

RUNNING HEAD: EVIDENCE-BASED ASSESSMENT FOR ASD

Evidence-Based Assessment for Autism Spectrum Disorders
Jessica Weeks
University of Utah

Superheroes Social Skills Training, Rethink Autism Internet Intervention, Parent Training, Evidence-Based Practices, Classroom Training, Functional Behavior Assessment: An Autism Spectrum Disorder, Evidence-Based Practices Training Track for School Psychologists

US Office of Education 84.325K
H325K12306

Principal Investigators:

William R. Jenson, PhD
Elaine Clark, PhD

Grant Director:

Julia Hood

University of Utah
Department of Educational Psychology
School Psychology Program

Manuscript Edited By: Alicia Stephens _____ Date: 10/15/2013 _____

Introduction

The assessment of Autism Spectrum Disorders is a complex and imperative issue. Appropriate and relevant assessment measures are fundamental to evidence-based assessment and practice requirements. Key features of acceptable assessment guidelines include the use of valid and reliable measures along with relevance of the assessment for the goals of the individual (Ozonoff, Goodlin-Jones, & Solomon, 2005). It is critical that assessment decisions are made to best inform intervention procedures. Especially regarding the assessment and diagnosis of autism spectrum disorder, utilization of psychometrically sound tools and applicable information gathering is crucial for optimum intervention at the soonest point possible. With the development of various assessment methods and diagnostic scales, it is the responsibility of clinicians (psychologists) to be informed of valid and reliable measures to ensure the most accurate assessment procedure. Additionally, the long-term benefit of the child should always guide this practice.

Evidence-Based Assessment

Best Practice procedures in various fields require the adherence to evidence-based practice. In assessment, methods that meet evidence-based standards are both valid and reliable. Diagnostic validity of a measure refers to its ability to accurately predict group membership, or the correct diagnosis. The correct identification of an individual having a disorder relates to a scale's sensitivity, with more sensitive scales missing fewer true cases. Specificity, on the other hand, refers to the percentage of cases without a disorder that screen negative, or accurately identify those that do not have a disorder. Recommended levels of sensitivity are between .70 and .80 or higher, and .80 or higher

for specificity, as the mistake of misidentification (also referred to as a ‘false negative’) can have significant and lasting consequences for an individual (Norris & Lecavalier, 2010). The predictive value of screening tools and assessments should also guide the diagnostic process as it describes how well a screening tool performs. The positive predictive value (PPV) is the percentage of all identified positive cases that in fact have the disorder. The negative predictive value (NPV) is the percentage of all cases that screened negative that in fact do not have the disorder. The prevalence rate of a disorder also has an effect on the PPV and NPV of a measure. It follows that when the prevalence rate of a disorder in any given population is low, the PPV is also low. Alternately, when the measure is used in a population where the prevalence of the disorder is higher, the PPV follows suit (Wilkinson, 2011). As previously noted, it is crucial that the clinician is familiar with the psychometric properties of any given scale or test being used. Misdiagnosis can tragically waste critical time for intervention.

Autism Developments

Autism research has made great progress over the past couple of decades in terms of understanding prevalence rates and developing accurate assessment measures to meet updated diagnostic criteria. Recently, a noticeable increase in the prevalence rate of autism spectrum disorder has made some question the accuracy of the methods used as well as the clinicians responsible for assigning diagnoses. In reality, a better understanding of the broader autism phenotype, along with a renewed emphasis on early screening methods, and a fuller conceptual grasp of less overt symptoms, contributes to the increase in ASD cases. Still, the overidentification (also referred to as false positives)

or misidentification of the disorder are very real threats, regardless of the progress the field has seen. As both over and under diagnosis of the disorder are problematic, a standard approach to assessment, employing relevant measures accurately, is of critical importance. Paralleling the development of autism assessment measures is the development of the actual diagnostic criteria. With the revisions of the Diagnostic and Statistical Manual of Mental Disorders, 4th Edition, Text Revised (DSM-IV-TR; American Psychiatric Association, 1994) the diagnostic criteria for autism spectrum disorder was more clearly laid out. Now, despite continued controversy with the DSM-5 (American Psychiatric Association, 2013), it aims to lessen the overdiagnosis even more by making criteria more stringent.

As this paper will review various assessment measures, including screening measures, checklists, and questionnaires, an initial caution implores that no single method should ever be used independently to diagnose Autism Spectrum Disorder, or any disorder. The following review will highlight the importance of parental interviews and the invaluable information one can gain through that process. In addition, the necessity of direct observation of the individual's behavior will be explained in terms of a structured assessment and the standardized conditions that are introduced. These methods combine a skilled clinician's keen eye for autism diagnosis as well as the knowledge that parents have of typical behaviors of their child. After compiling a host of information, a thorough assessment should include a description of the child's social behavior, language development, intellectual ability, comorbid medical, developmental, and psychological conditions (Ozonoff, Goodlin-Jones, & Solomon, 2005).

Also, due to intervention research demonstrating staggering improvement among individuals with ASD, it is well known that early intervention in evidence-based practice is fundamental in improving the trajectory of ASD (Lovaas, 1987; Rogers & Vismara, 2008). A critical window for intervention with children with ASD is suggested to be before the age of seven, with great improvement seen between the ages of 3 and 7. With this knowledge of the importance of early intervention, early and accurate diagnosis of ASD is essential. Researchers who understand this exigent need continue to work to develop more accurate, efficient, and reliable measures of autism assessment. This paper will review the established considerations of autism diagnosis, standardized screening procedures, specific autism assessment tools and the psychometric and clinical properties of these measures, along with the current gold standard for autism assessment.

