they suggest similar patterns (e.g., nonverbal IQ higher than verbal IQ, particularly in young children; Mayes & Calhoun, 2003b). One benefit of the Stanford-Binet is the very wide age range of individuals for whom it is appropriate (2 to 85 years). The recently revised fifth edition (Roid, 2003) included 108 children with autism in the normative sample and added entry items, improving measurement of young children, lower-functioning older children, and adults with mental retardation. It is appropriate for both verbal and nonverbal individuals, because half the subtests utilize a nonverbal mode of testing. The fifth edition of the Stanford-Binet may be a good choice when examiners must select an instrument before knowing a child's abilities or when longitudinal assessment is planned.

# Language Assessment

Expressive language level is, along with IQ, the other best predictor of long-term outcome, so it is an especially important characteristic to measure in terms of thinking about future prognosis (Lotter, 1974; Rutter, 1984; Stone & Yoder, 2001). A variety of general instruments, such as the Peabody Picture Vocabulary Test (Dunn & Dunn, 1997), Expressive One-Word Picture Vocabulary Test (Brownell, 2000), Clinical Evaluation of Language Fundamentals (Semel, Wiig, & Secord, 2003), and Preschool Language Scales (Zimmerman, Steiner, & Pond, 2002), have been used with children with ASD to measure receptive and expressive language abilities, but referral for a more comprehensive evaluation by a speech-language pathologist who can give detailed language recommendations is also often helpful (Filipek et al., 1999). Children with adequate spoken language, who score in the average range on these tests, may still exhibit deficits in the use of language in a social context. Pragmatic communication includes nonverbal behaviors (e.g., eye contact, gesture, facial expression, body language), turntaking, and understanding of inferences and figurative expressions. Tests that examine pragmatic language include the Test of Language Competence (Wiig & Secord, 1989) and the Children's Communication Checklist (Bishop & Baird, 2001).

## **Adaptive Behavior Assessment**

This domain makes up the final component of the core autism assessment. It is an essential component for three reasons. First, assessment of adaptive behavior should always accompany intellectual testing, because a diagnosis of mental retardation cannot be made unless functioning is compromised across both standardized tests of intelligence and real-life measures of adaptive function. Measuring adaptive behavior is also important for setting appropriate goals in treatment planning. Adaptive abilities largely determine whether an individual requires constant supervision or is capable of some independence. Finally, it is an important measure of outcome that has been used in many longitudinal and treatment studies (e.g., Freeman, Del'Homme, Guthrie, & Zhang, 1999; Szatmari, Bryson, Boyle, Streiner, & Duku, 2003). Children with autism consistently demonstrate adaptive behavior levels that are lower than their intelligence, and this pattern is most pronounced for higher functioning and normal IQ individuals (Boelte & Poustka, 2002).

The most widely used adaptive measure with children suspected of ASD (Luiselli et al., 2001) is the Vineland Adaptive Behavior Scales (Sparrow, Balla, & Cicchetti, 1984). The domains of functioning include communication, daily living skills, socialization, and, for children under 5, motor skills. The Vineland is completed during an interview with a parent or teacher and is appropriate for children up to age 19 and mentally retarded adults (separate norms are provided for each population). Supplementary norms for individuals with autism are available (Carter et al., 1998). A recent study found that the Vineland was moderately sensitive to changes due to developmental progress (Charman, Howlin, Berry, & Prince, 2004). The Vineland is currently undergoing restandardization and will include supplemental norms for children with ASD (S. Sparrow, personal communication, October 28, 2002). There are no published studies using other adaptive measures, such as the Scales of Independent Behavior (Bruininks, Woodcock, Weatherman, & Hill, 1996) or the Adaptive Behavior Assessment System (Harrison, & Oakland, 2003), with individuals with ASD, but these may be reasonable choices when time is a constraint, as they are questionnaires completed by parents, rather than interviews, and require little to no training to score and interpret.

# Additional Domains of Assessment: Beyond the Core Battery

Depending on the referral question(s), goals of the assessment, and practical constraints such as finances, insurance reimbursements, and waiting lists, a more

comprehensive evaluation might include a number of additional components.

## **Neuropsychological Assessment**

The neuropsychology of ASD has been studied extensively. As a group, persons with ASD exhibit spared rote, mechanical, and visual-spatial processes and deficient higher order conceptual processes, such as abstract reasoning (Minshew, Goldstein, & Siegel, 1997). They often perform acceptably on simple language, memory, and perspective-taking tasks but show deficits when tasks become more complex. Data from neuropsychological testing may be able to provide greater clarity about the individual's profile of strengths and weaknesses, an important foundation for treatment and educational planning. However, neuropsychological testing is costly and time consuming, and its use may be impacted by managed-care concerns (Piotrowski, 1999). The decision to carry out neuropsychological assessment, the choice of domains to evaluate, and the selection of instruments should be done thoughtfully, emphasizing those with most relevance for educational and treatment plans (Groth-Marnat, 1999; Klin & Shepard, 1994; Ozonoff, Dawson, & McPartland, 2002). Space issues preclude a comprehensive review of all domains of neuropsychology; in the following we discuss three areas of particular interest with this population. Neuropsychological assessment is not usually useful (or even possible) with nonverbal or mentally retarded children with ASD. It may be warranted for higher functioning individuals when there are unexplained discrepancies or weaknesses in school performance, behavioral difficulties that appear to stem from undiagnosed learning disorders, and suspected organic problems. For example, neuropsychological assessment of children with unexpected school failure or behavioral issues at school may reveal attention, flexibility, or organization problems that cause frustration, anxiety, or disorganization and significantly interfere with school function.

Attention. Children with ASD do not usually have problems with sustained attention (Garretson, Fein, & Waterhouse, 1990). They do, however, have difficulty with focused attention. In particular, they tend to over-focus their attention on extraneous details while missing meaning, a difficulty that has also been called impaired central coherence (Frith & Happe, 1994). Some children with ASD do exhibit classic attention deficit hyperactivity disorder symptoms of distractibility and hyperactivity (Noterdaeme, Amorosa, Mildenberger, Sitter, & Minow, 2001; R. Perry, 1998). For these children, a traditional attention deficit hyperactivity disorder work-up is indicated (see Pelham, 2005). Measures such as continuous performance tests

may be helpful to examine treatment response in such children (Aman et al., 2004).

Executive function. One of the most consistently replicated cognitive deficits in individuals with ASD is executive dysfunction (Pennington & Ozonoff, 1996; Russell, 1997). The executive function domain includes the many skills required to prepare for and execute complex behavior, such as planning, inhibition, organization, self-monitoring, cognitive flexibility, and set-shifting. Because executive functions are important to school success (Clark, Prior, & Kinsella, 2002), predict response to treatment (Berger, Aerts, van Spaendonck, Cools, & Teunisse, 2003) and longterm outcome (Szatmari, Bartolucci, Bremner, Bond, & Rich, 1989), and are associated with real-world adaptive skills (Clark et al., 2002; Gilotty, Kenworthy, Sirian, Black, & Wagner, 2002), they are important skills to measure.

The gold standard executive function task is the Wisconsin Card Sorting Test (Grant & Berg, 1948; Heaton, Chelune, Talley, Kay, & Curtiss, 1993), which measures cognitive flexibility and set-shifting. It is available in both an examiner-administered and a computer version. Persons with ASD often perform better on the computer version of the test (Ozonoff, 1995). If this executive function test is being given to document deficits for the purposes of treatment eligibility, it may therefore be best to use the examiner-administration format. If, however, the examiner wants to evaluate achievement under supportive conditions, to see how well the child is potentially capable of performing, then the computer-administration format may be preferable (Ozonoff, South, & Provencal, 2005). Computer administration is also more time- and cost-efficient, so when evaluators face such practical constraints, as they often do (Groth-Marnat, 1999), it may be an acceptable choice.

