Measurement of Autism: A Review of Four Screening Measures

Nishalie Fernandopulle

ABSTRACT

Using a key word search on electronic databases, two relevant journals and relevant review references, four measures of autistic symptoms (Childhood Asperger Syndrome Test, Autism Behavior Checklist, Social Communication Questionnaire, Social Responsiveness Scale) were reviewed with reference to their ability to discriminate the three major components of autism and measure across the whole spectrum of autism. None of the reviewed measures were able to effectively tap into and differentiate between all points on the spectrum. Further work is required to assess the sensitivity of the measures to independence between domains. The development of a measure sensitive to symptom change/continuity across development may be useful.

Key words: Autism, autism behavior checklist, childhood asperger syndrome test, measurement, social communication questionnaire, social responsiveness scale

INTRODUCTION

The concept of autism has undergone transition since Kanner first delineated it in 1943. The boundaries of the disorder have widened with time, and it is increasingly conceptualized as a spectrum of related diagnostic categories as opposed to an all-or-none disorder. Consequently, it is now believed that what is transmitted genetically is not classic autism but a more varied phenotype of social, communication and/or behavioral difficulties. To screen for these difficulties there is a need for a valid and reliable measure that has a broader conceptualization of the spectrum and its domains. Prior to the development of such a measure, however, it is important that existing screening measures, which have played facilitatory roles in this transition of autism are reviewed; and the requirement, or alternatively the non-requirement, for developing further measures that tap into and discriminate between the entire range of the spectrum and its domains be identified. Hence this review aims to provide a synthesis of published research on four measures of autistic symptoms, with emphasis on:

• Ability to discriminate the three major components of autism
• Ability to measure across the whole spectrum of autism

MATERIALS AND METHODS

Although the lines delineating existing measures are vague, the following criteria were adhered to in selecting those for review:

Inclusion criteria

The review was restricted to journal-published validation studies for rating-scales/checklists/questionnaires designed to measure autistic symptoms in clinical and
educational settings. Only measurements completed by parents/teachers were selected.

Exclusion criteria
Since the focus was on measures targeting autism spectrum disorders (ASD), those not specifically designed for the spectrum and those specifically predesigned for a specific point (e.g. only asperger syndrome) on the spectrum were excluded as the objective was to identify tools with the ability to discriminate across the spectrum. Measures focusing purely on adults/adolescents/infants were also excluded as the measures should have the potential to measure continuity/change of symptoms over time. Interview and observation-based measures and measures targeting response to treatment were excluded on the basis that they are not efficient or suitable for screening purposes. A list of non-selected measures is presented in Table 1.

Search method
A literature search was conducted incorporating electronic databases, a hand search, and reference lists.

Electronic databases
Initially, a search was conducted in the following databases:
• PsycINFO
• Ovid MEDLINE
• Web of Knowledge
• Cochrane Library

The key words used were:
• Autism; Autistic disorders; Asperger Syndrome (AS)
• Measurement; Questionnaire; Scale; Test; Assessment; Screening