Assessment Considerations

Experts on autism assessment have provided specific considerations to inform the appropriate process for this assessment. The assessment relies on two main sources of information: (1) description of course of development and current behavior patterns from caregivers and (2) direct observation of behavior (Lord, Rutter, Goode, Heemsbergen, Jordan, Mawhood, & Schopler, 1989). Throughout this process, there are three cardinal features of autism that clinicians should note: (1) impairment in social interaction, (2) impairment in communication, and (3) repetitive and stereotyped patterns of behavior and interest (American Academy of Neurology, 2000; Ozonoff, Goodlin-Jones, & Solomon, 2010). Throughout the assessment, it is necessary to remember the developmental perspective of autism. The disorder is lifelong and characterized by uneven development

across areas. Also, multiple sources of information should be used to inform the process and obtain a thorough picture of the child's development and functioning. It is also ideal for the assessment to be multidisciplinary whenever possible. The expertise of many fields can contribute greatly to creating a more comprehensive picture of the individual that will lead to better rounded intervention methods to address the various impairments characteristic of the disorder (Ozonoff, Goodlin-Jones, & Solomon, 2010; Norris & Lecavalier, 2010).

Screening and Diagnosis of Autism

Level One: Routine Developmental Surveillance

Due to the prevalence rate of autism along with the importance of identifying children early enough to intervene during the critical period of development, the American Academy of Neurology (Filipek et al., 2000) and the American Academy of Child and Adolescent Psychiatry (Volkmar, Cook, Pomeroy, Realmuto, & Tanguay, 1999) recommend that developmental surveillance be performed at all pediatric well-child visits from infancy through school-age. These screenings should take place at least at 18 months and 24 or 30 months; however, the typical well-child check-up schedule includes visits at 1, 2, 4, 6, 9, 12, 18, 24, and 30 months of age. Along with physical health concerns, including the child's height and weight, the pediatrician should pay special attention to when the child reaches specific developmental milestones. If concerns arise in typical development, social acceptance, learning, or behavior, more narrow screening methods should be employed. Appropriate screening instruments for this level

include the Ages and Stages Questionnaire, the BRIGANCE Screens, the Child Development Inventories (CDI), and the Parents' Evaluation of Developmental Status (PEDS) (Filipek, Accardo, & Baranek, 1999). These measures aid in identifying any kind of atypical development; the process then continues in order to specifically identify those at risk for autism.

Further evaluation is necessary for those children who fail to babble by 12 months, use single words by 16 months, use two-word spontaneous phrases by 24 months (not echolalic), and those children who display any loss of language or social skills at any age. Special attention should be paid to siblings of children with autism spectrum disorder, as the genetic contribution of the disorder is high. The prevalence rate for siblings of children with autism is 4%, 40 times greater than that of the general population (0.1%-0.16%). One should watch for social, communication, and play skills as well as the introduction of any form of maladaptive behavior. If a child fails one of the specified routine developmental surveillance procedures, specific autism screening should take place using either the Checklist for Autism in Toddlers (CHAT), (Baron-Cohen, et al., 1992, 1996) or the Autism Screening Questionnaire (Berument, Rutter, Lord, Pickles, & Bailey, 1999), both of which are validated instruments. If scores on one of these screening measures are elevated, level two diagnosis and evaluation of autism should be implemented (California Department of Developmental Services, 2002; Filipek et al. 1999; Johnson & Myers, 2007; Ozonoff et al. 2005; Volkmar et al. 1999).

Level Two: Diagnosis and Evaluation of Autism

At this level of assessment, a comprehensive autism evaluation needs to be conducted by a clinician specifically trained in autism spectrum disorders. The assessment of autism at any age should involve a methodical, careful procedure. Especially at very young ages, it is sometimes exceptionally difficult to discriminate an autism spectrum disorder from other developmental disorders of similar presentation, including intellectual disability, hearing impairments, speech and language disorders, and profound neglect (Rogers, 2001). For this reason, only those familiar with autism, having received training in the tests and measures, should continue the assessment.

As many assessment measures have been developed to date, a number of individuals in the field of autism assessment have agreed on a general best practice procedure for a comprehensive autism evaluation (California Department of Developmental Services 2002; Filipek et al. 1999; Johnson & Myers, 2007; National Research Council 2001; Ozonoff et al. 2005; Volkmar, Cook, Pomeroy, Realmuto, & Tanguay, 1999; Wilkinson, 2010). As previously mentioned, an autism evaluation necessitates the use of interviews and observation schedules. These pieces of the assessment provide the base of the evaluation. From this point, interdisciplinary assessment of social behavior, language and communication, adaptive behavior, motor skills, sensory issues, atypical behaviors, and cognitive functioning is conducted (National Research Council, 2001; Ozonoff et al., 2005). To obtain this information in an organized manner, listed below are the core autism assessment domains for best practice procedures:

- a. Record review
- b. Developmental and medical history

- c. Medical screening and/or evaluation
- d. Parent/caregiver interview
- e. Parent/teacher ratings of social competence
- f. Direct child observation
- g. Cognitive assessment
- h. Academic assessment
- i. Adaptive behavioral assessment
- j. Communication and language assessment

(Wilkinson, 2010).