The Delis-Kaplan Executive Function System (Delis, Kaplan, & Kramer, 2001) provides a battery of tests that assess cognitive flexibility, concept formation, planning, impulse control, and inhibition in children and adults. This measure was standardized on a sample of more than 1,700 children and adults ages 8 to 89. Most of its nine subtests are adaptations of traditional research measures of executive function that have been refined to examine skills more precisely, with fewer confounding variables. Subtests include Trail Making, Verbal Fluency, Design Fluency, Color-Word Interference (similar to a Stroop test), Sorting (similar to the Wisconsin Card Sorting Test), Twenty Questions, Tower (similar to the Towers of Hanoi or London), Word Context, and Proverbs. There are not yet any published studies using this instrument with children with autism, but its clinical use is increasing. The NEPSY (Korkman, Kirk, & Kemp, 1998) is another test that includes several measures of executive function, and it can be used with younger children (ages 3 to 12) than the Delis-Kaplan Executive Function System.

The Behavioral Rating Inventory of Executive Function (Gioia, Isquith, Guy, & Kenworthy, 2000) is a parent- or teacher-rated questionnaire for children ages 5 to 18 years that has 86 questions and takes about 10 min to complete. Clinical scales measure inhibition, cognitive flexibility, organization, planning, metacognition, emotional control, and initiation. Specific items tap everyday behaviors indicative of executive dysfunction that may not be captured by performance measures, such as organization of the school locker or home closet, monitoring of homework for mistakes, or trouble initiating leisure activities. Thus, this measure may have more ecological validity than other executive function tests. It can be especially useful to document the impact of executive function deficits on the child's real-world functioning and to plan treatment and educational accommodations. Correlational analyses with other behavior rating scales and executive function tests provide evidence of both convergent and divergent validity (Gioia et al., 2000), and it has been used empirically with samples with autism (Gilotty et al., 2002).

Academic functioning. Assessment of academic ability, even in younger children, is helpful for the purposes of educational decision making. It is often an area of strength that can go unrecognized. Many children with ASD have precocious reading skills and can decode words at a higher level than others of the same age and functional ability. Reading and other academic strengths can be used to compensate for weaknesses, as when a written schedule is provided to facilitate transitions (Bryan & Gast, 2000) or written directions are supplied to improve compliance. The good memory of children with ASD may mean that spelling lists and multiplication tables will be learned more easily (Mayes & Calhoun, 2003a). Conversely, specific areas of weakness also exist, with the most consistently demonstrated one being in reading comprehension. This academic profile is quite different from the problem patterns most teachers and school psychologists are trained to detect (e.g., the poor decoding but good comprehension of dyslexia). Thus, it is important that appropriate test batteries that highlight both academic strengths and weaknesses are included in the comprehensive evaluation, the learning patterns they suggest are interpreted in the feedback to parents and the written report, and appropriate educational recommendations are made. For young children, the Bracken Test of Basic Concepts (Bracken, 1998), the Young Children's Achievement Test (Hresko, Peak, Herron, & Bridges, 2000) and the Psychoeducational Profile (Schopler, Reichler, Bashford, Lansing, & Marcus, 1990) are useful instruments that highlight both the strengths and the

challenges typical of ASD. For older children who are verbal, the most often used academic tests are the Woodcock–Johnson Test of Achievement (Mather & Woodcock, 2001) and the Wechsler Individual Achievement Test (Wechsler, 2002).

Some children with ASD may exhibit a so-called nonverbal learning disability profile (Rourke, 1995). Children with this diagnosis have difficulties in tactile perception, psychomotor coordination, mathematic reasoning, visual-spatial organization, and nonverbal problem solving. They have well-developed rote verbal skills, as well as strong verbal memory and auditory linguistic capabilities. Some children with AS and high-functioning autism display a nonverbal learning disability profile (Klin et al., 1995). They may require additional interventions, such as occupational therapy and math tutoring. A nonverbal learning disability is an academic diagnosis that does not take the place of the primary ASD diagnosis, which is a more complete description of the full range of the child's behavioral and developmental limitations.

## **Psychiatric and Other Comorbidities**

Over the course of development, children with ASD may develop new symptoms and behaviors that disrupt their daily functioning. Behavioral changes can include problems with sleep, appetite, mood, anxiety, activity level, anger management, and aggression. Many factors influence the presentation of psychiatric disorders in individuals with ASD and complicate their diagnosis. The decrement in functioning associated with having an ASD means that the baseline is already lower than average and that a change in behavior has to be relatively marked to be identifiable. Autism, by itself, causes a variety of psychosocial deficits and maladaptive behaviors and their presence may "mask" other psychiatric symptoms or make them difficult to identify. Cognitive limitations may mean that the range and quality of symptoms differ. For example, anxiety may be manifest as obsessive questioning or insistence on sameness, rather than rumination or somatic complaints. Individuals with ASD may not demonstrate certain symptoms, such as the feelings of guilt often seen in depression or the grandiosity and inflation of self-esteem typical of mania. The diminished ability to think abstractly, communicate effectively, and be aware of and describe internal states also means that interview and self-report measures are often of less use. People with autism may lack the self-insight to recognize symptoms or the motivation and social relatedness needed to report them (D. W. Perry, Marston, Hinder, Munden, & Roy, 2001). Thus, the assessment of coexisting psychiatric illness can be quite tricky. Nevertheless, it is important to add to an evaluation whenever significant behavioral issues outside the autism spectrum (e.g., inattention, mood instability, anxiety, sleep

disturbance, aggression, and so on) are evident or when major changes in behavior from the typical baseline are reported. Comorbidity should also be carefully investigated when severe or worsening symptoms are present that are not responding to traditional methods of treatment (Lainhart, 1999).

Depression is one of the most common coexisting syndromes observed in individuals with ASD, particularly higher functioning individuals who can describe their difficulties (Lainhart & Folstein, 1994). Anxiety is also frequently reported (Kim, Szatmari, Bryson, Streiner, & Wilson, 2000). Assessment of these problems is challenging, because no specific tools for the autism spectrum have been developed. The validity of existing inventories (e.g., Children's Depression Inventory [Kovacs, 1992]; Multidimensional Anxiety Scale for Children [March, 1997]) is uncertain, because they require self-report. Given the limited selfinsight of ASD, reports of "no problems" should be interpreted with caution, and careful interviews of parents should be included in the assessment. No empirical studies of the use of these instruments with ASD have been performed.

The revised Child Behavior Checklist (Achenbach & Rescorla, 2001) is widely used to identify child behavioral and mental health issues but has only rarely been used with children with ASD. It does not provide an autism factor, but a few studies have suggested that certain patterns, such as high scores on the Social Problems and Thought Problems scales, may be associated with an ASD diagnosis (Bolte, Dickhut, & Poustka, 1999; Duarte, Bordin, de Oliveira, & Bird, 2003). The Child Behavior Checklist's utility in identifying comorbid internalizing and externalizing problems in children with ASD is not yet known, but it may be useful as a screening tool given its excellent psychometric properties.

Another measure for assessing several symptom profiles simultaneously is the Behavioral Assessment System for Children. These scales include parent-report, teacher-report, and self-report questionnaires for children ages 8 to 18 years (Kamphaus, Reynolds, & Hatcher, 1999). There are scales for internalizing, externalizing, and adaptive behaviors. Subscales assess school, clinical, and personal adjustment. The self-report form also measures "sense of inadequacy" and "sense of atypicality," which in our experience are helpful for understanding the struggles of children with AS and high-functioning autism who can validly report on their internal states (Ozonoff, Provencal, & Solomon, 2002). These subscales may also prove helpful for measuring treatment effects in ASD (Ozonoff, Provencal, et al., 2002). Importantly, each form provides caution indexes to inform the clinician of overly positive or negative responses and to provide a measure of the consistency of the respondent's profile.