Table 1: Measures that did not meet the inclusion criteria of this review

| Measure                                                      | Reason for exclusion                                                                 |
|--------------------------------------------------------------|-------------------------------------------------------------------------------------|
| Asperger syndrome (and high-functioning autism) diagnostic interview | Interview-based. Targets Asperger Syndrome and high-functioning autism only         |
| Autism diagnostic interview – Revised                        | Interview-based                                                                     |
| Autism diagnostic observation schedule – Generic             | Observatory measure                                                                |
| Autism spectrum disorder in adults screening questionnaire    | Targets only adults                                                                 |
| Autism spectrum screening questionnaire                       | Targets Asperger Syndrome and high-functioning autism only                           |
| Behavior function inventory                                  | Ratings by clinicians/professionals                                                 |
| Checklist for autism in toddlers                             | Targets only infants                                                                |
| Childhood autism rating scale                                | Ratings by clinicians/professionals                                                 |
| Children’s social behavior questionnaire                     | Targets milder variants of the spectrum                                             |
| Developmental, dimensional and diagnostic interview – 3di    | Interview                                                                          |
| Gilliam autism rating scale – Second edition                  | No published reports retrieved to document the psychometric properties of the 2nd Edition |
| Modified checklist for autism in toddlers                     | Targets only infants                                                                |
| Parent interview autism                                       | Interview                                                                          |
| Pervasive developmental disorder behavior inventory          | Assesses responsiveness to intervention                                              |
| Pervasive behavior developmental disorder screening test      | No published reports retrieved to document the psychometric properties retrieved   |
| Screening tool for autism in two-year-olds                    | Targets only infants                                                                |
| Rimland’s diagnostic checklist for behavior-disturbed children | Developed pre-1980                                                                  |
Childhood asperger syndrome test
The childhood asperger syndrome test (CAST) is a parental questionnaire to screen school-age populations for ASD. Although by name it targets the asperger syndrome (AS) by scope it targets the full spectrum, including typical autism and pervasive developmental disorder-not otherwise specified (PDD-NOS); and thus falls within the boundaries of this review.

Table 2 provides a description of the measures reviewed.

DISCUSSION

The results of the reviewed studies will be discussed with focus on the abilities of the four measures to encompass the entire spectrum of autistic disorders and its components.

Ability to discriminate the three major components of autism
Traditionally, autism has been characterized by severe difficulties in three domains; reciprocal social interaction, communication and repetitive stereotyped behavior. Yet, are impairments in all three domains as intrinsic as once believed? Data suggests differently. In fact the genetic influences on individual domains seem independent of one another. Therefore, it is important that each item on a scale/questionnaire should relate to one domain only, allowing the quantification of domain independence.

Although the SRS consists of items individually relating to no more than one domain, it is also based on the belief that difficulties in all three domains are attributable to a singular underlying deficit in reciprocal social behavior (RSB). Hence, whilst the other domains are included independently it is emphasizing their impact on RSB. This focus on RSB almost to a point of exclusion of other domains confounds the SRS’s ability to effectively explore domain independence and identify empirically if RSB is in reality the core deficit. The SCQ was split into four factors: Social interaction, communication, abnormal language and stereotyped behavior, together encompassing the three domains of autism. Individual questionnaire items are linked to one of the three domains. The ABC items were grouped into sensory, relating, body and object use, language, and social and self-help. However, subsequent work grouped ABC items as non-responsive, infant-like, aggressive and stereotypical behaviors, and echolalic speech. The latter study was more empirically based; hence the factors arguably more objective. Nevertheless, the relationship between these factors and the domains remains unclear. For the CAST no factor analytic work was available. Such work may be useful, in view of the CAST’s promise as a screening questionnaire in high-risk populations. Information on domain independence of items is necessary for both the CAST and ABC. The SCQ and ABC should also be administered on community samples and the scores on the different factors/domains examined with regard to their independence/interdependence. By definition and diagnosis a clinical autistic sample will include individuals with difficulties in all three domains. Therefore, to explore independence of domains community samples should be studied. Furthermore, providing separate scores and cutoff points per domain may be useful. It may also be useful to determine if different ASD differ with regard to factor profiles. Such information may lead to clearly defined diagnostic criteria for PDD-NOS.