The inclusion of various measures and tests in each evaluation should be based on the relevance and application of the individual client's needs and what services are necessary for the person to best function with the disorder. The specific approach of the assessment depends on the goal for the individual (Mash & Hunsley, 2005; Ozonoff et al., 2005). That being said, of the plethora of assessment tools, the psychometric properties and correlation data are available and should also inform assessment choices. Based on consistent research findings, the "Gold Standard" for autism assessment includes the combination of the Autism Diagnostic Interview-Revised (ADI- R; Lord, Rutter, & Le Couteur, 1994) and the Autism Diagnostic Observation Schedule-2 (ADOS-2; Lord, Rutter, DiLavore, Risi, Gotham, & Bishop, 2012).

GOLD STANDARDS

Standardized Parent Interview

The Autism Diagnostic Interview, Revised Edition (ADI-R; Rutter, Le Couteur, & Lord, 2003b) is the most reliable standardized measure to obtain an early developmental history of an individual suspected of having an autism spectrum disorder. It is considered the “gold standard” for a parent interview and identifies symptoms linked to the Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition, Text Revision (DSM-IV-TR; American Psychiatric Association, 1994) and ICD-10 criteria (International Statistical Classification of Diseases and Related Health Problems-10th Revision; World Health Organization [WHO], 1992; Lord and Corsello, 2005). The ADI-R is also the most widely used diagnostic instrument in autism research, and is itself research-derived. The ADI-R comprises 93 questions that are summed into 3 functional domains, the same domains used as diagnostic criteria in the DSM-IV-TR. They are:

- 1.) Language/Communication
- 2.) Reciprocal Social Interactions
- 3.) Restricted, Repetitive, and Stereotyped Behavior and Interest

The ADI-R interview questions mainly focus on the developmental time interval from age four to five of the individual suspected as having an autism spectrum disorder. This time period is of value due to the social-focus of many of the developmental norms of this age. By ages four and five, children are introduced to an increasing amount and complexity of social situations. Many children this age go to preschool, have play dates, and normally have enough language to carry on a conversation with age-mates as well as adults. The nature of these social situations highlights delays and discrepancies in language ability and social behaviors that may not have been as apparent before.

The particular information gathered from the interview assists a skilled interpreter in discriminating between ASD and other disorders such as severe receptive language disorders and general developmental delays. Some limitations of the ADI-R have been suggested. With the previous DSM-IV-TR diagnostic break down of Autistic Disorder, Asperger's Syndrome, and Pervasive Developmental Disorder, the ADI-R did not provide established scoring thresholds for the ASD subtypes. Although the ADI-R algorithm produces independent scores for the diagnostic domains of reciprocal social interaction, communication, and restricted, repetitive, and stereotyped patterns of behavior, the scores are then only measured against one diagnostic cutoff score. The interview algorithm does not suggest cutoffs for Asperger's Syndrome or Pervasive Developmental Disorder, which was previously viewed as a limitation; however, the diagnostic updates of the DSM-5 remove this specific limitation as autism spectrum disorder now exists on a continuum of severity depending on the specific support required by the deficits each individual experiences (American Psychiatric Association, 2013). Research has also shown that the ADI-R is not sensitive to differences among children with mental ages below 20 months or Intelligence Quotients (IQs) below 20 and the ADI-R is not at all recommended for use with individuals with a mental age below 18 (Cox, Klein, Charman, Baird, Baron-Cohen, Swettenham, Drew & Wheelwright, 1999; Lord, Pickles, McLennan, Rutter, Bregman, Folstein, Fombonne, Leboyer, & Minshew, 1997). In individuals with IQs above 20, however, the ADI-R shows great specificity in identifying autism over intellectual disability (de Bildt, Systema, Ketelaars, Kraijer, Mulder, Volkmar, & Minderaa, 2004).

Curious as to whether the ADI-R brought to light differences in autism across gender, Pilowsky, Yirmiya, Shulman, and Dover (1998) matched males and females on chronological age and mental ability and reviewed their ADI-R profiles. In this study, the only significant difference between genders was demonstrated in mental age, or the age at which a child is performing intellectually. As a group, the mental age of males with autism was higher than that of females with autism. Other than this difference, the clinical presentation of males and females with autism on the ADI-R appeared comparable; however, further research concerning gender differences is greatly needed.

The structure of the diagnostic interview is to confirm an initial diagnosis of autism, but does not assess change throughout development. Due to the quality of the interview and research support of the ADI-R, it is not a surprise that administration is lengthy and labor intensive, requiring 2 ½-3 hours to complete. Also, while some view the training required to administer the ADI-R as a limitation, this limit ensures more careful use of the measure and insists on specific training and familiarity with the diagnosis of autism.

Standardized Observation Schedule

The Autism Diagnostic Observation Schedule, second edition (ADOS-2) is the second half of the “gold standard” in autism evaluation. While the ADI-R covers developmental history of an individual, the ADOS-2 is a direct observation assessment. In a standardized context, the administrator conducts a number of “presses” that elicit spontaneous behaviors from the client. The ADOS-2 includes five Module options to choose from depending on the language level of the individual being assessed. With the

introduction of the Toddler module, the ability level of the ADOS-2 now spans all levels of functioning. The modules are as follows:

-Module 1- Children 31 months and older who do not consistently use phrase speech.

