



## Review

## Diagnosing autism in neurobiological research studies

Rebecca M. Jones\*, Catherine Lord

Weill-Cornell Medical College, Center for Autism and the Developing Brain, New York Presbyterian Hospital/Westchester Division, 21 Bloomingdale Road, White Plains, NY 10605 United States

## H I G H L I G H T S

- Presentation of autism symptoms varies upon developmental level, language ability and IQ.
- Methodological considerations for neurobiological autism research are addressed.
- A clinical perspective is used to outline common tools to characterize autism in a research setting.
- Advantages and disadvantages of diagnostic psychometric instruments are explored.
- Independent sources of information offer greater comprehension of the autism phenotype.

## A R T I C L E I N F O

## Article history:

Received 16 August 2012  
Received in revised form 18 October 2012  
Accepted 22 October 2012  
Available online 12 November 2012

## Keywords:

Autism  
Diagnosis  
Neurobiology  
Neuroimaging  
ADOS  
ADI-R

## A B S T R A C T

Autism Spectrum Disorder (ASD) is by definition a complex and heterogeneous disorder. Variation in factors such as developmental level, language ability and IQ further complicate the presentation of symptoms. Clinical research and basic science must continue to inform each other's questions to help address the heterogeneity inherent to the disorder. This review uses a clinical perspective to outline the common tools and best practices for diagnosing and characterizing ASD in a research setting. We discuss considerations for classifying research populations, including language ability and IQ and examine the advantages and disadvantages of different psychometric measurements. Ultimately, the contribution of multiple sources of data representing different perspectives is crucial for interpreting and understanding the ASD phenotype.

© 2012 Elsevier B.V. All rights reserved.

## Contents

1. Introduction .....	114
2. Population characterization .....	114
2.1. Autism phenotyping .....	114
2.1.1. Autism Diagnostic Observation Schedule (ADOS) .....	114
2.1.2. ADOS severity (comparison) scores .....	114
2.1.3. Diagnosis with neuroimaging or EEG? .....	115
2.1.4. Questionnaires .....	116
2.1.5. Autism Diagnostic Interview-Revised (ADI-R) .....	117
2.2. Cognitive testing – intelligence quotient .....	117
2.3. Adaptive skills .....	118
2.4. Language ability .....	118
2.5. Behavior problems .....	119
2.6. Sensory processing .....	119
2.7. Additional psychiatric disorders .....	119

\* Corresponding author. Tel.: +1 914 997 5993; fax: +1 914 997 4398.  
E-mail address: [rej2004@med.cornell.edu](mailto:rej2004@med.cornell.edu) (R.M. Jones).

3. Other research considerations .....	119
3.1. Comparison populations .....	119
3.2. Combining multiple sources of information .....	120
References .....	121

## 1. Introduction

Autism Spectrum Disorder (ASD) is a complex disorder of neurodevelopment that is currently characterized by impairments in social reciprocity, communication and unusual or restricted behaviors [1]. An ASD diagnosis is based upon behavioral history and observations. Yet, the increasing focus on the constellation of behaviors that we call ASD brings to the forefront how complex the ASD phenotype and associated behaviors are. Rarely do two children with ASD present with identical symptoms, and factors such as developmental level, language ability and IQ further complicate the presentation of symptoms.

Heterogeneity is at the forefront of any ASD research question (either as a confound or a starting point). Most research designs compare cases to controls and average data within groups. Averaging data within an ASD sample may result in bypassing crucial differences within the group. This problem is not new, and has been discussed in greater detail elsewhere [2], but to address this difficulty, researchers must rely on instruments to characterize individual components of the ASD phenotype. The measurements then become a crucial part of the study. These components range from measurements that further classify participants (language abilities, IQ) to other techniques such as neuroimaging or experimental behavior tasks (for example see [3–5]).

The goal of this review is to outline the common tools and best practices for diagnosing and characterizing ASD in a research setting from a clinical perspective. For any given research study, one must decide on which sources of diagnostic information most efficiently and reliably define the ASD phenotype and then determine which pieces of independent information (IQ, behavior tasks, neuroimaging) will address the experimental question. We begin by outlining different instruments that can be used to diagnose individuals with ASD. This is followed by a discussion of the advantages and disadvantages of other common psychometric instruments that provide more general information (such as language abilities and adaptive skills). We discuss these measurements within the context of human neurobiological research, where classifying the ASD phenotype is usually not the central goal of the experiment. Therefore, we consider advantages and disadvantages to the various instruments and highlight issues that may be particularly relevant to neuroimaging.

Overall, the issues that we address about the instruments in this review should not be significantly impacted by the proposed changes in the DSM 5. The details and rationale of the changes have been outlined elsewhere [2,6,7], but we want to stress the continuity of the instruments. In the text we will note any specific discrepancies with DSM 5.

## 2. Population characterization

As outlined in Fig. 1, there are many independent pieces of information that can be used to characterize a research sample. A standardized diagnosis forms the core of ASD clinical research. We begin by outlining different sources of information for diagnosing ASD. We then discuss other considerations including IQ testing, language ability, adaptive skills and sensory processing abnormalities that offer an additional level of perspectives on the ASD phenotype.

### 2.1. Autism phenotyping

#### 2.1.1. Autism Diagnostic Observation Schedule (ADOS)

The ADOS is a semi-structured, standardized assessment that allows examiners to observe behaviors relevant to ASD [8,9] and can be completed in approximately 30–45 min. There are explicit standards for establishment of research reliability in its administration and scoring which, when upheld, results in relatively uniform scores even across countries and translations [10]. The test consists of carefully planned social interactions and communication opportunities to elicit spontaneous behaviors within specific contexts [11]. By presenting individuals with various prompts, isolated skills can be assessed. The ADOS has strong predictive validity against best estimate diagnoses [12], and is considered by many to be the “gold standard” for classifying ASD along with the ADI-R (described below) and consensus best estimate clinical diagnosis [13].

The ADOS has five modules and the choice of which module is administered is based upon the development and language level of the participant [14]. The toddler module is intended for children up to 30–36 months of age with a nonverbal ability of 12 months [9,15]. Module 1 is intended for those who do not consistently use phrase speech, Module 2 is intended for those who use some phrase speech, but are not verbally fluent, Module 3 is intended for verbally fluent children and Module 4 is intended for verbally fluent adolescents and adults [8]. Individual behaviors are scored on a 4-point scale, with 0 representing ‘no abnormality of type specified’ and 3 representing ‘moderate to severe abnormality.’ Items are summed into two algorithms: ‘Social Affect’ (social and communication items) and ‘Restricted and Repetitive Behaviors’ [12,16]. The total algorithm scores are compared to thresholds to determine whether the child meets criteria for “Autism”, “Autism spectrum” or “nonspectrum.” A new version of the ADOS (ADOS-2) became available in 2012, providing an update to the manuals and protocols and it incorporates revised algorithms for Modules 1, 2 and 3 and includes the Toddler Module.

The ADOS is ultimately constrained by time and the specific context (e.g. an office visit with an unfamiliar adult) to complete the tasks, and therefore it is unlikely that any significant changes to the test will be made for the DSM 5. While many of the criteria changes in the DSM 5 are still well captured by the ADOS [17], several criteria are not, and never have been assessed by the ADOS, including peer interactions and relationships with others.

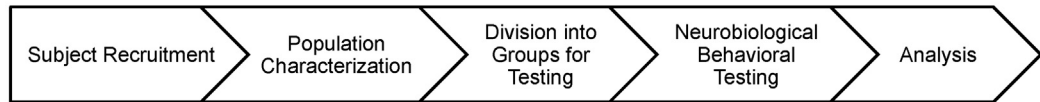
Meeting a cutoff on the ADOS does not necessarily mean that an individual should receive a diagnosis of ASD. Research has shown that diagnoses based on a combination of history/caregiver information and clinical observation are significantly more stable over time [18,19] than results from any single instrument. A more detailed discussion about considerations when choosing and combining diagnostic instruments follows in later sections.

#### 2.1.2. ADOS severity (comparison) scores

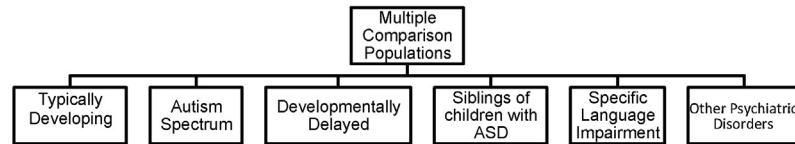
Individuals with ASD range in the severity of their symptoms. Though the authors have noted that using the ADOS raw scores, as a continuous measurement of symptom severity is not appropriate, researchers by necessity have used these scores for this purpose. One reason why these raw scores, either the algorithm totals or scores from individual behavior items, should not be used in this manner is that they are strongly influenced by age and IQ. In other words, there are different distributions for scores across items and

# Autism Researchers Roadmap

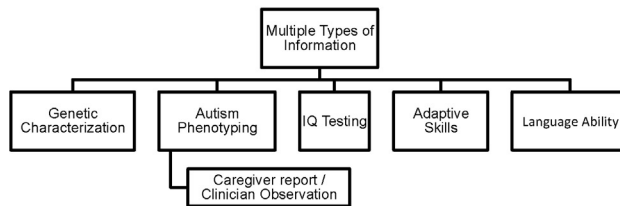
## Overview



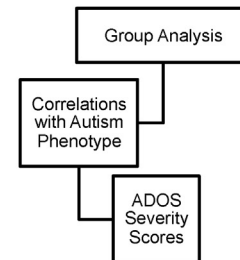
## A. Subject Recruitment



## B. Population Characterization



## C. Analysis



**Fig. 1.** Autism researchers roadmap. (A) Different comparison populations. (B) Multiple independent sources of information that be collected from the populations of study. (C) Considerations for analyzing data in terms of the ASD phenotype.

scores across developmental levels, and these differences vary in degree in terms of how much they correlate with IQ and age. However, it is possible to use the scores controlling for IQ, language level and/or age [20].