Another multisymptom scale often used with ASD is the Aberrant Behavior Checklist (Aman, Singh, Stewart, & Field, 1985). It is a global behavior checklist completed by a caregiver or a teacher familiar with the child in different settings. It was initially designed as a scale for rating inappropriate and maladaptive behavior of mentally retarded individuals in residential settings (Aman et al., 1985). However, the scale has been used often to monitor the effects of a variety of pharmacological, behavioral, dietary, and other treatments that may be expected to alter behavior and is sensitive to change in ASD (Arnold et al., 2000). There are five subscales (Irritability, Lethargy, Stereotypy, Hyperactivity, and Inappropriate Speech). The published validity and reliability studies report excellent test-retest reliability, internal consistency, and construct validity (Aman et al., 1985).

Another method of assessing problem behaviors that often coexist with ASD, such as aggression, destructiveness, tantrums, stereotypies, or self-injury, is functional analysis (Horner, 1994; O'Neill et al., 1997). Such challenging behaviors are rarely random and usually serve a purpose. Functional analysis is a systematic approach to determine the function of the behavior, that is, what the child is trying to communicate through the behavior. Some common functions of problem behaviors include getting access to a desired object; asking for help or attention; escaping a situation (e.g., schoolwork); and expressing a sensation (e.g., hunger, illness), emotion, or state (e.g., confusion, frustration). The ultimate goal of functional analysis is to provide the child with a more appropriate means of expressing the message (also called functional communication training; Carr & Durand, 1985). Although the functions of a particular problem behavior may seem obvious, the perceptions of informants who work with the child may not be confirmed through direct observations and analog probes that replicate the environmental antecedents of the problem behavior (Calloway & Simpson, 1998). Thus, functional assessment may require referral to a professional trained in these methods, such as a certified behavior analyst, who will also be able to assist in development of a behavioral support plan.

### **The School Context**

Because the goal of assessment should be to understand how ASD affects individuals in the course of daily life, when feasible it is helpful to augment the evaluation by obtaining information from teachers or others who interact with the child in the challenging and relatively unstructured school setting (Klin, Sparrow, Marans, Carter, & Volkmar, 2000). Teachers can be excellent sources of information about the child's adaptive, social, and emotional functioning outside of home and thereby enrich the clinician's understanding

of the child. For example, it has been shown in typically developing children, as well as those with attention deficit hyperactivity disorder, that teacher reports of peer relationships correspond more closely to ratings completed by peers than do parent ratings (Glow & Glow, 1980; Hinshaw & Melnick, 1995). Information from the school setting can be obtained through interviews, questionnaires, and direct clinician observations. Measures include the classroom and teacher editions of the Vineland Adaptive Behavior Scales (Sparrow et al., 1984), the PDD Behavior Inventory (Cohen et al., 2003), the Behavioral Assessment System for Children, and the Aberrant Behavior Checklist. Although not specifically designed for children on the autism spectrum, the teacher report form of the Social Skills Rating System (Gresham & Elliott, 1989) has been used successfully in research to assess social skills in children with ASD (Bauminger, 2002). In addition to questionnaires, school-based observations may yield a richer perspective on child social functioning or may be part of a functional analysis of behavioral problems (see Dunlap & Kern, 1993, and Wood, 1995, for examples).

When information on a child is collected from multiple sources, there may be disagreements in reports of the severity of the disorder, the level of daily adaptive behaviors, and the level of compliance or disruptive behaviors (Offord et al., 1996; Szatmari, Archer, Fisman, & Streiner, 1994). High levels of stress experienced by families appear to contribute to higher parent than teacher reports of autistic behavior (Szatmari et al., 1994). Because these well-known discrepancies exist and may well reflect setting-dependent expression of symptoms, our recommendation is to conceptualize them as separate types of information, without attempting to reconcile them by considering one more or less accurate than another, as suggested by Offord et al. (1996).

#### **Family and Community Context**

Assessment of the family system and community resources is important to service delivery and may also be related to outcome (Hauser-Kram, Warfield, Shonkoff, & Krauss, 2001). Many studies have documented increased stress and depression in parents of children with ASD (Bristol, 1984; Wolf, Noh, Fisman, & Speechley, 1989) that exceed that of parents of children with other disabilities (Olsson & Hwang, 2001). Stress levels are strongly correlated with severity of the child's disorder (Tobing & Glenwick, 2002).

There are several clinical instruments that measure the impact of a disabled child on the family. Those with established psychometric properties that have been used with the ASD population include the Parenting Stress Index (Abidin, 1995), the Questionnaire on Resources and Stress (Holroyd, 1974; Konstantareas, Homatidis, & Plowright, 1992), the Stress Index for Parents of Adolescents (Sheras, Abidin, & Konold, 1998), and the Beck Depression Inventory (Beck, 1987). Scales that measure family and community resources include the Family Support Scale (Dunst, Jenkins, & Trivette, 1984) and the Family Needs Scale (Dunst, Cooper, Weeldreyer, Snyder, & Chase, 1988), which assesses the availability of resources, such as education, leisure, employment, finances, and transportation. Also potentially important to measure is quality of life, although no scales designed for the autistic population exist. The Quality of Life Questionnaire (Schalock & Keith, 1993) has been used in a study of adults with mental retardation, a small proportion of whom also had autism (Kraemer, McIntyre, & Blacher, 2003). For research purposes, it has also been suggested that family and community risk and opportunity factors (e.g., socioeconomic status, residential and marital stability) are important to measure (Wolery & Garfinkle, 2002), although this is rarely done.

## **Evaluation of Response to Treatment**

One of the most functional contributions of assessment is the planning and evaluation of intervention (Hayes et al., 1987). This purpose was highlighted frequently in this article, with certain domains (e.g., language, adaptive behavior) and specific instruments within those domains evaluated in terms of their sensitivity to change. There are, however, no widely agreed-on skills that must change, or any degree of change, for treatment effects to be considered clinically significant. There is a general consensus that intervention outcomes should have social validity, making a genuine (practical, noticeable) difference in everyday life for the person treated or those who live with him or her (Foster & Mash, 1999; Kazdin, 1999). There is less agreement on the magnitude of change that must be shown to be considered clinically significant. It is often measured as an effect size, a percentage decline in symptoms, a change from baseline (pretreatment), or an attainment of functioning within the normal range (Kendall, Marrs-Garcia, Nath, & Sheldrick, 1999). Kazdin, however, argued that change that does not meet such criteria may be meaningful if it helps the person become more functional, even if symptoms remain well outside the normal range. Even no change in symptoms at all may be significant if the treatment improves coping skills (i.e., the ability to deal with the symptoms). These examples are particularly relevant to ASD, a lifelong condition in which functioning within the normative range may be possible for only a small proportion of affected individuals. Change in quantity, quality, or severity of autistic behaviors may be minimal, and many treatment studies do not consider this domain the primary target of the therapy (Kasari, 2002). Many drug studies, for example, focus

on change in aberrant behaviors, such as irritability and aggression, or other target symptoms that make life difficult for individuals and families (e.g., Arnold et al., 2000, 2003). Therefore, domains of central importance in the evaluation of response to treatment are adaptive behavior, comorbid symptoms, quality of life, and family functioning (Wolery & Garfinkle, 2002).