-Module 2- Children of any age who use phrase speech but are not verbally fluent. Focus on social interest, joint attention, communicative behaviors, symbolic play, and atypical behaviors (excessive sensory interests and hand mannerisms)

-Module 3- Verbally fluent children and young adolescents.

-Module 4- Verbally fluent older adolescents and adults.

Focus on conversational reciprocity, empathy, insight into social relationships, and special interests.

-Toddler Module- Children between 12 and 30 months of age who do not consistently use phrase speech.

The algorithm produces a total score after accounting for social affect, including communication and reciprocal social interaction, and restricted and repetitive behavior. Items such as conversation, use of gestures, facial expression directed to others, quality of social response, and stereotyped/idiosyncratic use of words or phrases are scored and measured by two empirically defined cutoff points: Autistic Disorder and Broader Autism Spectrum Disorder.

As a gold standard measure, the ADOS-2 has impressive psychometric properties. Due to the extensive training required to administer the schedule, excellent inter-rater reliability exists. By requesting that those who use the ADOS-2 go through administration and scoring of the observation several times with someone previously

trained, true scoring procedures are passed along. Internal consistency measures and test-retest reliability are also promising, meaning that the observation maintains consistency in accurately measuring the construct it is supposed to measure, and that there is consistency in the test results over time if the individual is given the assessment more than once. Unfortunately, as the ADOS-2 is brand new, the majority of correlational study data is from previous versions of the schedule. For example, using the ADOS-G in studying a sample of 78 individuals with autism, 69 with PDD-NOS, and 66 with non-PDD diagnoses, the ADOS-G correctly identified 95 percent of those on the autism spectrum, and correctly identified individuals as not having autism spectrum disorder 92 percent of the time (Lord et al., 2000). This provides evidence for high PPV.

As they both meet DSM-IV-TR and ICD (World Health Organization [WHO], 1993) diagnostic criteria, have excellent validity and reliability, and are research-based, the ADI-R and ADOS-2 compliment each other well. The ADI-R asks parents to report what may not be seen in the brief time given to administer an ADOS-2. On the same note, the ADOS-2 may allow the space for a child to show behaviors that the parent has grown used to, forgotten during the interview, or not notice as a trained administrator might. Although the ADI-R and ADOS-2 have different measurement types, they work together and use a parallel research-informed structure. This ensures correlation between the two measures and links the scores of individuals to the DSM criteria. Because of this, the ADI-R and ADOS-2 have strong evidence for validity in assessing classic autism and differentiating autism from intellectual disability or language delay. A study employing the ADOS-G and ADI-R focused on differentiating children with autism from children with intellectual disability. Results of the study found that both instruments performed

well and revealed that the ADI-R had even greater specificity for identifying autism than the ADOS-G (de Bildt et al., 2004). Further, in a specialty clinic, the ADOS led to approximately 75% agreement with ADI-R and team diagnoses (Mazefsky & Oswald, 2006).

Rating Scales and Checklists

SRS-2 Social Responsiveness Scale, Second Edition

(Constantino & Gruber, 2012).

The SRS-2 is defined by its author as a quantitative approach to measuring autistic symptomology across the entire range of severity. The scale takes 15-20 minutes to complete, consists of 65 items that are rated in regard to the past 6 months.

Constantino et al., (2000) tried to reduce the number of items in the original SRS in hopes of making it less time consuming; however, this greatly affected the scale's ability to distinguish patients with ASD from other psychiatric conditions. The SRS-2 measures social behavior across five main domains: social awareness, social information processing, capacity for reciprocal social responses, social anxiety/avoidance, and characteristic autistic preoccupations/traits. The overall picture created from this score shows the child's ability to engage in emotionally appropriate reciprocal social interactions, or reciprocal social behavior. This information is gathered from parents or teachers who have seen the child interacting in naturalistic social settings. In a study by Constantino et al., (2003) the SRS did not clearly distinguish individuals diagnosed with autism from those who carried expert clinician diagnoses of PDD-NOS or Asperger's

disorder. In other words, some higher functioning autistic patients had lower scores on the SRS than did more severely impaired patients with PDD-NOS or Asperger's disorder. Coefficients for these measures ranged from .65-.77 between maternal report SRS and the ADI-R algorithm for each subdomain (social deficits, verbal communication, non-verbal communication and stereotypic behavior/restricted interest). Although the SRS was not able to distinguish between various pervasive developmental disorders in the DSM-IV-TR, with the DSM-5 this limitation would need to be evaluated differently. SRS scores were unrelated to IQ and they widely differentiate individual subjects along a continuum of severity of social impairment. In addition, long-term stability was shown with 15 participants who had previous SRS scores; after two-years, the stability of these scores was .83. The ADI-R to SRS correlations were just as strong for older participants as for younger ones, and mother, father, and teacher correlations ranged from .75 to .91. It appears that the SRS-2 is a valid instrument for the assessment of clinically significant autistic traits.

CARS-2 Childhood Autism Rating Scale, Second Edition

(Schopler, Reichler, DeVellis, & Daly, 1980)

The Children's Autism Rating Scale (CARS) (Schopler, Reichler, and Renner, 1988) is a 15 item structured observation instrument. After observation the child is rated on each item using a 7-point scale. The ratings are along the guidelines of:

Normal for child's age

Mildly abnormal

Moderately abnormal

Severely abnormal

The score range is 0 to 60, with scores above 30 meeting an autism diagnosis.