Currently, the one continuously scaled variable that can be derived from the ADOS is the severity score [21], referred to in the ADOS-2 as “comparison scores”. The comparison score is a single score derived from the two dimensions of behavior assessed by the ADOS. These scores are calibrated from raw ADOS totals using separate distributions for different age and language-level groups. They have been independently validated [22,23]. Thus, unlike raw ADOS scores, severity scores can be compared across modules and time and are relatively independent of verbal IQ [21]. Specifically, verbal IQ accounts for 43% of the variance in raw ADOS totals, but only accounts for 10% of the variance in the severity scores [21]. With the calibrated scores, the difference between a 4 and 5 is the same difference between an 8 and a 9. Since the calibrated scores have a more uniform distribution across both age and language abilities, this makes them optimal for use in neurobiological as well as other types of research. Other measurements of ASD severity, discussed in greater in the following sections, include the Autism Behavior Checklist [24], the Child Autism Rating Scale [25], the Gilliam Autism Rating Scale [26], the Autism Diagnostic Interview-Revised [27] and the Social Responsiveness Scale [28].

There are ADOS severity scores separating the two diagnostic domains (i.e. social communication and repetitive and restricted behaviors) [20]. An important issue, particularly in the field of neurobiological research, is whether varied neural differences measured by MRI or EEG in the ASD sample reflect quantifiable behavioral differences. Many functional neuroimaging studies utilize tasks that target specific areas of symptoms (i.e. social interactions), severity scores that divide social and repetitive behaviors

should be helpful for these types of studies. Analyses that take into account symptom severity can therefore go beyond differences in group means and provide more insight into how a neural signature relates to an autistic behavior.

### 2.1.3. Diagnosis with neuroimaging or EEG?

Though it is generally agreed that ASD is currently most reliably diagnosed based upon behavioral observations and descriptions, there is still a considerable amount of hope that eventually neural genetic or biological markers will provide more objective information for individuals and families [29–33]. There is growing interest in using neuroimaging methods and electroencephalography (EEG) as supplementary tools to determine biomarkers that will predict a diagnosis of ASD [29,34–41]. Magnetic Resonance Imaging (MRI) offers direct measurement of gross brain structures, Diffusion Tensor Imaging (DTI) measures the integrity of white matter tracts in the brain, and functional neuroimaging (fMRI) provides a measurement of blood flow in the brain that can be assessed during different task demands. EEG, a less costly and more portable method, than MRI, records electrical activity from the scalp. Also portable is Near Infrared Spectroscopy (NIRS), a technique that measures light absorbance in the brain to assess changes in blood flow. Less common and rarely used in young children due to the use of radioisotopes, Positron Emission Tomography (PET) measures metabolic changes in the brain. These measurements provide a means to study neural signatures of ASD [39,42,43]. However, the utility of their diagnostic potential is still being explored.

It appears that white matter tracts have different growth trajectories from 6 to 24 months of age in infants who later develop ASD compared to those children who do not [36]. Specifically, fractional anisotropy, a measurement of white fiber integrity, was higher at 6 months of age in individuals who later developed ASD.

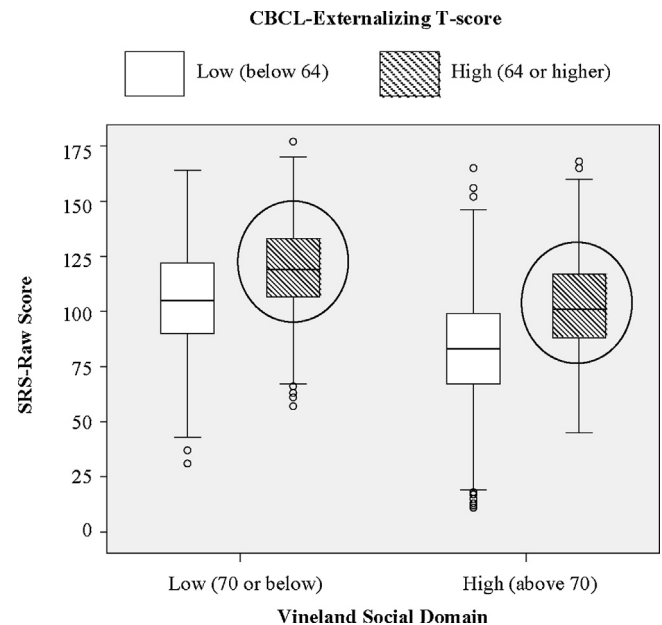
but lower at 24 months, demonstrating abnormal neural growth trajectories for those who developed ASD. These findings are exciting and suggest that neural abnormalities may precede behavioral symptoms which are not reliably characterized until approximately 24–36 months [9,44–46]. However, it is important to note that this research does not translate into a direct diagnostic tool. As the authors point out, more research is necessary to study typically developing individuals, as all of the individuals in the study were at high risk for developing ASD because they had a sibling with ASD [36]. More research is also necessary to follow the outcomes of these children whose final assessments were at 24 months. Lastly, replication is particularly important to ensure that differences are accounted for by the trajectory, rather than measurement problems at one or multiple time points. Another recent study with EEG [29], found abnormal ERP responses to dynamic eye gaze shifts at 6–10 months in children who later developed ASD. It will be important, as well, to follow these children longitudinally as well as to replicate these results in other laboratories. Together, studies examining neural biomarkers during the first year of life have the potential to provide great insight into underlying mechanisms, but their applicability as a diagnostic tool, above and beyond a behavioral measurement is not yet known.

Given the interest in using brain data as a diagnostic tool as well as to determine differences in brain structure and activity among individuals with ASD, an area of research has concentrated on using data reduction methods with EEG and neuroimaging data [47,48], specifically machine learning methods, to reliably distinguish an individual with ASD [49–52]. While some of these algorithms have proven to be highly reliable (90%) in detecting ASD among individuals, as has previously been pointed out [34,47] the use of diagnostic algorithms on neuroimaging data can only reach the same level of accuracy as the behavioral diagnostic assessment, it can never be higher. This is because the algorithm is based upon the difference between ASD and controls in the training data, which is reliant upon the behavioral measurements. Given the current understanding of ASD, it is likely that the underlying biology is heterogeneous, thus increasing reliance on behavioral observations. In addition, treatments for ASD, even medications, first consist of observations and then trying to change specific behaviors. Thus the clinical utility of a biomarker can only be as valuable as the degree to which it adds to behavioral information (which has to be obtained regardless to establish diagnosis). Furthermore, behavioral observations have shown to predict changes over time [18] and response to treatment [53], which biological markers have not done yet. If in fact biomarkers, including fMRI, become helpful in the treatment of autism then there will be a shift in perspectives about the effectiveness of these tools from research to clinical. On the other hand, even if not yet clinically useful, biomarkers may offer an intermediate step between biological and behavioral definitions.

#### 2.1.4. Questionnaires

Compared to a more complex face-to-face clinical assessment, questionnaires are more flexible and can assess additional aspects of the ASD phenotype. Clinical observation tools such as the ADOS, the Screening Tool for ASD in Toddlers and Young Children (STAT) [54], the Autism Diagnostic Interview Revised (ADI-R) and the Diagnostic Interview for Social and Communication Disorders (DISCO) [55] provide an important snapshot of clinical features. However, the benefits of using questionnaires compared to clinical diagnostic instruments is that they can be: (1) accomplished in relatively short amount of time; (2) completed by both controls and other comparison populations as well as those with ASD; (3) completed outside of a laboratory setting and (4) do not require advanced training for administration.

There are many caregiver questionnaires. Certain questionnaires are designed as a screener for ASD (i.e. Social Communication



**Fig. 2.** Influences on SRS scores. Adapted from [62]. Individuals with significant externalizing behaviors (the two circles above), have similar scores on the SRS, regardless of their social skills, measured by the Vineland Social domain. This data suggests that the SRS does not discriminate between poor social skills and significant externalizing behaviors.

Questionnaire), while others provide a more general measure (i.e. Social Responsiveness Scale) of autistic traits. Other questionnaires are used to understand both general and specific behaviors and adaptive skills (i.e. Child Behavior Checklist, Repetitive Behavior Scale-Revised or Vineland Adaptive Behavior Scales – discussed in the following section). In this section, we focus on the Social Responsiveness Scale (SRS) as it is one of the more common questionnaires used in research, we also mention other measures and relevant issues for their use in research.