Ability to measure across the whole spectrum of autism
Recent studies have suggested that autistic symptoms are continuously distributed in the population forming a broad spectrum. Hence, the disorder may take many forms from mild-moderate-severe with a variety of impairments. This increasing heterogeneity of the disorder can cause difficulty in measurement. Measures should be able to quantify subtle differences in the degree of impairment across the spectrum and alert clinicians to the severity of symptoms, including those that may be sub-threshold. Such measures may

Table 2: Description of measures reviewed

| Measure                          | Type of measure | Number of items | Subscales/factors                                      |
|---------------------------------|-----------------|-----------------|--------------------------------------------------------|
| Social responsiveness scale     | Questionnaire   | 65              | Overall score (on reciprocal social behavior)           |
|                                 |                 |                 | Social interaction                                     |
|                                 |                 |                 | Language and communication                             |
|                                 |                 |                 | Repetitive and stereotyped patterns of behavior        |
|                                 |                 |                 | Self-injurious behavior                                |
|                                 |                 |                 | (Current language functioning)                         |
| Social communication questionnaire | Questionnaire   | 40              | Sensory                                               |
|                                 |                 |                 | Non-responsive                                        |
|                                 |                 |                 | Infant-like                                            |
|                                 |                 |                 | Aggressive                                             |
|                                 |                 |                 | Stereotypical (behavior)                               |
|                                 |                 |                 | Echolalic speech                                       |
| Autism behavior checklist       | Rating scale    | 57 (Weighted score per item) | Sensory                                               |
|                                 |                 |                 | Non-responsive                                        |
|                                 |                 |                 | Infant-like                                            |
|                                 |                 |                 | Aggressive                                             |
|                                 |                 |                 | Stereotypical (behavior)                               |
| Childhood autism screening test | Questionnaire   | 37              | Overall score                                         |
|                                 |                 |                 | Social                                                 |
|                                 |                 |                 | Echolalic speech                                       |
also facilitate a clearer concept of PDD-NOS. With reference to the measures reviewed, although they purported to target the full spectrum, it may be that they only ‘claim’ to do so. Therefore, an important question to focus upon is, ‘Are the reviewed measures able to dimensionally tap into and contrast between all points on this continuous spectrum?’

The SRS, which conceptualizes RSB as a continuous variable, will probably be sensitive to any degree of RSB impairment. In fact, it has been observed that SRS scores were continuously distributed in clinical and school samples,[9] supporting the idea of autism as a spectrum condition rather than an all-or-none diagnosis. Nevertheless, it is prudent to be aware that this may merely imply that deficiencies in RSB are continuously distributed, as the SRS cannot justifiably be called a scale that measures the severity of deficiency in all three domains of impairment. Further, it has been observed that the SRS did not clearly distinguish individuals diagnosed with autism from those with PDD-NOS, especially amongst higher functioning autistic cases.[3] The SCQ and ABC are more categorical measures, in that they are generally designed to assign ‘caseness’ as opposed to severity. Hence, they may not tap into the milder cases on the spectrum. For instance, it has been observed that the SCQ may not be suitable for identifying less clear-cut ASD cases.[11] Similarly, the ABC may not reliably screen for high-functioning autistic individuals.[10] The CAST, like the SRS assumes that social and communication difficulties fall on a continuous distribution; and is based on a dimensional conceptualization of the spectrum. The advantage of this dimensionality is that arbitrary cutoff points can be imposed upon the continuum to delineate categories of behavior qualitatively different from ‘normalcy’. [6] However, while the CAST may be sensitive to the less extremely impaired on the spectrum, it also shows a high proportion of false positives. Hence, it may also be tapping onto difficulties outside the clinical spectrum. Regardless of their ability to tap into the entire spectrum of autism and provide independent ratings per domain, these measures would be redundant with weak psychometric properties. Hence, the reliability and validity data on the measures are briefly overviewed.