30-36: Mildly-Moderately Autistic

37-60: Severely Autistic

The CARS-2 was developed to correlate highly with the ADI-R; however, certain cautions need to be addressed. In actual correlational studies with the ADI-R, the CARS-2 over-identifies autism. The scale has also demonstrated poor discrimination between intellectual disability and autism after multiple studies (Perry, Condillac, Freeman, Dunn-Geier, & Belair, 2005; Lord, 1995; NRC, 2001). Some researchers hypothesize that the CARS-2 limitations are largely due to it being based on pre-DSM-IV-TR criteria of autism.

It is tempting to allow the less intensive and demanding nature of the CARS-2 to influence assessment-measure decision-making. Although the CARS-2 is significantly more time efficient than the ADOS-2 and requires less specific training to administer, using this measure in place of another tool with higher specificity is a disregard for evidence-based assessment and best-practice.

GARS-2 Gilliam Autism Rating Scale, Second Edition

(Gilliam, 2006)

The GARS-2 is a 42-item Likert Scale measure with three subscales: stereotyped behaviors, communication, and social interaction. The subscale standards sum to make a norm-referenced total score- the Autism Index. The GARS-2 was developed in hopes of addressing some of the limitations of the original GARS. The original GARS was

conceptually derived and loaded with internal validity issues. Mazefsky and Oswald (2006) found that the Autism Quotient (AQ) of the GARS did not correlate as highly as expected with the sub-domains of the Autism Diagnostic Observation Schedule (ADOS; Lord et al. 2002) and the Autism Diagnostic Interview- Revised (ADI-R; Rutter et al. 2003), which are reliable and valid diagnostic instruments. The authors found that the AQ underestimated the likelihood of autism. Additionally, Sikora, Hall, Hartley, Gerrand-Morris, & Cagle (2008) found that the Withdrawn and Pervasive Developmental Problems subscales of the Child Behavior Checklist Achenbach & Rescorla, 2001) were generally superior to the GARS in identifying children with autism, although specificity for both measures was consistently low. Wanting to fix these problems, Gilliam (2006) created the GARS-2 reporting that it measured the diagnostic features of autism as reflected in the DSM- IV-TR (American Psychiatric Association 2000) and the definition from the Autism Society of America (2003). Pandolfi, Magyar, & Dill (2010) hoped that a factor analysis of the GARS-2 would identify constructs assessed by the measure to inform clinical practice and research. However, exploratory and confirmatory factor analysis from matched groups of the standardization sample did not support the 3-subscale structure. Four factors were identified: stereotyped/repetitive behavior, stereotyped/idiosyncratic language, word use problems, and social impairment. Word use problems are not autism specific and identifies many other psychiatric conditions. Unfortunately, it seems that the GARS-2 remains limited in its clinical utility and the AQ should be interpreted with caution.

ABC- Autism Behavior Checklist

(Krug, Arick, & Almond, et al., 1978, 1980)

The Autism Behavior Checklist consists of 57 clinician-rated items that are sub-categorized into five groups: sensory, relating, body and object use, language, and social interaction and self-help. Ratings 67 and higher indicate a high level of autism, whereas scores under 53 support a low possibility of autism. However, many studies report that the ABC does not adequately distinguish individuals with autism spectrum disorders from other cases of developmental disorders (Rellini, Tortolani, Trillo, Carbone, & Montecchi, 2004; Nordin & Gillberg, 1996; Volkmar, Cicchetti, Dykens, Sparrow, Leckman, & Cohen, 1988; Wadden, Bryson, & Rodger, 1991). Actually, in the study by Rellini et al. (2004) the ABC performed even worse than the CARS in differentiating individuals with autistic disorders from other cases of developmental disorders. Using 53 as the cutoff, the number of false negatives was 46%. From these results, the author suggested not using this scale for diagnosis purposes (Rellini et al., 2004).

Supplemental Assessments

Cognitive Assessment

Intellectual functioning is associated with the severity of autistic symptoms and is one of the most important predictors of long-term outcome for individuals with ASD (Lovaas, 1987). The intellectual profile of individuals with ASD requires careful consideration as uneven abilities across cognitive domains is characteristic of the disorder (Klin, Saulnier, Tsatsanis, & Volkmar, 2005). To ensure that the intellectual assessment benefits the child and does not simply highlight weaknesses, the examiner should pay special attention to what the child does relatively well. With this information, one can

create a profile of the individual's strengths and weaknesses to aid in educational planning. A caution arises with intelligence testing in populations of ASD and one should remember to not use the results of cognitive assessments to confirm an autism diagnosis or differential diagnosis of ASD subtypes.

Academic Assessment

Academic measures are necessary for educational decision making and planning as this is where the child will most likely receive the most services and individual help. Once again, results from this assessment can provide a profile of academic strengths and weaknesses. In some individuals, this process will also reveal splinter skills, e.g., hyperlexia, above average memory, or calculation skills. Examples for possible academic measures to employ are: Woodcock Johnson Tests of Achievement (WJ-III; Woodcock, McGrew, & Mather, 2001), Wechsler Individual Achievement Test, Second Edition (WIAT-II; Wechsler, 2001), and Kaufman Test of Educational Achievement, Second Edition (KTEA-II; Kaufman & Kaufman, 2004).