### Conclusion

In this article, we reviewed the components of both a minimal assessment battery and a more comprehensive evaluation of suspected ASD. These components were selected empirically; that is, they have been demonstrated to be relevant to identification, differential diagnosis, service delivery, evaluation of outcome, or a combination of these in multiple empirical investigations. We covered a wide range of assessment strategies and tools, not only those with empirical support but also those in wide use that may not be supported by data, and new instruments that have not yet been studied. There is not always good correspondence between clinical practices and research (Luiselli et al., 2001). Some of the domains and instruments we reviewed are nearly universal in research studies but are rarely used in clinical practice (e.g., ADI-R), whereas others are widely used in clinical but not research settings (e.g., family needs surveys). Very few studies have directly compared different instruments, and thus there is little empirical basis to guide practitioners who are selecting among different assessment tools. Similarly, the incremental validity of the assessment domains we conceptualize outside the core battery has not been examined. But even though challenges remain, we have come a long way in the past few decades. As recently as 10 years ago, autism was considered a rare disorder; few clinicians knew how to evaluate it, or even considered it in a differential diagnosis. Autism was diagnosed through subjective clinical opinion, without the use of objective measures of development or behavior. As consensus about the diagnosis has been achieved, a number of standardized interviews and observational measures have been developed. Competent clinical evaluation now assumes the use of objective measures; funding and publication require it. This article is a first attempt at reviewing the evidence basis for tools currently in existence.

# References

Abidin, R. R. (1995). *Parenting Stress Index* (3rd ed.). Odessa, FL: Psychological Assessment Resources.

Achenbach, T. M., & Rescorla, L. A. (2001). Manual for the ASEBA school-age forms and profiles. Burlington: University of Vermont, Research Center for Children, Youth, and Families.

- Aman, M. G., Novotny, S., Samango-Sprouse, C., Lecavalier, L., Gadow, K. D., King, B. H., et al. (2004). Outcome measures for clinical drug trials in autism. CNS Spectrum, 9, 36–47.
- Aman, M. G., Singh, N. N., Stewart, A. W., & Field, C. J. (1985). The Aberrant Behavior Checklist: A behavior rating scale for the assessment of treatment effects. American Journal of Mental Deficiency, 89, 485–491.
- American Psychiatric Association. (2000). Diagnostic and statistical manual of mental disorders (4th ed., text rev.). Washington DC: Author.
- Arnold, L. E., Aman, M. G., Martin, A., Collier-Crespin, A., Vitiello, B., Tierney, E., et al. (2000). Assessment in multi-site randomized clinical trials of patients with autistic disorder: The autism RUPP network. *Journal of Autism and Developmental Disorders*, 30, 99-111.
- Arnold, L. E., Vitiello, B., McDougle, C., Scahill, L., Shah, B., Gonzalez, N. M., et al. (2003). Parent-defined target symptoms respond to risperidone in the RUPP autism study: Customer approach to clinical trials. *Journal of the American Academy of Child & Adolescent Psychiatry*, 42, 1443-1450.
- Bailey, A., LeCouteur, A., Gottesman, I., Bolton, P., Simonoff, E., Yuzda, E., et al. (1995). Autism as a strongly genetic disorder: Evidence from a British twin study. *Psychological Medicine*, 25, 63-77.
- Bailey, A., Luthert, P., Dean, A., Harding, B., Janota, I., Montgomery, M., et al. (1998). A clinicopathological study of autism. *Brain*, 121, 889-905.
- Baron-Cohen, S. (1991). Precursors to a theory of mind: Understanding attention in others. In A. Whiten (Ed.), Natural theories of mind: The evolution, development and simulation of everyday mindreading (pp. 233-251). Oxford, England: Blackwell.
- Bauminger, N. (2002). The facilitation of social-emotional understanding and social interaction in high-functioning children with autism: Intervention outcomes. *Journal of Autism and De*velopmental Disorders, 32, 283–298.
- Bayley, N. (1993). The Bayley Scales of Infant Development (2nd ed.). San Antonio, TX: Psychological Corporation.
- Beck, A. T. (1987). Beck Depression Inventory manual. San Antonio, TX: Psychological Corporation.
- Berger, H. J. C., Aerts, F. H. T. M., van Spaendonck, K. P. M., Cools, A. R., & Teunisse, J. P. (2003). Central coherence and cognitive shifting in relation to social improvement in high-functioning young adults with autism. *Journal of Clinical and Experimental* Neuropsychology, 25, 502-511.
- Bertrand, J., Mars, A., Boyle, C., Bove, F. Yeargin-Allsopp, M., & Decoufle, P. (2001) Prevalence of autism in the United States population: The Brick Township New Jersey investigation. *Pediatrics*, 108, 1155-1161.
- Berument, S. K., Rutter, M., Lord, C., Pickles, A., & Bailey, A. (1999). Autism screening questionnaire: Diagnostic validity. *British Journal of Psychiatry*, 175, 444–451.
- Bishop, D. V. M., & Baird, G. (2001). Parent and teacher report of pragmatic aspects of communication: Use of the Children's Communication Checklist in a clinical setting. *Developmental Medicine and Child Neurology*, 43, 809-818.
- Bishop, D. V. M., & Norbury, C. F. (2002). Exploring the border-lands of autistic disorder and specific language impairment: A study using standardized instruments. *Journal of Child Psychology and Psychiatry*, 43, 917-930.
- Bolte, S., Dickhut, H., & Poustka, F. (1999). Patterns of parent-reported problems indicative of autism. *Psychopathology*, 32, 93-97.
- Bolte, S., & Poustka, F. (2000). Diagnosis of autism: The connection between current and historical information. *Autism*, 4, 382–390.
- Bolte, S., & Poustka, F. (2002). The relation between general cognitive level and adaptive behavior domains in individuals with autism with and without co-morbid mental retardation. Child Psychiatry and Human Development, 33, 165-172.