**Reliability**

**Test-retest reliability**

As presented in Table 3, long-term measurement stability is indicated for the SRS and the German version of the SCQ, albeit in a small sample of 17 subjects.[12] No test-retest data was available for the ABC but given the nature of the scale, such data may be redundant.[13] The CAST seems to have moderately good temporal validity.[6]

**Internal consistency**

Internal consistency for the SRS in data from a school sample was high. Alpha reliability coefficients for the SCQ were also high for both the English and German versions. Split-half reliability data for the ABC were discrepant when two separate studies were compared.[3,10] This may be because the original study used the Spearman-Brown prophesy formula.[5] The overall classification rate on the ABC was virtually unchanged when only the more highly weighted items on the scale were utilized, suggesting that it could be shortened.[14]

**Interrater reliability**

High interrater reliability has been obtained for the SRS,[3] implying universality in the expression of symptoms, and the utility of the measure across multiple informants. ABC data seems discrepant, with recent studies observing lower rates than the original validation data. However, the original study did not incorporate corrections for chance agreement. Further, subsequent studies[10,13] modified the presentation of the scale to blind raters to its aim. Overall, the reliability of the ABC seems unclear and for further reliability data [Table 3].

**Validity**

Most of the validation studies focused on the discriminative ability of the measures. Further, the measures were validated against established diagnostic tools for autism to identify construct-validity. The SRS is capable of distinguishing children with ASD from those with other child psychiatric disorders, including mental

---

**Table 3: Reliability of measures reviewed**

| Measure                                           | Test-retest | Internal consistency | Interrater reliability |
|---------------------------------------------------|-------------|----------------------|------------------------|
| Social responsiveness scale                       | 3 months: 0.88[9] | Cronbach’s Alpha: 0.97[9] | Teachers-fathers: 0.75 |
|                                                   | 27 months: 0.83[3] |                        | Teachers-mother: 0.82   |
| Social communication questionnaire                 | 12-24 months: 0.74[12] | Cronbach’s Alpha: 0.90[9] | Mothers-fathers: 0.91[3] |
| Autism behavior checklist                         | No data retrieved | Split-half reliability: 0.94 (Spearman-Brown prophesy formula)[9] | No data retrieved 95%[9] |
| Childhood asperger syndrome test                  | Good temporal validity[4] | 70% (Kappa ≥.40)[19] | No data identified      |

---

[9,10,12,13]

---

[4] Indian Journal of Psychological Medicine | Jan - Jun 2011 | Vol 33 | Issue 1
Between autism and the other disorders within the spectrum, although significant differentiation was observed, it was less clear-cut and there was substantial overlap. The SRS compares favorably with the ADI-R. Overall, the SRS seems to be a valid instrument for the measurement of clinically significant autistic traits. Understandably, the discriminatory value of the SCQ was significantly correlated with the ADI upon which it was based. Further, the SCQ showed statistically significant differentiation between ASD and other disorders. Interestingly, a relatively weak diagnostic differentiation was provided by the factor representing repetitive stereotyped behaviors. Thus a challenge for researchers is to identify an effective means of measuring this domain. There is also some questionability regarding the SCQ’s ability to differentiate autism from mental retardation. In addition, the SCQ may not be able to differentiate between different ‘points’ on the autism spectrum, and may not be an effective screening tool in a non-clinical population. Results from studies on the discriminative validity of the ABC are wide-ranging. Data indicates that total scores in autistics are consistently higher than in severely mentally retarded, emotionally disturbed, deaf-blind, learning impaired or normal. However, when the more current, broader criteria was used to define autism, a high negative predictive value was obtained, suggesting a need to validate a new cutoff score for the ABC defined by current diagnostic criteria. Similarly, the ABC has been found to be less discriminative with a high number of false-negatives, particularly in mild-moderate autism. With high-functioning autism, when retrospective ratings were used, the ABC was sensitive to its diagnosis. A pilot study demonstrated that the CAST differentiates well between children with AS/autism and normally developing children. However, there is a lack of clarity as to whether the discrimination was between those with autism/AS and typical children, or simply between AS and typical children. A subsequent pilot study on mainstream school samples showed that the CAST has good specificity (.98) on a cutoff point of 15. However, positive predictive value was poor (.64). In addition, the sample studied may have been biased, as there was a poor response rate (17%). Hence, researchers were unable to reliably calculate the sensitivity of the scale. In a subsequent study the CAST demonstrated good accuracy for use as a screening test in a high-risk population, with high sensitivity and specificity. However, once again, it had a low positive predictive value, leading to a high number of false-positives, which could have detrimental implications and for further validity data [Table 4].