The SRS [28,56] is a 65-item rating scale that has questions that pertain to autistic behaviors as well as more general behavior problems over the previous 6 months and is on a scale from 0 ('never true') to 3 ('almost always true'). Higher scores on the SRS are meant to discriminate children with or without ASD. Both sensitivity (correctly identifying individuals with ASD) and the specificity (correctly identifying individuals who do not have ASD) is somewhat mixed among studies using the SRS with values ranging from .41 to .95 [57–59]. The SRS has been shown to have a variable distribution among individuals who do not have ASD [60,61], which has been suggested to indicate that the measure captures autistic traits in the general population.

However, recent findings suggest that scores on the SRS are influenced by a broad range of factors that have to do with general levels of impairment rather than with ASD symptoms. This has important implications not only for the use of this questionnaire as an assessment of ASD but also as a continuous measurement of autistic traits and ASD severity. Specifically, Hus et al. [62] examined raw and deviation SRS scores from individuals with ASD and their non-autistic siblings and found that in individuals with ASD, higher SRS scores were associated with greater problem behaviors, higher age and more impaired language and cognitive skills. In those who did not have ASD, higher SRS scores were predicted by more behavior problems (see Fig. 2). Overall, these findings suggested that behavior problems as well as age, expressive language and cognitive level have a significant impact on SRS scores. Thus the SRS measures more general impairments than just ASD. If researchers want to use the SRS as a metric for ASD severity, this



can be done [62] but would require the use of additional covariates of behavior problems (i.e. CBCL [63] to control for their effects on the SRS).

Another screening instrument for ASD is the Social Communication Questionnaire (SCQ: [64]), which is a caregiver questionnaire based upon the ADI-R. The SCQ has good specificity, particularly when used in conjunction with the ADOS [65], for diagnosing ASD. It does not yield meaningful continuous scores but the SCQ has been used as a means to ensure the absence or presence of autistic symptoms in neurotypical participants [66]. It is not designed to measure general ASD traits in a typical population.

The ASD Quotient (AQ) [67], is a 50-item questionnaire that can be divided into five different domains and can measure autistic traits in both typically developing individuals as well as those with ASD, its psychometrics are somewhat variable [68]. Neuroimaging studies have correlated percent signal change in the brain during behavior tasks with the number of autistic traits as measured by the AQ in participants with ASD [69] as well as typically developing individuals [70–72].

The Childhood Autism Rating Scale (CARS) [25] is a 15-item behavior rating scale that is typically completed by a clinician based upon observations and/or caregiver reports. Intellectual ability and language level are included as part of the total score, making these constructs a potential confound. Rarely have studies correlated structural or functional changes in the brain with symptoms on the CARS [73], but the CARS is very commonly used in clinical settings and in studies of diagnosis and clinical change [74,75].

Lastly, a questionnaire that targets repetitive behaviors is the Repetitive Behavior Scale-Revised (RBS-R) [76]. The RBS-R has five subscales that target different types of repetitive and ritualistic behaviors and where they often occur (i.e. at home, during play etc). Recent findings suggest that the RBS-R measures the similar construct of repetitive behaviors as the ADI-R [77]. This is promising for researchers who might not have the time to complete the ADI-R. However, the RBS-R is not designed as a diagnostic tool, and therefore focuses on very specific repetitive behaviors.

There are inherent limitations to administering questionnaires over using behavioral observations. Some of these limitations are specific to ASD research and others apply to any research domain. ASD-specific questionnaires are not diagnostic instruments. This is in part because many of the questions cover broader issues than formal diagnostic instruments. Second, many of the questionnaires are limited to certain age ranges and language levels or are not validated to be independent of age or IQ. These factors are critical to consider before determining an ASD diagnosis. Lastly, parent accounts do not always correlate with behavioral assessments [78]. Many of the defining concepts of ASD are not easily communicated to a parent or other informant in a single phrase or sentence as is necessary in a questionnaire. Another possibility is developing self-report instruments (for the person with ASD reporting), but this is very complicated as part of having ASD is having difficulties with social awareness.

#### 2.1.5. Autism Diagnostic Interview-Revised (ADI-R)

The Autism Diagnostic Interview-Revised is a comprehensive interview designed to obtain historical and current information from a parent or caregiver of the patient being assessed. It is typically completed in 1.5–2.5 h. Similar to the ADOS, it is conducted by an experienced clinical interviewer who has inter-rater reliability on the measurement. The standardization of the measure relies on the interviewer, and thus the questions must be asked in a precise but clinically sensitive manner. The interview focuses on three domains of functioning (as specified in the DSM-IV); information collected from the informant is systematic and detailed [27]. The ADI-R has been demonstrated to be reliable with subjects whose mental age is above 2 years; recently algorithms were developed

for 12–47 months of age for children whose mental age equivalents went down to 10 months [79]. Behaviors are scored on a 4-point scale, with 0 representing that the behavior is not present and 3 representing severity sufficient to interfere with other areas of functioning. Also similar to the ADOS, thresholds have been developed to yield classifications of “Autism,” “Autism spectrum” or “nonspectrum” [10]. According to the proposed changes in the DSM 5, the ADI-R may need to be altered to address the need for more “current” social information particularly about older children and adults, but most items as they stand will continue to measure core ASD features. The measurement of the behavioral domains should still be valid even though cut-offs for what is ASD might shift slightly [17].

An important aspect of the ADI-R is that it focuses on behaviors that are rare in non-affected individuals, and thus it is truly designed as an ASD specific measurement. Recent research suggested that scores on the ADI-R are particularly influenced by the age, IQ and language level of the child [20,80]. Researchers may want to consider statistically controlling for these variables, before making inferences from their ADI-R data. A primary advantage of the ADI-R (also true for the ADOS) is that two separate scores are obtained, one in social communication that includes non-verbal and verbal communication and another for repetitive behaviors, can be generated [81,82]. Repetitive behaviors can be further broken down into Repetitive Motor Behaviors, Insistence on Sameness and Circumscribed Interests [83]. This could be beneficial as these scores can be analyzed separately.

How suitable an instrument is to test a particular hypothesis is a crucial question. For example, certain questionnaires are designed to have cutoffs where individuals who fall above or below a certain score do or do not have ASD (i.e. SCQ). In these cases, data from within a group will not necessarily reflect subtle differences in autistic traits as the range of values may be too limited to make meaningful interpretations. In the case of questionnaires that provide more continuous scores (e.g. T scores), it is important to consider that the data is not necessarily linear. Therefore, regardless of the measure, it is important for researchers to examine the distribution of scores across groups before drawing conclusions about general ASD symptoms from a single instrument.

## 2.2. Cognitive testing – intelligence quotient

Common practice among behavior and neurobiological studies is to test IQ in different groups and either match IQ across groups or use IQ as a covariate in analyses. Such methods allow description of ASD-specific behaviors that are theoretically independent of cognitive ability. Assessing IQ also allows researchers to understand how an individual may compare to others who are similar in developmental level (e.g., sometimes 10-year olds with ASD who score at an 8-year-old level on an IQ test, receiving IQs of approximately 80, are compared to typical 8-year olds). Yet, such matching does not take into account that brain development can be on a different trajectory than IQ. The neurodevelopment of an 8-year old may be different than that of a 10-year old, despite similarities in IQ. Furthermore, there is evidence that the relationship between IQ and cortical thickness varies with age [84].

The majority of neuroimaging research has focused on individuals who have IQs > 70 [3,85] and typically has matched groups based upon IQ and age. To be awake, lie still and complete a task in the scanner environment requires a certain level of cognitive ability, and thus functional neuroimaging tasks typically captured the most able individuals with ASD. Imaging studies during sleep [36] and EEG studies [35] are able to test less able and younger individuals, though those who have severe ASD are often not able to complete these tasks. We outline in the paragraphs below the

different options for testing intelligence in research samples and discuss some of the strengths and weakness of these assessments.

Which intelligence test to use typically depends on the developmental level of the individual as well as age. The most common tests used with children and adolescents with ASD include the Mullen Scales of Early Learning (MSEL) [86], the Differential Ability Scales (DAS-II) [87], the various age appropriate versions of the Wechsler Intelligence Scale for Children (WISC-IV) [88], the Stanford-Binet [89,90] and the Wechsler Abbreviated Scale of Intelligence (WASI-II). These tests all differ in terms of difficulty and length of testing.

The MSEL can be used to characterize young developmental ages (1–68 months) and thus does not require a child to talk or understand language. While the MSEL formally produces a single standardized general score and then separate T scores for each domain, in practice age-equivalent scores from the MSEL scales are often used to derive ratio and deviation IQ scores to provide separate estimates of verbal IQ and non-verbal IQ, and these scores have separate trajectories [18]. The MSEL has been used in individuals with ASD who have limited verbal and non-verbal skills.