- Bracken, B. A. (1998). *The Bracken Basic Concept Scale-Revised*. San Antonio, TX: Psychological Corporation.
- Bristol, M. M. (1984). Family resources and successful adaptation in families of autistic children. In E. Schopler & G. Mesibov (Eds.), The effects of autism on the family (pp. 289–310). New York: Plenum.
- Brownell, R. (2000). Expressive One-Word Picture Vocabulary Test-2000. Novato, CA: Academic Therapy.
- Bruininks, R. H., Woodcock, R. W., Weatherman, R. E., & Hill, B. K. (1996). Scales of Independent Behavior-Revised. Itasca, IL: Riverside.
- Bryan, L. C., & Gast, D. L. (2000). Teaching on-task and on-schedule behaviors to high-functioning children with autism via picture activity schedules. *Journal of Autism and Developmental Disorders*, 30, 553-567.
- Buitelaar, J. K., van der Gaag, R., Klin, A., & Volkmar, F. (1999). Exploring the boundaries of pervasive developmental disorder not otherwise specified: Analyses of data from the DSM-IV autistic disorder field trial. Journal of Autism and Developmental Disorders, 29, 33-43.
- Burack, J. A., Iarocci, G., Bowler, D., & Mottron, L. (2002). Benefits and pitfalls in the merging of disciplines: The example of developmental psychopathology and the study of persons with autism. Development and Psychopathology, 14, 225-237.
- Calloway, C. J., & Simpson, R. (1998). Decisions regarding functions of behavior: Scientific versus informal analyses. Focus on Autism and Other Developmental Disabilities, 13, 167-175.
- Carr, E. G., & Durand, V. M. (1985). Reducing behavior problems through functional communication training. *Journal of Applied Behavior Analysis*, 18, 111-126.
- Carter, A. S., Volkmar, F. R., Sparrow, S. S., Wang, J. J., Lord, C., Dawson, G., et al. (1998). The Vineland Adaptive Behavior Scales: Supplementary norms for individuals with autism. *Journal of Autism and Developmental Disorders*, 28, 287-302.
- Chakrabarti, S., & Fombonne, E. (2001). Pervasive developmental disorders in preschool children. *Journal of the American Medi*cal Association, 285, 3093–3099.
- Charman, T., Howlin, P., Berry, B., & Prince, E. (2004). Measuring developmental progress of children with autism spectrum disorders on school entry using parent report. Autism, 8, 89-100.
- Clark, C., Prior, M., & Kinsella, G. (2002). The relationship between executive function abilities, adaptive behavior, and academic achievement in children with externalizing behavior problems. *Journal of Child Psychology and Psychiatry*, 43, 785-796.
- Cohen, I. L. (2003). Criterion-related validity of the PDD Behavior Inventory. Journal of Autism and Developmental Disorders, 33, 47-54.
- Cohen, I. L., Schmidt-Lackner, S., Romanczyk, R., & Sudhalter, V. (2003). The PDD Behavior Inventory: A rating scale for assessing response to intervention in children with pervasive developmental disorder. *Journal of Autism and Developmental Disor*ders, 33, 31-46.
- Courchesne, E., Karns, C. M., Davis, B. S., Ziccardi, R., Carper, R. A., Tigue, Z. D., et al. (2001). Unusual brain growth patterns in early life in patients with autistic disorder: An MRI study. *Neurology*, 57, 245–254.
- Cox, A., Klein, K., Charman, T., Baird, G., Baron-Cohen, S., Swettenham, J., et al. (1999). Autism spectrum disorders at 20 and 42 months of age: Stability of clinical and ADI-R diagnosis. *Journal of Child Psychology & Psychiatry*, 40, 719-732.
- de Bildt, A., Sytema, S., Ketelaars, C., Kraijer, D., Mulder, E., Volkmar, F., et al. (2004). Interrelationships between the ADOS, ADI-R, and DSM-IV-TR classification in children and adolescents with mental retardation. *Journal of Autism and Developmental Disorders*, 34, 129-137.
- Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). Delis-Kaplan Executive Function System. San Antonio, TX: Psychological Corporation.

- DiLalla, D. L., & Rogers, S. J. (1994). Domains of the Childhood Autism Rating Scale: Relevance for diagnosis and treatment. Journal of Autism and Developmental Disorders, 24, 115–128.
- Duarte, C. S., Bordin, I. A. S., deOliveira, A., & Bird, H. (2003). The CBCL and the identification of children with autism and related conditions in Brazil: Pilot findings. *Journal of Autism and Developmental Disorders*, 33, 703-707.
- Dunlap, G., & Kern, L. (1993). Assessment and intervention for children within the instructional curriculum. In J. Reichler & D. Wacker (Eds.), Communication and language intervention series: Vol. 3. Communicative alternatives to challenging behavior: Integrating functional assessment and intervention strategies (pp. 106-126). San Diego, CA: College-Hill.
- Dunn, L. M., & Dunn, L. M. (1997). Peabody Picture Vocabulary Test (3rd ed.). Circle Pines, MN: American Guidance Service.
- Dunst, C. J., Cooper, C. S., Weeldreyer, J. C., Snyder, K. D., & Chase, J. H. (1988). Family Needs Scale. In C. J. Dunst, C. M. Trivette, & A. G. Deal (Eds.), Enabling and empowering families: Principles and guidelines for practice (pp. 149–163). Cambridge, MA: Brookline.
- Dunst, C. J., Jenkins, V., & Trivette, C. M. (1984). Family support scale: Reliability and validity. *Journal of Individual, Family* and Community Wellness, 1, 45-52.
- Eaves, R. C., & Milner, B. (1993). The criterion-related validity of the Childhood Autism Rating Scale and the Autism Behavior Checklist. *Journal of Abnormal Child Psychology*, 21, 481–491.
- Ehlers, S., & Gillberg, C. (1993). The epidemiology of Asperger syndrome: A total population study. *Journal of Child Psychol*ogy and Psychiatry, 34, 1327–1350.
- Ehlers, S., Gillberg, C., & Wing, L. (1999). A screening questionnaire for Asperger syndrome and other high-functioning autism spectrum disorders in school age children. *Journal of Autism* and Developmental Disorders, 29, 129-141.
- Elliott, C. (1990). Differential Abilities Scale. San Antonio, TX: Psychological Corporation.
- Filipek, P. A., Accardo, P. J., Ashwal, S., Baranek, G. T., Cook, E. H., Jr., Dawson, G., et al. (2000). Practice parameter: Screening and diagnosis of autism. *Neurology*, 55, 468–479.
- Filipek, P. A., Accardo, P. J., Baranek, G. T., Cook, E. H., Jr., Dawson, G., Gordon, B., et al. (1999). The screening and diagnosis of autistic spectrum disorders. *Journal of Autism and De*velopmental Disorders, 29, 439–484.
- Fombonne, E. (2003). Epidemiological surveys of autism and other pervasive developmental disorders: An update. *Journal of Autism and Developmental Disorders*, 33, 365–382.
- Foster, S. L., & Mash, E. J. (1999). Assessing social validity in clinical treatment research: Issues and procedures. *Journal of Consulting and Clinical Psychology*, 67, 308–319.
- Freeman, B. J., Del'Homme, M., Guthrie, D., & Zhang, F. (1999). Vineland Adaptive Behavior Scale scores as a function of age and initial IQ in 210 autistic children. *Journal of Autism and Developmental Disorders*, 29, 379–384.
- Freeman, B. J., Rahbar, B., Ritvo, E. R., Bice, T. L., Yokota, A., & Ritvo, R. (1991). The stability of cognitive and behavioral parameters in autism: A 12-year prospective study. *Journal of the American Academy of Child & Adolescent Psychiatry*, 30, 479-482.
- Frith, U., & Happe, F. (1994). Autism: Beyond theory of mind. Cognition, 50, 115-132.
- Garfin, D. G., McCallon, D., & Cox, R. (1988). Validity and reliability of the Childhood Autism Rating Scale with autistic adolescents. *Journal of Autism and Developmental Disorders*, 18, 367–378.
- Garretson, H. B., Fein, D., & Waterhouse, L. (1990). Sustained attention in children with autism. *Journal of Autism and Developmental Disorders*, 20, 101-114.
- Gilliam, J. E. (1995). Gilliam Autism Rating Scale. Austin, TX: PRO-ED.