CONCLUSIONS AND FUTURE DIRECTIONS

All the measures reviewed within the limits of this analysis are relatively brief [Table 2], and efficient for use in busy clinical and educational settings. Responses are obtained firsthand from informants, allowing greater diagnostic input from primary care-providers who have direct observation of the child’s behavior.

None of the measures reviewed seem potentially useful in differentiating between disorders within the autistic spectrum. The usefulness of the SCQ and ABC may be limited to screening for the more clinically defined ends of the spectrum. In contrast, the CAST and SRS may be more effective to identify individuals

| Measure | Content | Criterion | Construct |
|---------|---------|-----------|-----------|
| Social responsiveness scale | Expert opinion | Differentiates between children with ASD and typical children/children with other psychiatric disorders | Compares favorably with the ADI-R (coefficients >0.64) |
| Social communication questionnaire | Autism diagnostic interview-revised algorithm used for International Classification of disease-10 and diagnostic and statistical manual-IV | Differentiates between children with ASD and other psychiatric disorders | Compares favorably with the ADI-R (coefficient 0.71) |
| Autism behavior checklist | Rimland’s behavior rating instrument for autistic children; British working party’s checklist; Behavior rating instrument for autistic children; Rendle-short and Clancy’s checklist; Lotter’s checklist; Lorda’s and Kanner’s descriptions; Expert opinion; Statistical analysis | Differentiates between children with ASD and typical children/children with other disorders (learning disabled, emotionally disturbed, deaf-blind) | |
| Childhood autism screening test | DSM-IV; ICD-10; Autism spectrum Screening questionnaire items, Pervasive developmental disorder-questionnaire items | Differentiates between ASD and typical children in pilot and community samples | None reported |

The chief drawback in much of the validation data resulted from limitations in sample generalizability and size. Samples may have been biased or sensitized (prior experience on previous measures etc.). Overall there is a need for more validation and normative data in both clinical and non-clinical samples of adequate size to tap into all points of the spectrum.
with difficulties which may/may not be indicative of autism. More information regarding the specificity, sensitivity and positive and negative predictive values of the measurements at varying cutoff points (possibly providing within-spectrum discrimination) will be useful.

It may also be useful to study the distribution of scores in a population, especially in a non-clinical one. The SRS provides limited data in this regard. Further work will enhance current understanding on continuities or discontinuities within the spectrum. An important question requiring further exploration is: ‘Are the differences within this spectrum qualitative or merely quantitative?’.

If autism is indeed a spectrum of continuity, developing dimensional and continuous measures will be advantageous, especially with regard to genetic causation studies. However, if these dimensions measure the severity of symptoms, how should severity be defined? Is it merely the number of symptoms? Is not continuity of symptoms over time an important aspect of severity? Hence, a tool that can measure symptom change/continuity over development would be useful. From the scales reviewed, the SCQ emphasizes qualitative deviance rather than developmental delay/impairment, focusing on lifetime manifestations. The ABC takes into account the ‘age factor’ and provides separate profile charts for different chronological ages. However, none of the measures focus on symptom change/continuity over the life course.

With regard to this review, there is a possibility of a publication bias, especially as examiners’ manuals/catalogues were not reviewed in detail. Further, measures pre-focusing on a defined point on the autistic spectrum were excluded. However, there have been reviews on measures targeting such groups.[19]

In conclusion, therefore, none of the reviewed measures were able to singularly tap into and differentiate between all points of the autistic spectrum. With regard to their sensitivity to measuring the three domains of autism as possibly existing independently from one another, further research is necessary.