For children who are reaching ceiling levels on the Mullen, the DAS is a more appropriate tool. It is standardized for ages 2.6–17.11 years. Typically 4–7 subtests are administered, although more are available. Because of its ease of use and reliability, the DAS has been the test of choice in most of the large ASD consortia research (see [17]). A more challenging cognitive test, more commonly used in typical populations, is the WISC-IV. Unlike the DAS, the instructions in the WISC-IV are more complex and the measurement of verbal comprehension more dependent on the child's expressive language. The WISC-IV is standardized for ages 6–16 years of age. For those who are over 16 years, the Wechsler Adult Intelligence Scale (WAIS) should be used. Individuals between 3 and 7 years can be tested with the Wechsler Preschool and Primary Scale of Intelligence (WPPSI-III). The decision to use the DAS or the WISC will depend on the abilities of the individual. The most recent version of the WISC-IV has fifteen subtests. Given the length of the WISC-IV, often the WASI-II is used in research settings. The WASI-II can be administered as two or four subtests, but studies have found the same individuals when tested with the WASI versus the WAIS-III [91] had higher WASI scores. The test should be used with caution. Convergent validity has been established between the DAS and the WISC-IV [92] as well as the MSEL and the DAS [93]. Lastly, more widely used in Europe is the Griffiths Mental Development Scales that has six subscales. It can be used from 0 to 8 years of age [94].

### 2.3. Adaptive skills

Adaptive skill measurements provide an important and reliable source of information about the day-to-day skills of the individual. These assessments characterize everyday behaviors and to some degree, “outcome” as defined by adaptive independence, as opposed to intelligence tests that are intended to measure learning potential [95]. Adaptive skills are important to consider in task behavior and imaging research. Determining how observed task and brain changes in ASD compared to other individuals correspond to differences in practical behaviors (for example, social functioning in the real-world) contributes important links between how ASD symptoms present themselves and behavior/brain relationships.

One of the most common measurements of adaptive skills is the Vineland Adaptive Behavior Scales [96,97]. The Vineland [97] is administered to a parent or caregiver (either in interview or survey forms). It has three domains: Communication, Daily Living Skills, and Socialization and can be used in individuals from birth to 90 years [97–99]. A fourth motor domain is only used for children up to age 7. The three domains can be further divided into subdomains that target specific areas (e.g. receptive, expressive and written language in the Communication domain or interpersonal

relationships, play and leisure time and coping skills in the Daily Living Skills domain). Maladaptive Behaviors can also be scored for individuals 3 and older.

The Vineland can be a useful tool to measure differences between ASD and other comparison groups, such as developmental delay (DD). IQ and age are critical when looking at Vineland data. As children with ASD become older, their adaptive skills are more impaired relative to age matched peers [100]. Even though raw adaptive scores increase with IQ, the relative difference between the children with ASD and others differs as a function of IQ [101]. When individuals with ASD were divided into low IQ (<70) and high IQ (>70), those in the low IQ group had Vineland scores that were above their IQ score, whereas those in the high IQ group had Vineland scores that were below their IQ score [101]. Within groups of people with ASD, studies have also shown weak correlations with the specific subdomains on the Vineland and domains on the ADOS [100] and ADOS severity scores [101]. Thus, different social measures target different aspects of behavior. This makes sense given that the Vineland is parent report about everyday contexts and the ADOS is based on behavioral observations of an experienced clinician during a single structured session. Individuals with ASD may develop compensatory abilities to help navigate in familiar environments. On the other hand, individuals may be able to function better in a one-to-one setting such as an ADOS than in more demanding, peer environments.

Neuroimaging studies have sometimes used the Vineland along with other measurements to characterize ASD symptoms [102,103]. Other measures of adaptive behaviors include the Adaptive Behavior Assessment System (ABAS-II) [104] and the Scales of Independent Behavior Revised (SIB-R) [105]. Researchers who are interested in studying a specific subset of abilities, such as social interactions in a behavioral task or in a task used during fMRI, may find assessing adaptive skills useful. For example, tasks that measure social interactions in individuals with ASD versus controls might use the Interpersonal Relationships subdomain of the Vineland to further understand the data and how differences they see may be related to daily skills.

### 2.4. Language ability

Language level of the individual influences many of the instruments described above. Even though language processing and abilities may not be central to the research question, they are crucial components for how ASD symptoms present themselves.

Some of the common measurements of language ability (although lengthy) include the Clinical Evaluation of Language Fundamentals (CELF) [106] and its preschool version [107]. Both examine expressive and receptive language abilities. Often used in younger populations (3 years of age and older) is the Comprehensive Assessment of Spoken Language (CASL) [108] that measures lexical and syntactic abilities as well as supralinguistics (complex language) and pragmatic language. The Preschool Language Scale (PLS-5) [109], which goes down to even younger ages than the CASL, measures auditory comprehension and expressive language. Receptive vocabulary can directly be tested with the Peabody Picture Vocabulary Test (PPVT) [110] and expressive vocabulary with the Expressive Vocabulary Test (EVT) [111]. Time constraints often limit the amount of measurements that can be acquired in a given study. Therefore, if researchers cannot use the more detailed measurements described above, expressive language ability can also be inferred by the following: ADOS module, the ‘overall level of language level’ item from the ADI-R, or from the Vineland expressive communication subdomain [20]. Ultimately researchers should consider measurements that separate expressive and receptive language. While correlated with each other, they have somewhat different associations with diagnosis and developmental

trajectories [112,113] and with other general social and motor abilities [114].

### 2.5. Behavior problems

Mentioned previously, collecting information on problem behaviors is important because such behaviors may interact with how ASD symptoms present themselves on other clinical diagnostic measurements. One of the most common measurements to assess problem behaviors is the Child Behavior Checklist (CBCL), a caregiver questionnaire that measures maladaptive and emotional problems [63]. There are two versions, one for children up to 5 years of age and another for ages 6–18. They both have syndrome scales that can be summed into internalizing or externalizing behavior difficulties. Children and adolescents with ASD had higher scores on attention problems, social problems and thought problems scales, but these were not independent of IQ and age [115]. The CBCL differentiated children with ASD from typical children based upon the thought problems scale [116] and social problems [117] and from children referred for non-ASD psychiatric problems on the Withdrawn, Social Problems and Thought Problems scales [118]. Another questionnaire, the Aberrant Behavior Checklist (ABC) was originally designed to assess treatment effects in people with intellectual disability, ages 6–54 [119]. It has five subscales [120]. The ABC distinguishes ASD from both typically developing individuals and those with Down Syndrome [121]. These measurements provide valuable information that may explain certain aspects of behavioral and/or neural differences in the ASD population compared to other groups that might ordinarily be overlooked.

### 2.6. Sensory processing

Assessing the sensory processing profile of an individual with ASD is particularly useful for neurobiological studies that target the general domain of repetitive and restricted behaviors. Currently, sensory interests are a new inclusion in the proposed revisions of the DSM 5 and are not part of the DSM-IV. This inclusion will most likely increase the research focus on sensory behaviors.

The ADOS, ADI-R, SCQ and SRS all address sensory symptoms, but there are specific questionnaires that isolate different aspects of sensory processing. The Short Sensory Profile [122] is a 38-item caregiver questionnaire that can be administered from birth to adulthood and uses items from the Sensory Profile [123] with the highest predictive power. It takes approximately 10 min to administer [124] and thus is well suited for research settings. There are a total of 7 sections (Tactile Sensitivity, Taste/Smell Sensitivity, Movement Sensitivity, Under-responsive/Seeks Sensation, Auditory Filtering, Low Energy/Weak and Visual/Auditory Sensitivity). The Sensory Profile has 9 sections. A total score of all the sections is the most sensitive indicator of sensory dysfunction, but these sections can be useful to separate subgroups for further research [125]. Children with ASD performed differently on 92% of the items compared to typically developing children [124] and sensory abnormalities were not correlated with IQ [126]. Studies have used these questionnaires to understand differences in the pathophysiology of the thalamus using PET, MRI [127], brainstem volume [128] and somatosensory responses with MEG [129]. However, research has shown the questionnaire did not reliably differentiate between children with ASD and ADHD [130].

Another sensory processing caregiver questionnaire is the Social Experiences Questionnaire (SEQ) [131]. It is aimed for very young children (5–72 months) and measures five sensory domains (Tactile, Auditory, Visual, vestibular-Proprioceptive and Gustatory-Olfactory). There are four subscale scores as well as totals that measure hyper- and hyporesponsiveness patterns of behavior. The

SEQ, like the SSP is brief and reliably measures sensory abnormalities in children with ASD [132].

### 2.7. Additional psychiatric disorders

Given that there is overlap with ASD and other psychiatric disorders including ADHD [133,134] and anxiety [135], it is important to consider testing for co-morbidities as they may alter the presentation of ASD symptoms [134] and even brain structures [136]. The most comprehensive research assessment is the Kiddie Schedule for Affective Disorders and Schizophrenia-Present state and Lifetime version (K-SADS-PL) [137]. It is a semi structured diagnostic interview that covers disorders for children and adolescents in the DSM-IV. Other measurements described above, particularly those that also measure problem behaviors also target symptoms that are associated with other psychiatric disorders as well, but the K-SADS-PL is the most comprehensive.