Adaptive Behavior Assessment

Ability to display age-appropriate behaviors necessary to function safely and appropriately in daily life is an essential piece of the autism trajectory. Once again, this is a highly important predictor of long-term outcome as well as ability to live independently. Functioning in this domain is not always consistent with intellectual functioning as exhibited on a cognitive measure. As explained earlier, the diagnosis of autism comes with uneven development. Intellectual disability is only met when both cognitive and adaptive skills are compromised. It is important to note any discrepancies the individual

may demonstrate between cognitive level and adaptive scores in the areas of social skills and daily living. This information is essential for intervention and treatment recommendations as it most largely determines the possibility of the individual living alone, with family, or in assisted living. Measures of adaptive behavior that are appropriate for an autism assessment include: Vineland Adaptive Behavior Scales, Second Edition (Vineland-II; Sparrow, Cicchetti, & Balla, 2005), Developmental Profile, Third Edition (DP-3; Alpern, 2007), and the Adaptive Behavior Assessment System, Second Edition (ABAS-II; Harrison & Oakland, 2003).

Communication and Language Assessment

Evaluation of communication and language is important in a comprehensive assessment for ASD, as delays in this domain are inherent to the diagnosis. Also, similar to intelligence and adaptive behavior, the level of expressive language is an important predictor of long-term outcome for individuals with ASD. For the assessment, information should be obtained in both verbal and nonverbal domains of communication. Language areas in need of consideration are: the social communicative functions of language, nonverbal skills that regulate interaction, and communication. Varying levels of impairment can be exhibited through language, as children with high functioning autism often exhibit strengths in formal language; however, they demonstrate difficulty in communicating socially. Some measures of expressive and receptive language that may be used are: Peabody Picture Vocabulary Test, Fourth Edition (PPVT-4), Expressive One-Word Picture Vocabulary Test (EOWPVT), and Clinical Evaluation of Language Fundamentals CELF-4.

The evaluation of social communicative functions of language should be considered whether or not more basic speech and language delays are present. Some children and adolescents with less-severe ASD diagnoses do not demonstrate speech sound disorders or difficulty with language form or content; however, their use of language pragmatics may be markedly impaired (Tager-Flusberg, Paul, & Lord, 2005). Typical deficits in pragmatic language for this population include, difficulty taking turns in conversation, perseveration on specific topics, inability to shift topics when provided with conversational cues, and a lack of use of transitional cues and phrases when shifting topics (Paul & Wilson, 2009). There are many assessments that evaluate this ability: Test of Pragmatic Skills, Comprehensive Assessment of Spoken Language (CASL), Test of Pragmatic Language (TOPL-2), Test of Language Competence, Children's Communication Checklist, and Pragmatic Language Skills Inventory.

Conclusion

As the area of evidence-based practice for autism spectrum disorders continues to flourish, the same needs to happen for evidence-based assessment measures for the disorder. The contributions made by Lord et al. (1991, 2002, 2012) are promising and continue to focus on standardization, validity, reliability, and clinical implications. As the validity and diagnostic value of the ADOS-2 and the ADI-R improve, it is discouraging to see that various rating scales and checklists do not show stronger correlations with our "Gold Standard" measures. As assessment should inform practical and applicable intervention for the individual, only accurate and reliable measures should be utilized. Although the training for some measures may seem intensive and the measures

themselves lengthy, a method to ensure consistency across professionals is necessary in order to improve research and clinical practice (Klinger & Renner, 2000). With the information we have from research-informed measures, we can continue to develop psychometrically sound and practical tools to correctly identify autism at the earliest stages.

References

- Achenbach, T. M., & Rescorla, L. A. (2001). *Manual for the ASEBA school-age forms & profiles*. Burlington, VT: University of Vermont, Research Center for Children, Youth, and Families.
- Alpern, G. D. (2007). *Developmental Profile 3*. Lutz, FL: Western Psychological Services.
- American Academy of Neurology (2000). Practice parameter: Screening and diagnosis of autism: Report of the quality standards subcommittee of the American Academy of Neurology and the child neurology society. *Neurology*, *55*, 468-479.
- American Psychiatric Association (2000). *Diagnostic and Statistical Manual of Mental Disorders*, (4th ed., text revision). Washington, DC: American Psychiatric Association.
- American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders*, Fifth Edition. Arlington, VA, American Psychiatric Association, 2013.
- Baron-Cohen, S., Allen, J., & Gillberg, C. (1992). Can autism be detected at 18 months? The needle, the haystack and the CHAT. *British Journal of Psychiatry*, *138*, 839–843.
- Baron-Cohen, S., Cox, A., Baird, G., Swettenham, J., Nightingale, N., Morgan, K., Drew, A., & Charman, T. (1996). Psychological markers in the detection of autism in infancy in a large population. *British Journal of Psychiatry*, *168*, 158–163.
- Berument, S. K., Rutter, M., Lord, C., Pickles, A., & Bailey, A. (1999). Autism screening questionnaire: Diagnostic validity. *British Journal of Psychiatry*, *175*, 444–451