REFERENCES

1. Lord C, Volkmar F. Genetics of childhood disorders: XLII. Autism part 1: Diagnosis and assessment in autistic spectrum disorders. J Am Acad Child Adolesc Psychiatry 2002;41:1134-6.
2. Volkmar FR, Lord C, Bailey A, Schultz RT, Klin A. Autism and pervasive developmental disorders. J Child Psychol Psychiatry 2004;45:135-70.
3. Constantino JN, Davis SA, Todd RD, Schindler MK, Gross MM, Brophy SL, et al. Validation of a Brief Quantitative Measure of Autistic Traits: Comparison of the Social Responsiveness Scale with the Autism Diagnostic Interview-Revised. J Autism Dev Disord 2003;33:427-33.
4. Berument SK, Rutter M, Lord C, Pickles A, Bailey A. Autism screening questionnaire: Diagnostic validity. British J Psychiatry 1999;175:444-51.
5. Krug DA, Arick J, Almond P. Behavior checklist for identifying severely handicapped individuals with high levels of autistic behavior. J Child Psychol Psychiatry 1980;21:221-9.
6. Williams J, Scott F, Stott C, Allison C, Bolton P, Baron-Cohen S, et al. The CAST (Childhood Asperger Syndrome Test): Test accuracy. Autism 2005;9:45-68.
7. Ronald A, Happe F, Pomin R. The genetic relationship between individual differences in social and nonsocial behaviours characteristic of autism. Dev Sci 2005;8:444-58.
8. Miranda-Linne FM, Melin L. A factor analytic study of the Autism Behavior Checklist. J Autism Dev Disord 2002;32:181-8.
9. Constantino JN, Przybeck T. Friesen D, Todd RD. Reciprocal social behavior in children with and without pervasive developmental disorders. J Dev Behav Pediatr 2000;21:2-11.
10. Volkmar FR, Cicchetti DV, Dykens E, Sparrow SS. An evaluation of the Autism Behavior Checklist. J Autism Dev Disord 1988;18:81-97.
11. Scott FJ, Baron-Cohen S, Bolton P, Brayne C. The CAST (Childhood Asperger Syndrome Test): Preliminary development of a UK screen for mainstream primary-school-age children. Autism 2002;6:9-31.
12. Bolte S, Crecelius K, Pousta M. The Questionnaire on Behavior and Social Communication (VSK): An autism screening instrument for research and practice. Diagn 2000;46:149-55.
13. Parks SL. The assessment of autistic children: A selective review of available instruments. J Autism Dev Disord 1983;13:255-67.
14. Wadden NP, Bryson SE, Rodger RS. A closer look at the Autism Behavior Checklist: Discriminant validity and factor structure. J Autism Dev Disord 1991;21:529-41.
15. Szatmari P, Archer L, Fisman S, Steiner DL. Parent and teacher agreement in the assessment of pervasive developmental disorders. J Autism Dev Disord 1994;24:703-17.
16. Teal MB, Wiebe MJ. A validity analysis of selected instruments used to assess autism. J Autism Dev Disord 1986;16:485-94.
17. Rellini E, Tortolani D, Trillo S, Carbone S, Montecchi F. Childhood Autism Rating Scale (CARS) and Autism Behavior Checklist (ABC) Correspondence and Conflicts with DSM-IV Criteria in Diagnosis of Autism. J Autism Dev Disord 2004;34:703-8.
18. Yirmiya N, Sigman M, Freeman BJ. Comparison between diagnostic instruments for identifying high-functioning children with autism. J Autism Dev Disord 1994;24:281-91.
19. Campbell JM. Diagnostic Assessment of Asperger’s Disorder: A Review of Five Third-Party Rating Scales. J Autism Dev Disord 2005;35:25-35.

How to cite this article: Fernandopulle N. Measurement of autism: A review of four screening measures. Indian J Psychol Med 2011;33:5-10.
Source of Support: Nil, Conflict of Interest: None.