## 3. Other research considerations

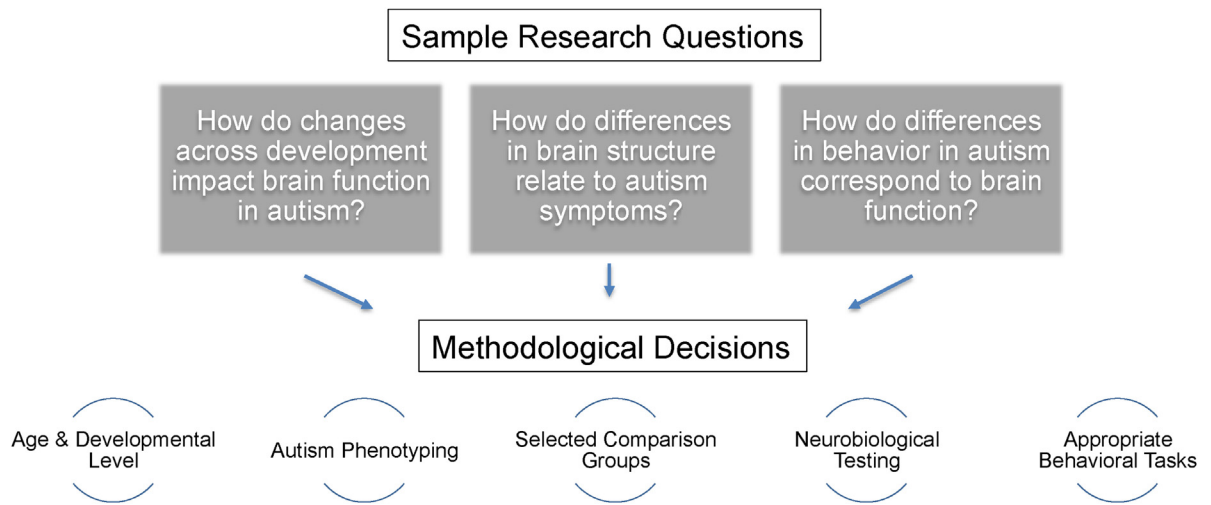
### 3.1. Comparison populations

The most common ASD studies of behavior and brain structure/function have two groups of study, individuals with ASD and those who are typically developing. However, the complex nature of how ASD symptoms present themselves on a spectrum, has forced researchers to think beyond a two-group design and to be more creative to understand what is unique to ASD.

Recently, neuroimaging and EEG studies have used an important third comparison group: siblings of children with ASD. In these studies, unaffected siblings of children with ASD [4,138–140] were compared to children with ASD and children who are typically developing. Many of these studies found a unique neural signature in the siblings of children with ASD, despite showing no differences in behavior from the case controls. These findings suggest that there may be neural compensatory mechanisms in siblings; this research furthers understanding of the neurobiology behind the spectrum of ASD symptoms.

As mentioned earlier, a second sibling design is longitudinal [29,36]. These studies specifically recruit children who are at higher risk for developing ASD because they are younger siblings of a child with ASD. This research provides the opportunity to prospectively collect data from siblings who do and do not develop ASD. Thus, this data is instrumental in charting out the neural trajectories of the autistic brain prior to the manifestation of behavioral symptoms it may also help provide potential biomarkers for the disorder beyond sibling risk.

In addition to siblings, there are other potentially important comparison groups (see Fig. 1A). Given the complex genetic findings in ASD [32], it is clear that the etiologies of ASD overlap with many other forms of psychopathology and other developmental disorders. Individuals who are developmentally delayed [141] have a specific language impairment [142] or have another disorder such as ADHD [143] may provide valuable insight into what may be unique or similar to ASD versus other developmental disorders. Our perspective comes from research in developing ASD diagnostic tools. Many of these measurements are developed and ultimately characterized by their ability to distinguish individuals with ASD from those with developmental delays and other disorders as well as neurotypical individuals. Clinically, it is rare for there to be a diagnostic question of whether a child has ASD or is typical; usually if ASD is not relevant, some other psychopathology and/or developmental delay (often both) are present. However, because neuroimaging research is expensive and time consuming a third comparison group, beyond typical development, is often



**Fig. 3.** Combining multiple sources of information. In order to address common research questions, multiple independent sources of information are necessary, impacting methodological decisions.

not considered in an initial study. Even after an initial finding in which ASD is compared to typical development, researchers become eager to build upon prior research with new research questions, rather than replicate previous findings with more appropriate additional comparison groups. Nevertheless, differentiating ASD from other disorders offers significant advances to our understanding of the ASD phenotype [144,145] and the specificity of potential biomarkers.

### 3.2. Combining multiple sources of information

As outlined throughout this paper, there are a number of methodological considerations for accumulating data on an ASD phenotype and for other behavior characteristics. In this final section, we discuss different issues when combining these multiple sources of information. First, we discuss making decisions between different diagnostic instruments and then we explore various issues when combining independent pieces of information. Fig. 3 provides sample research questions. We describe several independent methodological strategies to address these questions. The purpose of illustrating these components is to show that there are many facets to a neurobiological study of ASD and that, the more independent components that are accumulated, the more power there is to interpret the data.

Researchers must pick and choose what is most appropriate for their experimental design in terms of ASD phenotyping. As discussed in detail in Kim and Lord [19], there are certain cases that may be on the extremes (very low scores or very high scores) where it can be sufficient to administer just one instrument and rely on it for diagnosis (ADOS and ADI-R positive and negative screening estimates). Kim and Lord found that a total score on the toddler ADI-R algorithm above 18–22 or 18–25 on the ADOS in children less than 4 years of age resulted in 100% probability of an ASD classification on the other measurement. Conversely, scores of under 4–5 on the toddler ADI-R algorithm or under 8–11 on the ADOS-T resulted in less than 5% probability of receiving an ASD diagnosis on the other instrument. Using these scoring boundaries could be of great help to researchers, as Kim and Lord reported that these cut-offs applied to approximately 72% of clinic referrals in their dataset. However, one diagnostic strategy is only useful if researchers are not using both the ADOS and ADI-R for phenotype description. It may be more practical for researchers to have a standard sequence where one instrument is administered and then, only if results are uncertain, another is added. Ultimately, it will be important for

future research to look into the utility of using positive and negative screening estimates for neurobiological research studies.

Previous research has shown that correlations between parent report from the ADI-R and clinical observations from the ADOS within ASD samples are somewhat low [146–148]. Studies have consistently shown that combining information from these measurements [19,146] along with other sources of information [65] enhances diagnostic accuracy. It is important to note that the strategy described in the paragraph above, using one diagnostic test, pertains only to individuals with very high or very low scores and when researchers are satisfied with descriptive data from just one instrument. Overall, relying on both an observation and a caregiver report, along with additional specific questionnaires will result in the most accurate ASD diagnosis.

Once researchers have accumulated data on independent levels, there are many factors to explore. Below we outline various components, as listed in Fig. 3 and the ways in which they may interact with each other and how they can be harnessed to delve further into understanding the ASD phenotype and underlying neurobiology.

Age and developmental level are important considerations of ASD symptom presentation. As mentioned throughout this review, they have a significant role in how they are characterized by different instruments. Recent studies have begun to look at how typical developmental trajectories interact with ASD trajectories [36,38]. We know there are significant structural changes that occur in the brain during typical development [149,150], as well as across IQ [151]. Therefore, studies that consider age and developmental level, whether by constraining the age range [103] or measuring a broad cross section of ages [152], offer important information beyond case control comparisons.

Heterogeneity is an important consideration in research design and should not necessarily be ignored in data analysis. The differences that relate to symptom presentation within an ASD group may provide valuable data that can help to further understand disorder etiology. Therefore, researchers should consider reporting distributions of appropriate metrics, like ADOS severity scores, Vineland scores or the SRS and CBCL, to see how subtle differences in ASD symptoms or general behaviors may map onto observed task behavior or neural changes. For example, researchers could determine whether variability within the percent signal change in a certain brain region measured by fMRI in the ASD group is due to symptom severity or a tangentially related behavior to ASD such as externalizing symptoms measured on the CBCL or the level of self-help skills as measured on the Vineland. Assuming there



is sufficient power, researchers could also stratify the ASD group into smaller groups based upon one of these measurements and see if there are behavior differences measured by a psychological task and/or neural differences. All of these strategies harness the heterogeneity in ASD.

Measuring changes in task behavior (i.e. detecting facial expressions, shifting attention, predicting the receipt of rewards) are a key component of ASD research. As ASD symptoms are currently classified by standardized behavioral observations, research studies that examine psychological changes in behavior offer great insight into the disorder. By linking differences in task behaviors within an ASD sample to symptom presentation, we can further understand the nature of the impairment inherent to a symptom. Studies that use tasks to examine reaction times [153], accuracy [154], eye tracking [155] or skin conductance [156] can provide important complementary data to any ASD behavioral observation assessment. The new DSM 5 criteria classify ASD along two behavioral dimensions ([www.dsm5.org](http://www.dsm5.org)). Behavioral tasks will be a crucial component of future research to categorize individuals along continuums. These tasks will allow researchers to quantify how individuals differ along a single continuum. They may ultimately be expanded to research that assesses how different behavior continuums interact within individuals.