- Gilliam, J. E. (2001). Gilliam Asperger Disorder Scale. Austin, TX: PRO-FD
- Gilotty, L., Kenworthy, L., Sirian, L., Black, D. O., & Wagner, A. E. (2002). Adaptive skills and executive function in autism spectrum disorders. *Child Neuropsychology*, 8, 241–248.
- Gioia, G. A., Isquith, P. K., Guy, S. C., & Kenworthy, L. (2000). Behavior rating inventory of executive function. Lutz, FL: Psychological Assessment Resources.
- Glow, R. A., & Glow, P. H. (1980). Peer and self-rating: Children's perception of behavior relevant to hyperkinetic impulse disorder. *Journal of Abnormal Child Psychology*, 8, 397–404.
- Goldstein, S. (2002). Review of the Asperger Syndrome Diagnostic Scale. Journal of Autism and Developmental Disorders, 32, 611-614.
- Grant, D. A., & Berg, E. A. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigle-type card sorting problem. *Journal of Experimental Psychology*, 32, 404-411.
- Gresham, F. M., & Elliot, S. N. (1989). Social skills assessment technology for the learning disabled student. *Learning Disability Quarterly*, 12, 141–152.
- Groth-Marnat, G. (1999). Financial efficacy of clinical assessment: Rational guidelines and issues for future research. *Journal of Clinical Psychology*, 55, 813–824.
- Harris, S. L., & Handleman, J. S. (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 Dis*orders, 30, 137-142.
- Harrison, P., & Oakland, T. (2003). Adaptive Behavior Assessment System (2nd ed.). San Antonio, TX: Psychological Corporation.
- Hauser-Kram, P., Warfield, M. E., Shonkoff, J. P., & Krauss, M. W. (2001). Children with disabilities: A longitudinal study of child development and parent well-being. Monographs of the Society for Research in Child Development, 66, 1-131.
- Hayes, S. C., Nelson, R. O., & Jarrett, R. B. (1987). The treatment utility of assessment: A functional approach to evaluating assessment quality. *American Psychologist*, 42, 963-974.
- Heaton, R. K., Chelune, G. J., Talley, J. L., Kay, G. G., & Curtiss, G. (1993). Wisconsin Card Sorting Test manual: Revised and expanded. Odessa, FL: Psychological Assessment Resources.
- Hinshaw, S. P., & Melnick, S. M. (1995). Peer relationships in boys with attention deficit hyperactivity disorder with and without comorbid aggression. *Developmental Psychopathology*, 7, 627-647.
- Holroyd, J. (1974). The Questionnaire on Resources and Stress: An instrument to measure family response to a handicapped member. *Journal of Humanistic Psychology*, 37, 1–11.
- Horner, R. H. (1994). Functional assessment: Contributions and future directions. *Journal of Applied Behavior Analysis*, 27, 401–404.
- Howlin, P. (2003). Outcome in high-functioning adults with autism with and without early language delays: Implications for the differentiation between autism and Asperger syndrome. *Journal of Autism and Developmental Disorders*, 33, 3–14.
- Hresko W. P., Peak, P. K., Herron, S. R., & Bridges, D. L. (2000). Young Children's Achievement Test. Austin, TX: PRO-ED.
- International Molecular Genetic Study of Autism Consortium. (2001). A genomewide screen for autism: Strong evidence for linkage to chromosomes 2q, 7q, and 16p. American Journal of Human Genetics, 69, 570–581.
- Joseph, R. M., Tager-Flusberg, H., & Lord, C. (2002). Cognitive profiles and social-communicative functioning in children with autism spectrum disorders. *Journal of Child Psychology and Psychiatry*, 43, 807–821.
- Kamphaus, R. W., Reynolds, C. R., & Hatcher, N. M. (1999). Treatment planning and evaluation with the Behavior Assessment System for Children. In M. E. Maruish (Ed.), The use of psychological testing for treatment planning and outcomes assessment

- (2nd ed., pp. 563-597). Mahwah, NJ: Lawrence Erlbaum Associates. Inc.
- Kanner, L. (1943). Autistic disturbances of affective content. Nervous Child, 2, 217–250.
- Kasari, C. (2002). Assessing change in early intervention programs for children with autism. *Journal of Autism and Developmental Disorders*, 32, 447–461.
- Kazdin, A. E. (1999). The meanings and measurement of clinical significance. *Journal of Consulting and Clinical Psychology*, 67, 332-339.
- Kendall, P. C., Marrs-Garcia, A., Nath, S. R., & Sheldrick, R. C. (1999). Normative comparisons for the evaluation of clinical significance. *Journal of Consulting and Clinical Psychology*, 67, 285-299.
- Kim, J. A., Szatmari, P., Bryson, S. E., Streiner, D. L., & Wilson, F. J. (2000). The prevalence of anxiety and mood problems among children with autism and Asperger syndrome. *Autism*, 4, 117-132.
- Klin, A., & Shepard, B. A. (1994). Psychological assessment of autistic children. Child and Adolescent Psychiatric Clinics of North America, 3, 53-70.
- Klin, A., Sparrow, S. S., Marans, W. D., Carter, A., & Volkmar, F. R. (2000). Assessment issues in children and adolescents with Asperger syndrome. In A. Klin, F. R. Volkmar, & S. S. Sparrow (Eds.), Asperger syndrome (pp. 309–339). New York: Guilford.
- Klin, A., Volkmar, F. R., Sparrow, S. S., Cicchetti, D. V., & Rourke, B. P. (1995). Validity and neuropsychological characterization of Asperger syndrome: Convergence with nonverbal learning disabilities syndrome. *Journal of Child Psychology and Psychi*atry, 36, 1127–1140.
- Klinger, L. G., & Renner, P. (2000). Performance-based measures in autism: Implications for diagnosis, early detection, and identification of cognitive profiles. *Journal of Clinical Child Psychol*ogy, 29, 479–492.
- Koegel, L. K., Koegel, R. L., & Smith, A. (1997). Variables related to differences in standardized test outcomes for children with autism. *Journal of Autism and Developmental Disorders*, 27, 233-243.
- Konstantareas, M. M., Homatidis, S., & Plowright, C. M. (1992). Assessing resources and stress in parents of severely dysfunctional children through the Clarke modification of Holroyd's Questionnaire on Resources and Stress. *Journal of Autism and Developmental Disorders*, 22, 217–234.
- Korkman, M., Kirk, U., & Kemp, S. (1998). NEPSY manual: A developmental neuropsychological assessment. San Antonio, TX: Psychological Corporation.
- Kovacs, M. (1992). Children's Depression Inventory. North Tonawanda, NY: Multi-Health Systems.
- Kraemer, B. R., McIntyre, L. L., & Blacher, J. (2003). Quality of life for young adults with mental retardation during transition. *Mental Retardation*, 41, 250–262.
- Krug, D. A., Arick, J. R., & Almond, P. J. (1988). Autism Behavior Checklist. Austin, TX: PRO-ED.
- Lainhart, J. E. (1999). Psychiatric problems in individuals with autism, their parents and siblings. *International Review of Psychiatry*, 11, 278–298.
- Lainhart, J. E., & Folstein, S. E. (1994). Affective disorders in people with autism: A review of published cases. *Journal of Autism* and *Developmental Disorders*, 24, 587-601.
- Lincoln, A. J., Allen, M. H., & Kilman, A. (1995). The assessment and interpretation of intellectual abilities in people with autism. In E. Schopler & G. B. Mesibov (Eds.), Learning and cognition in autism (pp. 89–117). New York: Plenum.
- Lord, C. (1995). Follow-up of two-year-olds referred for possible autism. *Journal of Child Psychology and Psychiatry*, 36, 1365-1382.
- Lord, C. (1997). Diagnostic instruments in autism spectrum disorders. In D. J. Cohen & F. R. Volkmar (Eds.), Handbook of au-