- California Department of Developmental Services. (2002). Autism spectrum disorders: Best practice guidelines for screening, diagnosis, and assessment. Sacramento: California DDS.
- Cohen, D., Pichard, N., Tordjman, S., Baumann, C., Burglen, L., & Excoffier, E. (2005). Specific genetic disorders and autism: Clinical contribution towards their identification. *Journal of Autism and Developmental Disorders*, *35*, 103–116.
- Constantino, J. N. (2002). *The Social Responsiveness Scale*. Los Angeles: Western Psychological Services.
- Constantino, J. N., Przybeck, T., Friesen, D., & Todd, R. D. (2000). Reciprocal social behavior in children with and without pervasive developmental disorders. *Journal of Developmental and Behavioral Pediatrics*, *21*, 2–11.
- Constantino, J. N., Davis, S. A., Todd, R. D., Schindler, M. K., Gross, M. M., Brophy, S. L., Metger, L. M., Shoushtari, C. S., Splinter, R., & Reich, W. (2003). Validation of a brief quantitative measure of autistic traits: Comparison of the social responsiveness scale with the autism diagnostic interview-revised. *Journal of Autism and Developmental Disorders*, *33*(4), 427-433
- Cox, A., Klein, K., Charman, T., Baird, G., Baron-Cohen, S., Swettenham, J., Drew, A., & Wheelwright, S. (1999). Autism spectrum disorders at 20 and 42 months of age: Stability of clinical and ADI-R diagnosis. *Journal of Child Psychology and Psychiatry*, *40*, 719-732.
- de Bildt, A., Systema, S., Ketelaars, C., Kraijer, D., Mulder, E., Volkmar, F., & Minderaa, R. (2004). Interrelationship between Autism Diagnostic Observation Schedule—Generic (ADOS-G), Autism Diagnostic Interview—Revised (ADI-R), and the

- Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR) classification in children and adolescents with mental retardation. *Journal of Autism and Developmental Disorders*, 34, 129– 137.
- Filipek, P. A., Accardo, P. J., Baranek, G. T. (1999). The screening and diagnosis of autistic spectrum disorders. *Journal of Autism and Developmental Disorders*, 29, 437– 482.
- Gilliam, J. E. (1995). Gilliam autism rating scale. Austin, TX: PRO-ED
- Gilliam, J. E. (2006). Gilliam autism rating scale (2nd ed.). Austin, TX: PRO-ED.
- Gray, K. M., Tonge, B. J., & Brereton, A. V. (2006). Screening for autism in infants, children, and adolescents. *International Review of Research in Mental Retardation*, 32, 197-227.
- Harrison, P. L., & Oakland, T. (2003). *Adaptive Behavior Assessment System- Second Edition*. San Antonio, TX: Psychological Corporation.
- Johnson, C. P. & Myers, S. M. (2007). Identification and evaluation of children with autism spectrum disorders. *Pediatrics*, 120(5), 1183–1215.
- Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman Test of Educational Achievement, Second Edition*. San Antonio, TX: Psychological Corporation.
- Klin, A., Saulnier, C., Tsatsanis, K., & Volkmar, F. R. (2005). Clinical evaluation in autism spectrum disorders: Psychological assessment within a transdisciplinary framework. In F.R. Volkmar, R. Paul, A. Klin, & D. J. Cohen (Eds.), *Handbook of autism and pervasive developmental disorders: Vol. 2. Assessment, interventions, and policy* (3rd ed., pp. 772-798). Hoboken, NJ: Wiley.

- Klinger, L. G. & Renner, P. (2000). Performance-based measures in autism: Implications for diagnosis, early detection, and identification of cognitive profiles. *The Journal of Clinical Child Psychology, 29*: 479-492.
- Krug, D. A., Arick, J. R., & Almond, P. J. (1978). Autism Screening Instrument for Educational Planning: Background and Development. In J. Gillam (Ed.), *Autism: Diagnosis, Instruction, Management and Research*. Austin (Texas): University of Texas Press.
- Krug, D. A., Arick, J. R., & Almond, P. J. (1980). Behavior checklist for identifying severely handicapped individuals with high levels of autistic behavior. *Journal of Child Psychology and Psychiatry, 21*, 221–229.
- Lovaas, I. O. (1987). Behavioral treatment and normal education and intellectual functioning in young autistic children. *Journal of Consulting and Clinical Psychology, 55*, 3-9.
- Lord, C. (1995). Follow-up of two-year-olds referred for possible autism. *Journal of Child Psychology and Psychiatry, 36*, 1365–1382. doi: 10.1111/j.1469-7610.1995.tb01669.x
- Lord, C. & Corsello, C. (2005). Diagnostic instruments in autistic spectrum disorders. In F. R. Volkmar, R. Paul, A. Klin, & D. J. Cohen (Eds.), *Handbook of Autism and Pervasive Developmental Disorders, Vol. 2: Assessment, Intervention, and Policy* (3rd ed. Pp. 730-771). Hoboken, NJ: Wiley.
- Lord, C., Pickles, A., McLennan, J., Rutter, M., Bregman, J., Folstein, S., Fombonne, E., Leboyer, M., & MInshew, N. (1997). Diagnosing autism: Analyses of data from

- the Autism Diagnostic Interview. *Journal of Autism Developmental Disorders*, 27, 501-517.
- 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. doi: 10.1023/A:1005592401947.
- Lord, C., Rutter, M., DiLavore, P. C., & Risi, S. (2002). *Autism Diagnostic Observation Schedule manual*. Los Angeles: Western Psychological Services.
- Lord, C., Rutter, M., Goode, S., Heemsbergen, J., Jordan, H., Mawhood, L., & Schopler, E. (1989). Autism diagnostic observation schedule: A standardized observation of communicative and social behavior. *Journal of Autism & Developmental Disorders*, 19(2), 185–212.
- Lord, C., Rutter, M., DiLavore, P. C., Risi, S., Gotham, K., & Bishop, S. L. (2012). *Autism Diagnostic Observation Schedule, Second Edition*. North Tonawanda, NY: Multi-Health Systems Inc.
- Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism Diagnostic Interview—Revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders*, 24, 659–685.
- 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.