While continua offer one direction for designing tasks to examine differences in behavior in ASD, another popular method is to parse autism symptoms into clusters. Focusing on understanding a specific set of behaviors such as repetitive behaviors, language difficulties, social interactions or sensory problems enables researchers to isolate abnormalities. These clusters can be broken down even further into more specific impairments, for example, reward processing [157], semantic processing [158], processing of facial features [159], biological motion [4], joint attention [160], moral judgment [161], and imitation [162]. Capturing the heterogeneity of autism behaviors within a single research task is complicated, and narrowing in on a symptom enables researchers to reduce the influence of other symptom confounds. Isolating behaviors and categorizing severity within the subset of behaviors also has utility for studies of genetics [163] and mouse models of autism [164]. One of the key concerns for dividing symptoms into smaller islets is that ultimately the symptoms within one domain influence those in another. Researchers who have *a priori* hypotheses about a certain behavior within a symptom cluster should not be deterred from this approach, but be mindful of the limitations and if possible, use the basic phenotyping data as covariates to ensure observed changes are not due to other features of ASD.

Using behavioral tasks in conjunction with NIRS, EEG or neuroimaging offers an additional level of complexity. Some studies in ASD have shown no behavioral differences (such as no group differences in reaction times or accuracy) in the presence of brain changes (for example see Dichter et al. [165]), while other studies show differences in behavior that correspond to brain changes [166]. The behavior and brain relationship is a pivotal focus for human ASD research with an increasing number of studies considering both sources of data [167]. Understanding the interplay between the two will not only further our understanding of ASD and offer clues into biomarkers, but also ultimately provide insight into typical neurodevelopment.

## References

- [1] Association AP. Task Force on DSM-IV. Diagnostic and statistical manual of mental disorders: DSM-IV-TR. 4th ed. Washington DC: American Psychiatric Association; 2000.
- [2] Lord C, Jones RM. Annual research review: re-thinking the classification of autism spectrum disorders. *Journal of Child Psychology and Psychiatry, and Allied Disciplines* 2012;53:490–509.

- [3] Di Martino A, Kelly C, Grzadzinski R, Zuo XN, Mennes M, Mairena MA, et al. Aberrant striatal functional connectivity in children with autism. *Biological Psychiatry* 2011;69:847–56.
- [4] Kaiser MD, Hudac CM, Shultz S, Lee SM, Cheung C, Berken AM, et al. Neural signatures of autism. *Proceedings of the National Academy of Sciences of the United States of America* 2010;107:21223–8.
- [5] Jones W, Carr K, Klin A. Absence of preferential looking to the eyes of approaching adults predicts level of social disability in 2-year-old toddlers with autism spectrum disorder. *Archives of General Psychiatry* 2008;65:946–54.
- [6] Swedo SE, Baird G, Cook Jr EH, Happe FG, Harris JC, Kaufmann WE, et al. Commentary from the DSM-5 workgroup on neurodevelopmental disorders. *Journal of the American Academy of Child and Adolescent Psychiatry* 2012;51:347–9.
- [7] Association AP. DSM-5; 2012.
- [8] Lord C, Rutter M, DiLavore PC, Risi S, Gotham K, Bishop SL. Autism diagnostic observation schedule, second edition (ADOS-2) manual (part 1): modules 1–4. Torrance, CA: Western Psychological Services; 2012.
- [9] Lord C, Luyster RJ, Gotham K, Guthrie W. Autism diagnostic observation schedule, second edition (ADOS-2) manual (part ii): toddler module. Torrance, CA: Western Psychological Services; 2012.
- [10] Lord C, Petkova E, Hus V, Gan W, Lu F, Martin DM, et al. A multisite study of the clinical diagnosis of different autism spectrum disorders. *Archives of General Psychiatry* 2012;69:306–13.
- [11] Lord C, Risi S, Lambrecht L, Cook Jr EH, Leventhal BL, DiLavore PC, et al. 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* 2000;30:205–23.
- [12] Gotham K, Risi S, Pickles A, Lord C. The Autism Diagnostic Observation Schedule: revised algorithms for improved diagnostic validity. *Journal of Autism and Developmental Disorders* 2007;37:613–27.
- [13] de Bildt A, Sytema S, Ketelaars C, Kraijer D, Mulder E, Volkmar F, et al. 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* 2004;34:129–37.
- [14] Hus V, Lord C. Adapted ADOS, modules 1 and 2. Unpublished manuscript; 2012.
- [15] Luyster R, Gotham K, Guthrie W, Coffing M, Petrak R, Pierce K, et al. The Autism Diagnostic Observation Schedule-toddler module: a new module of a standardized diagnostic measure for autism spectrum disorders. *Journal of Autism and Developmental Disorders* 2009;39:1305–20.
- [16] Gotham K, Risi S, Dawson G, Tager-Flusberg H, Joseph R, Carter A, et al. A replication of the Autism Diagnostic Observation Schedule (ADOS) revised algorithms. *Journal of the American Academy of Child and Adolescent Psychiatry* 2008;47:642–51.
- [17] Huerta M, Bishop SL, Duncan A, Hus V, Lord C. Application of DSM-5 criteria to 3 samples of children with DSM-IV diagnoses of PDD. *American Journal of Psychiatry* 2012;169:1056–64.
- [18] Lord C, Risi S, DiLavore PS, Shulman C, Thurm A, Pickles A. Autism from 2 to 9 years of age. *Archives of General Psychiatry* 2006;63:694–701.
- [19] Kim SH, Lord C. Combining information from multiple sources for the diagnosis of autism spectrum disorders for toddlers and young preschoolers from 12 to 47 months of age. *Journal of Child Psychology and Psychiatry, and Allied Disciplines* 2012;53:143–51.
- [20] Hus V, Lord C. Effects of child characteristics on the autism diagnostic interview-revised: implications for use of scores as a measure of ASD severity. *Journal of Autism and Developmental Disorders* 2012.
- [21] Gotham K, Pickles A, Lord C. Standardizing ADOS scores for a measure of severity in autism spectrum disorders. *Journal of Autism and Developmental Disorders* 2009;39:693–705.
- [22] de Bildt A, Oosterling IJ, van Lang ND, Sytema S, Minderaa RB, van Engeland H, et al. Standardized ADOS scores: measuring severity of autism spectrum disorders in a Dutch sample. *Journal of Autism and Developmental Disorders* 2011;41:311–9.
- [23] Shumway S, Farmer C, Thurm A, Joseph L, Black D, Golden C. The ADOS calibrated severity score: relationship to phenotypic variables and stability over time. *Autism Research: Official Journal of the International Society for Autism Research* 2012;5:267–76.
- [24] Krug DA, Arick J, Almond P. Behavior checklist for identifying severely handicapped individuals with high levels of autistic behavior. *Journal of Child Psychology and Psychiatry, and Allied Disciplines* 1980;21:221–9.
- [25] Schopler E, Reichler RJ, DeVellis RF, Daly K. Toward objective classification of childhood autism: Childhood Autism Rating Scale (CARS). *Journal of Autism and Developmental Disorders* 1980;10:91–103.
- [26] Gilliam JE. Gilliam autism rating scale: examiner's manual. Austin, TX: Pro-Ed.; 1995.
- [27] Rutter M, LeCouteur A, Lord C. Autism diagnostic interview-revised (ADI-R) manual. Los Angeles: Western Psychological Services; 2003.
- [28] Constantino JN. Social responsiveness scale second edition (SRS-2). Los Angeles: Western Psychological Services; 2012.
- [29] Elsabbagh M, Mercure E, Hudry K, Chandler S, Pasco G, Charman T, et al. Infant neural sensitivity to dynamic eye gaze is associated with later emerging autism. *Current Biology* 2012;22:1–5.