- tism and pervasive developmental disorders (2nd ed., pp. 460-483). New York: Wiley.
- Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Leventhal, B. L., DiLavore, P. C., et al. (2000). The Autism Diagnostic Observation Schedule-Generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders*, 30, 205-223.
- Lord, C., Rutter, M., DiLavore, P. C., & Risi, S. (2002). Autism Diagnostic Observation Schedule manual. Los Angeles: Western Psychological Services.
- Lord, C., Rutter, M., & LeCouteur, A. (1994). Autism Diagnostic Interview–Revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, 24, 659–685.
- Lord, C., & Schopler, E. (1989). The role of age at assessment, developmental level, and test in the stability of intelligence scores in young autistic children. *Journal of Autism and Developmental Disorders*, 19, 483-499.
- Lotter, V. (1966). Epidemiology of autistic conditions in young children: Prevalence. *Social Psychiatry, 1*, 124–137.
- Lotter, V. (1974). Factors related to outcome in autistic children. Journal of Autism and Childhood Schizophrenia, 4, 263–277.
- Luiselli, J. K., Campbell, S., Cannon, B., DiPietro, E., Ellis, J. T., Taras, M., et al. (2001). Assessment instruments used in the education and treatment of persons with autism: Brief report of a survey of national service centers. Research in Developmental Disabilities, 22, 389–398.
- Macintosh, K. E., & Dissanayake, C. (2004). The similarities and differences between autistic disorder and Asperger disorder: A review of the empirical evidence. *Journal of Child Psychology* and Psychiatry, 45, 410-434.
- Magiati, I., & Howlin, P. (2001). Monitoring the progress of preschool children with autism enrolled in early intervention programs: Problems in cognitive assessment. Autism, 5, 399-406.
- March, J. S. (1997). Multidimensional Anxiety Scale for Children. North Tonawanda, NY: Multi-Health Systems.
- Mash, E. J., & Hunsley, J. (2005). Evidence-based assessment of child and adolescent disorders: Issues and challenges. *Journal* of Clinical Child and Adolescent Psychology, 34, 362–379.
- Mather, N., & Woodcock, R. W. (2001). Woodcock-Johnson III Tests of Achievement. Chicago: Riverside.
- Mayes, S., & Calhoun, S. (2003a). Ability profiles in children with autism: Influence of age and IQ. Autism, 6, 83–98.
- Mayes, S. D., & Calhoun, S. L. (2003b). Analysis of WISC-III, Stanford-Binet: IV, and academic achievement test scores in children with autism. *Journal of Autism and Developmental Disorders*, 33, 329-341.
- Mildenberger, K., Sitter, S., Noterdaeme, M., & Amorosa, H. (2001). The use of the ADI-R as a diagnostic tool in the differential diagnosis of children with infantile autism and children with receptive language disorder. European Child and Adolescent Psychiatry, 10, 248-255.
- Miller, J. N., & Ozonoff, S. (2000). The external validity of Asperger disorder: Lack of evidence from the domain of neuropsychology. *Journal of Abnormal Psychology*, 109, 227–238.
- Minshew, N. J., Goldstein, G., & Siegel, D. J. (1997). Neuropsychologic functioning in autism: Profile of a complex information processing disorder. *Journal of the International Neuro*psychological Society, 3, 303–316.
- Mullen, E. M. (1995). Mullen Scales of Early Learning. Circle Pines, MN: American Guidance Service.
- 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
- Myles, B. S., Bock, S. J., & Simpson, R. (2001). Asperger Syndrome Diagnostic Scale. Austin TX: PRO-ED.

- Noterdaeme, M., Amorosa, H., Mildenberger, K., Sitter, S., & Minow, F. (2001). Evaluation of attention problems in children with autism and children with a specific language disorder. European Child and Adolescent Psychiatry, 10, 58-66.
- Offord, D., Boyle, M., Racine, Y., Szatmari, P., Fleming, J., Sanford, M., et al. (1996). Integrating assessment data from multiple informants. *Journal of the American Academy of Child & Adolescent Psychiatry*, 35, 1078-1085.
- Olsson, M. B., & Hwang, C. P. (2001). Depression in mothers and fathers of children with intellectual disability. *Journal of Intellectual Disability Research*, 45, 535-543.
- O'Neill, R. E., Horner, R. H., Albin, R. W., Storey, K., & Srague, J. R. (1997). Functional assessment and program development for problem behavior (2nd ed.). Pacific Grove, CA: Brooks/Cole.
- Owley, T., McMahon, W., Cook, E. H., Laulhere, T. M., South, M., Mays, L. Z., et al. (2001). A multi-site, double-blind, placebo-controlled trial of porcine secretin in autism. *Journal of the American Academy of Child & Adolescent Psychiatry*, 40, 1293-1299.
- Ozonoff, S. (1995). Reliability and validity of the Wisconsin Card Sorting Test in studies of autism. *Neuropsychology*, 9, 491–500.
- Ozonoff, S., Dawson, G., & McPartland, J. (2002). A parent's guide to Asperger syndrome and high-functioning autism: How to meet the challenges and help your child thrive. New York: Guilford.
- Ozonoff, S., & Griffith, E. M. (2000). Neuropsychological function and the external validity of Asperger syndrome. In A. Klin, F. Volkmar, & S. S. Sparrow (Eds.), Asperger syndrome (pp. 72-96). New York: Guilford.
- Ozonoff, S., Provencal, S., & Solomon, M. (2002, October). The effectiveness of social skills training programs for autism spectrum disorders. Paper presented at the annual meeting of the American Academy of Child and Adolescent Psychiatry, San Francisco.
- Ozonoff, S., Rogers, S. J., & Pennington, B. F. (1991). Asperger syndrome: Evidence of an empirical distinction from high-functioning autism. *Journal of Child Psychology and Psychiatry*, 32, 1107-1122.
- Ozonoff, S., South, M., & Miller, J. (2000). DSM-IV-defined Asperger syndrome: Cognitive, behavioral and early history differentiation from high-functioning autism. Autism, 4, 29-46.
- Ozonoff, S., South, M., & Provencal, S. (2005). Executive functions. In F. R. Volkmar, A. Klin, & R. Paul (Eds.), *Handbook of autism and pervasive developmental disorders* (3rd ed., pp. 606–627). New York: Wiley.
- Pennington, B. F., & Ozonoff, S. (1996). Executive functions and developmental psychopathologies. *Journal of Child Psychology and Psychiatry*, 37, 51–87.
- Perry, D. W., Marston, G. W., Hinder, S. A., Munden, A. C., & Roy, A. (2001). The phenomenology of depressive illness in people with learning disability and autism. Autism, 5, 265-275.
- Perry, R. (1998). Misdiagnosed ADD/ADHD; rediagnosed PDD. Journal of the American Academy of Child & Adolescent Psychiatry, 37, 113–114.
- Piotrowski, C. (1999). Assessment practices in the era of managed care: Current status and future directions. *Journal of Clinical Psychology*, 55, 787-796.
- Prior, M. (2000). Guest editorial: Special issue on Asperger syndrome. Autism, 4, 5-8.
- Roid, G. H. (2003). Stanford–Binet Intelligence Scale manual (5th ed.). Itasca, IL: Riverside.
- Roid, G., & Miller, L. (1997). Leiter International Test of Intelligence-Revised. Chicago: Stoelting.
- Rourke, B. P. (Ed.). (1995). Syndrome of nonverbal learning disabilities: Neurodevelopmental manifestations. New York: Guilford.
- Russell, J. (Ed.). (1997). Autism as an executive disorder. New York: Oxford University Press.