- Mazefsky, C. A., & Oswald, D. P. (2006). The discriminative ability and diagnostic utility of the ADOS-G, ADI-R, and GARS for children in a clinical setting. *Autism, 10*(6), 533–549.
- McEachin, J. J., Smith, T., & Lovaas, I. O. (1993). Long-term outcome for children with autism who received early intensive behavioral treatment. *American Journal on Mental Retardation, 97*, 359-372.
- National Research Council (NRC). (2001). *Educating children with autism*. Committee on Educational Interventions for Children with Autism. Catherine Lord & James P. McGee (Eds.). Division of Behavioral and Social Sciences and Education. Washington, DC: National Academy Press.
- Nordin V, Gillberg C (1996). Autism spectrum disorders in children with physical or mental disability or both. Part I: Clinical and epidemiological aspects. *Developmental Medicine and Child Neurology 38*(4), 297-313.
- Norris, M. & Lecavalier, L. (2010). Screening accuracy of level 2 autism spectrum disorder rating scales: A review of selected instruments. *Autism, 14*, 263-284.
- Ozonoff, S. Goodlin-Jones, B., & Solomon, M. (2005). Evidence-based assessment of autism spectrum disorders in children and adolescents. *Journal of Clinical Child and Adolescent Psychology, 34*(3), 523-540.
- Pandolfi, V., Magyar, C. L., & Dill, C. A. (2010). Constructs assessed by the GARS-2: Factor analysis of data from the standardized sample. *Journal of Autism and Developmental Disorder, 40*, 1118-1130
- Paul, R. & Wilson, K. P. (2009). Assessing speech, language and communication in autism spectrum disorders. In S. Goldstein, J. A. Naglieri, & S. Ozonoff (Eds.),

- Assessment of Autism Spectrum Disorders* (pp. 335-364). New York, NY: Guilford Press.
- Perry, A., Condillac, R., Freeman, N., Dunn-Geier, J., & Belair, J. (2005). Multi-site study of the Childhood Autism Rating Scale (CARS) in five clinical groups of young children. *Journal of Autism and Developmental Disorders, 35*, 625–634. doi: 10.1007/s10803-005-0006-9.
- Pilowsky, T., Yirmiya, N., Shulman, C., & Dover, R. (1998). The autism diagnostic interview-revised and the childhood autism rating scale: Differences between diagnostic systems and comparison between genders. *Journal of Autism and Developmental Disorders, 28*(2), 143-151.
- Rellini, E., Tortolani, D., Trillo, S., Carbone, S., & Montecchi, F. (2004). Childhood autism rating scale (CARS) and autism behavior checklist (ABC) correspondence and conflicts with DSM-IV criteria in diagnosis of autism. *Journal of Autism and Developmental Disorders, 34*(6), 703-708.
- Rogers, S. J. (2001). Diagnosis of autism before the age of 3. *International review of research in mental retardation, 23*(23), 1–31.
- Rogers, S. J. & Vismara, L. A. (2008). Evidence-based comprehensive treatments for early autism. *Journal of Clinical & Adolescent Psychology, 37*(1), 8-38.
- Rutter, M., LeCouteur, A. L., & Lord, C. (2003). *Autism diagnostic interview—Revised*. Los Angeles: Western Psychological Services.
- Saemundsen, E., Magnusson, P., Smari, J., & Sigurdardottir, S. (2003). Autism diagnostic interview-revised and the childhood autism rating scale: Convergence

- and discrepancy in diagnosing autism. *Journal of Autism and Developmental Disorders*, 33(3), 319-328.
- Sikora, D. M., Hall, T. A., Hartley, S. L., Gerrard-Morris, A. E., & Cagle, S. (2008). Does parent report of behavior differ across ADOS-G classifications: Analysis of scores from the CBCL and GARS. *Journal of Autism and Developmental Disorders*, 38, 440–448.
- Schopler, E., Reichler, R., DeVellis, R., & Daly, K. (1980). Toward objective classification of childhood autism: The Childhood Autism Rating Scale (CARS). *Journal of Autism and Developmental Disorders*, 10, 91–103.
doi:10.1007/BF02408436.
- Sparrow, S. S., Cicchetti, D. V., & Balla, D. A. (2005). *Vineland Adaptive Behavior Scales, Second Edition*. Circle Pines, MN: American Guidance Service.
- Tager-Flusberg, H., Paul, R., & Lord, C. (2005). Language and communication in autism. In F. R. Volkmar, R. Paul, A. Klin, & D. J. Cohen (Eds.), *Handbook of autism and pervasive developmental disorders* (3rd ed., Vol. XX, pp. 335-364). Hoboken, NJ: Wiley.
- 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*, 18, 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 Disorders*, 21(4), 529–541.

Wilkinson, L. A. (2011, Aug. 27). *Assessment tools for ASD: Diagnostic Validity*.

Retrieved from <http://bestpracticeautism.blogspot.com/2011/08/assessment-tools-for-asd-diagnostic.html>.

Woodcock, R. W., McGrew, K. S., & Mather, N. (2001). *Woodcock-Johnson Tests of Achievement*. Rolling Meadows, IL: Riverside.

World Health Organisation. *The International Classification of Diseases and Disorders—10: Diagnostic Criteria for Research*. Geneva: World Health Organization, 1993.