- [82] Frazier TW, Youngstrom EA, Kubu CS, Sinclair L, Rezaei A. Exploratory and confirmatory factor analysis of the autism diagnostic interview-revised. *Journal of Autism and Developmental Disorders* 2008;38:474–80.
- [83] Lam KS, Bodfish JW, Piven J. Evidence for three subtypes of repetitive behavior in autism that differ in familiarity and association with other symptoms. *Journal of Child Psychology and Psychiatry, and Allied Disciplines* 2008;49:1193–200.
- [84] Shaw P, Greenstein D, Lerch J, Clasen L, Lenroot R, Gogtay N, et al. Intellectual ability and cortical development in children and adolescents. *Nature* 2006;440:676–9.
- [85] Ecker C, Suckling J, Deoni SC, Lombardo MV, Bullmore ET, Baron-Cohen S, et al. Brain anatomy and its relationship to behavior in adults with autism spectrum disorder: a multicenter magnetic resonance imaging study. *Archives of General Psychiatry* 2012;69:195–209.
- [86] Mullen EM. Mullen scales of early learning. Circle Pines, MN: American Guidance Services, Inc.; 1995.
- [87] Elliott CD. Differential ability scales. 2nd ed. Antonio, TX: Psychological Corporation; 1990.
- [88] Wechsler D. Wechsler intelligence scale for children. 4th ed. San Antonio, TX: Psychological Corporation; 2003.
- [89] Thorndike RL, Hagen EP, Sattler JM. The Stanford–Binet intelligence scale. 4th ed. Chicago, IL: The Riverside Publishing Co.; 1986.
- [90] Roid GH. Stanford–Binet intelligence scales, 5th ed. Rolling Meadows, IL: Riverside Publishing 2003.
- [91] Axelrod BN. Validity of the Wechsler abbreviated scale of intelligence and other very short forms of estimating intellectual functioning. *Assessment* 2002;9:17–23.
- [92] Elliott CD. Differential ability scales. 2nd ed. San Antonio, TX: Harcourt Assessment, Inc.; 2007.
- [93] Bishop SL, Guthrie W, Coffing M, Lord C. Convergent validity of the Mullen Scales of Early Learning and the differential ability scales in children with autism spectrum disorders. *American Journal on Intellectual and Developmental Disabilities* 2011;116:331–43.
- [94] Griffiths R. The Griffiths mental development scales. 1996 revision Association for Research in Infant and Child Development, Test Agency; 1996.
- [95] Kraijer D. Review of adaptive behavior studies in mentally retarded persons with autism/pervasive developmental disorder. *Journal of Autism and Developmental Disorders* 2000;30:39–47.
- [96] Sparrow SS, Balla DA, Cicchetti DV. Vineland adaptive behavior scales interview edition expanded form manual. Circle Pines, MN: American Guidance Services, Inc.; 1984.
- [97] Sparrow SS, Cicchetti DV, Balla DA. Vineland adaptive behavior scales. 2nd ed. Bloomington, MN: NCS Pearson, Inc; 2005.
- [98] Ray-Subramanian CE, Huai N, Ellis Weismer S. Brief report: adaptive behavior and cognitive skills for toddlers on the autism spectrum. *Journal of Autism and Developmental Disorders* 2011;41:679–84.
- [99] Volkmar FR, Sparrow SS, Goudreau D, Cicchetti DV, Paul R, Cohen DJ. Social deficits in autism: an operational approach using the Vineland Adaptive Behavior Scales. *Journal of the American Academy of Child and Adolescent Psychiatry* 1987;26:156–61.
- [100] Klin A, Saulnier CA, Sparrow SS, Cicchetti DV, Volkmar FR, Social Lord C. communication abilities and disabilities in higher functioning individuals with autism spectrum disorders: the Vineland and the ADOS. *Journal of Autism and Developmental Disorders* 2007;37:748–59.
- [101] Kanne SM, Gerber AJ, Quirimbach LM, Sparrow SS, Cicchetti DV, Saulnier CA. The role of adaptive behavior in autism spectrum disorders: implications for functional outcome. *Journal of Autism and Developmental Disorders* 2011;41:1007–18.
- [102] Schumann CM, Barnes CC, Lord C, Courchesne E. Amygdala enlargement in toddlers with autism related to severity of social and communication impairments. *Biological Psychiatry* 2009;66:942–9.
- [103] Munson J, Dawson G, Abbott R, Faja S, Webb SJ, Friedman SD, et al. Amygdalar volume and behavioral development in autism. *Archives of General Psychiatry* 2006;63:686–93.
- [104] Harrison PL, Oakland T. Adaptive behavior assessment system. 2nd ed. Minneapolis, MN: Pearson Assessment; 2003.
- [105] Bruininks R, Woodcock R, Weatherman R, Hill B. Scales of independent behavior-revised. Rolling Meadows, IL: Riverside Publishing; 1997.
- [106] Semel E, Wiig EH, Secord WA. Clinical evaluation of language fundamentals – fourth edition (CELF-4). Toronto, Canada: The Psychological Corporation/A Harcourt Assessment Company; 2003.
- [107] Wiig EH, Secord WA, Semel E. Clinical evaluation of language fundamentals-preschool, second edition (CELF Preschool-2). Toronto, Canada: The Psychological Corporation/A Harcourt Assessment Company; 2004.
- [108] Carrow-Woolfolk E. Comprehensive assessment of spoken language. Circle Pines, MN: American Guidance Service, Inc; 1999.
- [109] Zimmerman IL, Steiner VG, Pond RE, Karas GB, Marshall J, Boland J, et al. PLS-5: preschool language scales. Bloomington, MN: Pearson/PsychCorp; 2011.
- [110] Dunn LM, Dunn DM. Peabody picture vocabulary test. 4th ed. Minneapolis, MN: Pearson Education; 2007.
- [111] Williams KT. Expressive vocabulary test. 2nd ed. Toronto, Canada: Pearson Assessment Inc.; 2007.
- [112] Hudry K, Leadbitter K, Temple K, Slonims V, McConachie H, Aldred C, et al. Preschoolers with autism show greater impairment in receptive compared with expressive language abilities. *International Journal of Language & Communication Disorders/Royal College of Speech & Language Therapists* 2010;45:681–90.
- [113] Luyster R, Qiu S, Lopez K, Lord C. Predicting outcomes of children referred for autism using the MacArthur–Bates Communicative Development Inventory. *Journal of Speech, Language, and Hearing Research: JSLHR* 2007;50:667–81.
- [114] Luyster RJ, Kadlec MB, Carter A, Tager-Flusberg H. Language assessment and development in toddlers with autism spectrum disorders. *Journal of Autism and Developmental Disorders* 2008;38:1426–38.
- [115] Bolte S, Dickhut H, Poustka F. Patterns of parent-reported problems indicative in autism. *Psychopathology* 1999;32:93–7.
- [116] Duarte CS, Bordin IA, de Oliveira A, Bird H. The CBCL and the identification of children with autism and related conditions in Brazil: pilot findings. *Journal of Autism and Developmental Disorders* 2003;33:703–7.
- [117] Mazefsky CA, Anderson R, Conner CM, Minshew N. Child behavior checklist scores for school-aged children with autism: preliminary evidence of patterns suggesting the need for referral. *Journal of Psychopathology and Behavioral Assessment* 2011;33:31–7.
- [118] Biederman J, Petty CR, Fried R, Wozniak J, Micco JA, Henin A, et al. Child behavior checklist clinical scales discriminate referred youth with autism spectrum disorder: a preliminary study. *Journal of Developmental and Behavioral Pediatrics: JDBP* 2010;31:485–90.
- [119] Aman MG, Singh NN, Stewart AW, Field CJ. The aberrant behavior checklist: a behavior rating scale for the assessment of treatment effects. *American Journal of Mental Deficiency* 1985;89:485–91.
- [120] Aman MG, Singh NN. Aberrant behavior checklist (ABC). East Aurora, NY: Slosson Educational Publications; 1986.
- [121] Capone GT, Grados MA, Kaufmann WE, Bernard-Ripoll S, Jewell A. Down syndrome and comorbid autism-spectrum disorder: characterization using the aberrant behavior checklist. *American Journal of Medical Genetics Part A* 2005;134:373–80.
- [122] McIntosh DN, Miller LJ, Shyu V. Development and validation of the Short Sensory Profile. In: Dunn W, editor. Sensory profile manual. San Antonio, TX: Psychological Corporation; 1999.
- [123] Dunn W. The sensory profile: user's manual. San Antonio, TX: Psychological Corporation; 1999.
- [124] Tomchek SD, Dunn W. Sensory processing in children with and without autism: a comparative study using the short sensory profile. *The American Journal of Occupational Therapy: Official Publication of the American Occupational Therapy Association* 2007;61:190–200.
- [125] Lane AE, Dennis SJ, Geraghty ME. Brief report: further evidence of sensory subtypes in autism. *Journal of Autism and Developmental Disorders* 2011;41:826–31.
- [126] Rogers SJ, Hepburn S, Wehner E. Parent reports of sensory symptoms in toddlers with autism and those with other developmental disorders. *Journal of Autism and Developmental Disorders* 2003;33:631–42.
- [127] Hardan AY, Minshew NJ, Melhem NM, Srihari S, Jo B, Bansal R, et al. An MRI and proton spectroscopy study of the thalamus in children with autism. *Psychiatry Research* 2008;163:97–105.
- [128] Jou RJ, Minshew NJ, Melhem NM, Keshavan MS, Hardan AY. Brainstem volumetric alterations in children with autism. *Psychological Medicine* 2009;39:1347–54.
- [129] Marco EJ, Khatibi K, Hill SS, Siegel B, Arroyo MS, Dowling AF, et al. Children with autism show reduced somatosensory response: an MEG study. *Autism Research: Official Journal of the International Society for Autism Research* 2012;5:340–51.
- [130] Cheung PP, Siu AM. A comparison of patterns of sensory processing in children with and without developmental disabilities. *Research in Developmental Disabilities* 2009;30:1468–80.
- [131] Baranek GT, David FJ, Poe MD, Stone WL, Watson LR. Sensory Experiences Questionnaire: discriminating sensory features in young children with autism, developmental delays, and typical development. *Journal of Child Psychology and Psychiatry, and Allied Disciplines* 2006;47:591–601.
- [132] Little LM, Freuler AC, Houser MB, Guckian L, Carbine K, David FJ, et al. Psychometric validation of the Sensory Experiences Questionnaire. *The American Journal of Occupational Therapy: Official Publication of the American Occupational Therapy Association* 2011;65:207–10.
- [133] Grzadzinski R, Di Martino A, Brady E, Mairena MA, O'Neale M, Petkova E, et al. Examining autistic traits in children with ADHD: does the autism spectrum extend to ADHD. *Journal of Autism and Developmental Disorders* 2011;41:1178–91.
- [134] Yerys BE, Wallace GL, Sokoloff JL, Shook DA, James JD, Kenworthy L. Attention deficit/hyperactivity disorder symptoms moderate cognition and behavior in children with autism spectrum disorders. *Autism Research: Official Journal of the International Society for Autism Research* 2009;2:322–33.
- [135] Guttman-Steinmetz S, Gadow KD, DeVincenzi CJ, Crowell J. Anxiety symptoms in boys with autism spectrum disorder, attention-deficit hyperactivity disorder, or chronic multiple tic disorder and community controls. *Journal of Autism and Developmental Disorders* 2010;40:1006–16.
- [136] Toal F, Bloemen OJ, Deeley Q, Tunstall N, Daly EM, Page L, et al. Psychosis and autism: magnetic resonance imaging study of brain anatomy. *The British Journal of Psychiatry: The Journal of Mental Science* 2009;194:418–25.
- [137] Kaufman J, Birmaher B, Brent D, Rao U, Flynn C, Moreci P, et al. Schedule for Affective Disorders and Schizophrenia for School-Age Children–Present and Lifetime Version (K-SADS-PL): initial reliability and validity data. *Journal of the American Academy of Child and Adolescent Psychiatry* 1997;36:980–8.