- Rutter, M. (1984). Autistic children growing up. Developmental Medicine and Child Neurology, 26, 122–129.
- Rutter, M., Bailey, A., Berument, S. K., LeCouteur, A., Lord, C., et al. (2003). Social Communication Questionnaire (SCQ) manual. Los Angeles: Western Psychological Services.
- Rutter, M., LeCouteur, A., & Lord, C. (2003). Autism Diagnostic Interview— Revised manual. Los Angeles: Western Psychological Services.
- Saemundsen, E., Magnusson, P., Smari, J., & Sigurdardottir, S. (2003). Autism Diagnostic Interview-Revised and the Child-hood Autism Rating Scale: Convergence and discrepancy in diagnosing autism. *Journal of Autism and Developmental Disorders*, 33, 319-328.
- Schalock, R. L., & Keith, K. D. (1993). Quality of Life Questionnaire manual. Columbus, OH: IDS.
- Schopler, E., Reichler, R. J., Bashford, A., Lansing, M. D., & Marcus, L. M. (1990). Psychoeducational Profile-Revised (PEP-R). Austin, TX: PRO-ED.
- Schopler, E., Reichler, R., & Renner, B. (1988). The Childhood Autism Rating Scale (CARS). Los Angeles: Western Psychological Services.
- Schultz, R. T., Gauthier, I., Klin, A., Fulbright, R. K., Anderson, A. W., Volkmar, F., et al. (2000). Abnormal ventral temporal cortical activity during face discrimination among individuals with autism and Asperger syndrome. Archives of General Psychiatry, 57, 331–340.
- Semel, E., Wiig, E. H., & Secord, W. A. (2003). Clinical evaluation of language fundamentals (4th ed.). San Antonio, TX: Psychological Corporation.
- Sevin, J. A., Matson, J. L., Coe, D. A., Fee, V. E., & Sevin, B. M. (1991). A comparison and evaluation of three commonly used autism scales. *Journal of Autism and Developmental Disorders*, 21, 415-432.
- Sheras, P. L., Abidin, R. R., & Konold, T. R. (1998). The Stress Index for Parents of Adolescents: Professional manual. Odessa, FL: Psychological Assessment Resources.
- Siegel, D. J., Minshew, N. J., & Goldstein, G. (1996). Wechsler IQ profiles in diagnosis of high-functioning autism. *Journal of Autism and Developmental Disorders*, 26, 389-406.
- South, M., Williams, B. J., McMahon, W. M., Owley, T., Filipek, P. A., Shernoff, E., et al. (2002). Utility of the Gilliam Autism Rating Scale in research and clinical populations. *Journal of Autism and Developmental Disorders*, 32, 593-599.
- Sparrow, S., Balla, D., & Cicchetti, D. (1984). Vineland Adaptive Behavior Scales. Circle Pines, MN: American Guidance Service.
- Sponheim, E. (1996). Changing criteria of autistic disorders: A comparison of ICD-10 research criteria and DSM-IV with DSM-III-R, CARS, and ABC. Journal of Autism and Developmental Disorders, 26, 513-525.
- Sponheim, E., & Spurkland, I. (1996). Diagnosing childhood autism in clinical practice: An inter-rater reliability study of ICD-10, DSM-III-R, Childhood Autism Rating Scale, and Autism Behavior Checklist. Nordic Journal of Psychiatry, 50, 5-9.
- Starr, E., Szatmari, P., Bryson, S., & Zwaigenbaum, L. (2003). Stability and change among high-functioning children with pervasive developmental disorders: A 2-year outcome study. *Journal of Autism and Developmental Disorders*, 33, 15-22.
- Stevens, M. C., Fein, D. A., Dunn, M., Allen, D., Waterhouse, L. H., Feinstein, C., et al. (2000). Subgroups of children with autism by cluster analysis: A longitudinal examination. *Journal of the American Academy of Child & Adolescent Psychiatry*, 39, 346-352.
- Stone, W. L., Coonrod, E. E., Pozdol, S. L., & Turner, L. M. (2003). The Parent Interview for Autism-clinical version (PIA-CV): A measure of behavioral change for young children with autism. Autism, 7, 9-30.
- Stone, W. L., & Yoder, P. J. (2001). Predicting spoken language level in children with autism spectrum disorders. *Autism*, 5, 341–361.

- Szatmari, P., Archer L., Fisman, S., & Streiner, D. L. (1994). Parent and teacher agreement in the assessment of pervasive developmental disorders. *Journal of Autism and Developmental Disor*ders, 24, 703-717.
- Szatmari, P., Bartolucci, G., Bremner, R., Bond, S., & Rich, S. (1989). A follow-up study of high-functioning autistic children. Journal of Autism and Developmental Disorders, 19, 213-225.
- Szatmari, P., Bryson, S. E., Boyle, M. H., Streiner, D. L., & Duku, E. (2003). Predictors of outcome among high functioning children with autism and Asperger syndrome. *Journal of Child Psychol*ogy and Psychiatry, 44, 520-528.
- Tager-Flusberg, H., & Joseph, R. M. (2003). Identifying neurocognitive phenotypes in autism. *Philosophical Transactions of* the Royal Society of London, 358, 303-314.
- Tobing, L. E., & Glenwick, D. S. (2002). Relation of the Childhood Autism Rating Scale-Parent version to diagnosis, stress, and age. Research in Developmental Disabilities, 23, 211-223.
- Tsatsanis, K. D., Dartnall, N., Cicchetti, D., Sparrow, S. S., Klin, A., & Volkmar, F. R. (2003). Concurrent validity and classification accuracy of the Leiter and Leiter–R in low-functioning children with autism. *Journal of Autism and Developmental Disorders*, 33, 23–30.
- Van Bourgondien, M., Marcus, L., & Schopler, E. (1992). Comparison of DSM-III-R and Childhood Autism Rating Scale diagnoses of autism. *Journal Autism & Developmental Disorders*, 22, 493-506.
- Venter, A., Lord, C., & Schopler, E. (1992). A follow-up study of high-functioning autistic children. *Journal of Child Psychology* and Psychiatry, 33, 489-507.
- Volkmar, F. R., Cicchetti, D. V., Dykens, E., Sparrow, S. S., Leckman, J. F., & Cohen, D. J. (1988). An evaluation of the Autism Behavior Checklist. *Journal of Autism and Developmental Disorders*, 16, 81-97.
- Volkmar, F. R., Cook, E. H., Jr., Pomeroy, J., Realmuto, G., & Tanguay, P. (1999). Practice parameters for the assessment and treatment of children, adolescents, and adults with autism and other pervasive developmental disorders. *Journal of the*

- American Academy of Child & Adolescent Psychiatry, 38, 32S-54S.
- Wadden, N. P., Bryson, S. E., & Rodger, R. S. (1991). A closer look at the Autism Behavior Checklist: Discriminant validity and factor structure. *Journal of Autism and Developmental Disor*ders, 21, 529-541.
- Walker, D. R., Thompson, A., Zwaigenbaum, L., Goldberg, J., Bryson, S. E., Mahoney, W. J., et al. (2004). Specifying PDDNOS: A comparison of PDDNOS, Asperger syndrome, and autism. Journal of the American Academy of Child & Adolescent Psychiatry, 43, 172-180.
- Wechsler, D. (2002). Wechsler Individual Achievement Test (2nd ed.). San Antonio, TX: Psychological Corporation.
- Wechsler, D. (2003). Wechsler Intelligence Scale for Children (4th ed.). San Antonio, TX: Psychological Corporation.
- Wiig, E., & Secord, W. (1989). Test of Language Competence-expanded edition. San Antonio, TX: Psychological Corporation.
- Wing, L., & Gould, J. (1979). Severe impairments of social interaction and associated abnormalities in children. *Journal of Autism and Developmental Disorders*, 9, 11–29.
- Wolery, M., & Garfinkle, A. N. (2002). Measures in intervention research with young children who have autism. *Journal of Autism and Developmental Disorders*, 32, 463–478.
- Wolf, L., Noh, S., Fisman, S. N., & Speechley, M. (1989). Psychological effects of parenting stress on parents of autistic children. Journal of Autism and Developmental Disorders, 19, 157–166.
- Wood, M. (1995). Parent-professional collaboration and the efficacy of the IEP process. In R. L. Koegel & L. K. Koegel (Eds.), Teaching children with autism (pp. 147-174). Baltimore: Brookes.
- Zimmerman, I. L., Steiner, V. G., & Pond, R. E. (2002). Preschool Language Scales 4. San Antonio, TX: Psychological Corporation.

Received October 15, 2003 Accepted July 10, 2004 Copyright of Journal of Clinical Child & Adolescent Psychology is the property of Lawrence Erlbaum Associates and its content may not be copied or emailed to multiple sites or posted to a listsery without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.