- [138] Barnea-Goraly N, Lotspeich LJ, Reiss AL. Similar white matter aberrations in children with autism and their unaffected siblings: a diffusion tensor imaging study using tract-based spatial statistics. *Archives of General Psychiatry* 2010;67:1052–60.
- [139] Dalton KM, Nacewicz BM, Alexander AL, Davidson RJ. Gaze-fixation, brain activation, and amygdala volume in unaffected siblings of individuals with autism. *Biological Psychiatry* 2007;61:512–20.
- [140] Belmonte MK, Gomot M, Baron-Cohen S. Visual attention in autism families: 'unaffected' sibs share atypical frontal activation. *Journal of Child Psychology and Psychiatry, and Allied Disciplines* 2010;51:259–76.
- [141] Dawson G, Carver L, Meltzoff AN, Panagiotides H, McPartland J, Webb SJ. Neural correlates of face and object recognition in young children with autism spectrum disorder, developmental delay, and typical development. *Child Development* 2002;73:700–17.
- [142] Verhoeven JS, Rommel N, Prodi E, Leemans A, Zink I, Vandewalle E, et al. Is there a common neuroanatomical substrate of language deficit between autism spectrum disorder and specific language impairment? *Cereb Cortex* 2012;22:2263–71.
- [143] Brandimonte MA, Filippello P, Coluccia E, Altgassen M, Kliegel M. To do or not to do? Prospective memory versus response inhibition in autism spectrum disorder and attention-deficit/hyperactivity disorder. *Memory* 2011;19:56–66.
- [144] Wilson R, Pascalis O, Blades M. Familiar face recognition in children with autism; the differential use of inner and outer face parts. *Journal of Autism and Developmental Disorders* 2007;37:314–20.
- [145] Chawarska K, Macari S, Shic F. Context modulates attention to social scenes in toddlers with autism. *Journal of Child Psychology and Psychiatry, and Allied Disciplines* 2012;53:903–13.
- [146] Risi S, Lord C, Gotham K, Corsello C, Chrysler C, Szatmari P, et al. Combining information from multiple sources in the diagnosis of autism spectrum disorders. *Journal of the American Academy of Child and Adolescent Psychiatry* 2006;45:1094–103.
- [147] Le Couteur A, Haden G, Hammal D, McConachie H. Diagnosing autism spectrum disorders in pre-school children using two standardised assessment instruments: the ADI-R and the ADOS. *Journal of Autism and Developmental Disorders* 2008;38:362–72.
- [148] Ventola PE, Kleinman J, Pandey J, Barton M, Allen S, Green J, et al. Agreement among four diagnostic instruments for autism spectrum disorders in toddlers. *Journal of Autism and Developmental Disorders* 2006;36:839–47.
- [149] Raznahan A, Shaw P, Lalonde F, Stockman M, Wallace GL, Greenstein D, et al. How does your cortex grow? *The Journal of Neuroscience: the Official Journal of the Society for Neuroscience* 2011;31:7174–7.
- [150] Gogtay N, Giedd JN, Lusk L, Hayashi KM, Greenstein D, Vaituzis AC, et al. Dynamic mapping of human cortical development during childhood through early adulthood. *Proceedings of the National Academy of Sciences of the United States of America* 2004;101:8174–9.
- [151] Ramsden S, Richardson FM, Josse G, Thomas MS, Ellis C, Shakeshaft C, et al. Verbal and non-verbal intelligence changes in the teenage brain. *Nature* 2011;479:113–6.
- [152] Duerden EG, Mak-Fan KM, Taylor MJ, Roberts SW. Regional differences in grey and white matter in children and adults with autism spectrum disorders: an activation likelihood estimate (ALE) meta-analysis. *Autism Research: Official Journal of the International Society for Autism Research* 2012;5:49–66.
- [153] Larson MJ, South M, Clayson PE, Clawson A. Cognitive control and conflict adaptation in youth with high-functioning autism. *Journal of Child Psychology and Psychiatry, and Allied Disciplines* 2012;53:440–8.
- [154] Riby DM, Doherty-Sneddon G, Whittle L. Face-to-face interference in typical and atypical development. *Developmental Science* 2012;15:281–91.
- [155] Guiraud JA, Tomalski P, Kushnerenko E, Ribeiro H, Davies K, Charman T, et al. Atypical audiovisual speech integration in infants at risk for autism. *PLoS ONE* 2012;7:e36428.
- [156] Kylliaainen A, Wallace S, Coutanche MN, Leppanen JM, Cusack J, Bailey AJ, et al. Affective-motivational brain responses to direct gaze in children with autism spectrum disorder. *Journal of Child Psychology and Psychiatry, and Allied Disciplines* 2012;53:790–7.
- [157] Dichter GS, Felder JN, Green SR, Rittenberg AM, Sasson NJ, Bodfish JW. Reward circuitry function in autism spectrum disorders. *Social Cognitive and Affective Neuroscience* 2012;7:160–72.
- [158] Knaus TA, Silver AM, Lindgren KA, Hadjikhani N, Tager-Flusberg H. fMRI activation during a language task in adolescents with ASD. *Journal of the International Neuropsychological Society: JINS* 2008;14:967–79.
- [159] Kliemann D, Dziobek I, Hatri A, Baudewig J, Heekeren HR. The role of the amygdala in atypical gaze on emotional faces in autism spectrum disorders. *The Journal of Neuroscience: the Official Journal of the Society for Neuroscience* 2012;32:9469–76.
- [160] Redcay E, Dodel-Feder D, Mavros PL, Kleiner M, Pearrow MJ, Triantafyllou C, et al. Atypical brain activation patterns during a face-to-face joint attention game in adults with autism spectrum disorder. *Human Brain Mapping* 2012.
- [161] Moran JM, Young LL, Saxe R, Lee SM, O'Young D, Mavros PL, et al. Impaired theory of mind for moral judgment in high-functioning autism. *Proceedings of the National Academy of Sciences of the United States of America* 2011;108:2688–92.
- [162] Spengler S, Bird G, Brass M. Hyperimitation of actions is related to reduced understanding of others' minds in autism spectrum conditions. *Biological Psychiatry* 2010;68:1148–55.
- [163] Rudie JD, Hernandez LM, Brown JA, Beck-Pancer D, Colich NL, Gorrindo P, et al. Autism-associated promoter variant in MET impacts functional and structural brain networks. *Neuron* 2012;75:904–15.
- [164] Peca J, Feliciano C, Ting JT, Wang W, Wells MF, Venkatraman TN, et al. Shank3 mutant mice display autistic-like behaviours and striatal dysfunction. *Nature* 2011;472:437–42.
- [165] Dichter GS, Richey JA, Rittenberg AM, Sabatino A, Bodfish JW. Reward circuitry function in autism during face anticipation and outcomes. *Journal of Autism and Developmental Disorders* 2012;42:147–60.
- [166] Shafritz KM, Dichter GS, Baranek GT, Belger A. The neural circuitry mediating shifts in behavioral response and cognitive set in autism. *Biological Psychiatry* 2008;63:974–80.
- [167] Kaiser MD, Pelphrey KA. Disrupted action perception in autism: behavioral evidence, neuroendophenotypes, and diagnostic utility. *Developmental Cognitive Neuroscience* 2012;2:25–35.