A Critical Review of Screening and Diagnostic Instruments for Autism Spectrum Disorders in People with Sensory Impairments in Addition to Intellectual Disabilities

Gitta de Vaan, Mathijs P. J. Vervloed, Marella Hoevenaars-van den Boom, Anneke Antonissen, Harry Knoors & Ludo Verhoeven


To link to this article: https://doi.org/10.1080/19315864.2015.1119917
A Critical Review of Screening and Diagnostic Instruments for Autism Spectrum Disorders in People with Sensory Impairments in Addition to Intellectual Disabilities

Gitta de Vaan, Mathijs P. J. Vervloed, Marella Hoevenaars-van den Boom, Anneke Antonissen, Harry Knoors, and Ludo Verhoeven

Introduction

According to the American Psychiatric Association (2013), autism spectrum disorder (ASD) is a developmental disorder characterized by two major components: deficits in communication and social interaction, and repetitive and stereotyped patterns of behavior. The process of diagnosing ASD roughly consists of four general steps (National Institute for Health and Clinical Excellence, 2012; Nederlandse Vereniging voor Psychiatrie, 2009; Volkmar et al., 2014). The first step is to identify any problems or concerns. When there is reason to believe a person might have ASD, the second step is screening for the presence of ASD. Screening is done by talking to parents or caretakers, by studying medical and psychological information and...
history, by making observations, and through using specific screening instruments. When a person screens positive for ASD, the third step is the application of diagnostic instruments. The final step is to make an individual profile to guide treatment (National Institute for Health and Clinical Excellence, 2012; Nederlandse Vereniging voor Psychiatrie, 2009).

Important in this diagnostic process is to combine multiple instruments (Risi et al., 2006) and to incorporate multidisciplinary clinical judgments (Rutter, 2006; Volkmar et al., 2014). Steps two and three may be difficult to conduct in people with both sensory and intellectual disabilities. In this critical review we focus on these steps, the screening and diagnostic instruments commonly used in the assessment of ASD.

Although there is a broad range of instruments that can be used for screening for and diagnosing of ASD, these instruments may not be very valid and/or useful when people develop atypically because of motor, sensory, or intellectual disabilities. For instance, ASD typical behaviors are not only seen in people with ASD but also in people with visual impairments (Cass, 1998; Hobson, Lee, & Brown, 1999), auditory impairments (Knoors & Vervloed, 2011), intellectual disabilities (De Bildt, Sytema, Kraijer, & Minderaa, 2005; Matson & Shoemaker, 2009; Matson, Dempsey, LoVullo, & Wilkins, 2008; Vig & Jedysek, 1999) and also in people with a combination of these impairments (Dammeyer, 2011, 2013; De Vaan, Vervloed, Knoors, & Verhoeven, 2013; Hoevenaars-Van Den Boom, Antonissen, Knoors, & Vervloed, 2009; Rodbroe & Janssen, 2006). It is the latter group that is the focus of this review, people with an intellectual disability combined with a visual impairment or deafblindness. Deafblindness is broadly defined as any combination of both a visual and auditory impairment, and may be congenital or acquired. In this article, no boundaries are set for the severity of visual and auditory impairments (see: Hoevenaars-Van Den Boom et al., 2009; Larsen & Damen, 2014). Especially for this group, clinicians are often asked to assess the presence of ASD. The reason for this is that people with motor, sensory, and intellectual disabilities show many behaviors that topographically look the same as ASD symptoms, but reflect other underlying causes because they may be caused by the respective disabilities instead of by ASD. This is an example of equifinality, the possibility that different conditions may lead to the same outcome. As a result, ASD is both over- as well as under-diagnosed in people with multiple disabilities (Andrews & Wyver, 2005; Cass, 1998; Jure, Rapin, & Tuchman, 1991; Roper, Arnold, & Monteiro, 2003). This can be either because of diagnostic overshadowing, where symptoms are attributed to the most prominent disability (Carvill, 2001; Hoevenaars-Van Den Boom et al., 2009; Mason & Scior, 2004; Reiss, Levitan, & Szyszko, 1982), or because of diagnostic underrepresentation, which refers to falsely missing relevant behaviors. In the current case, some behaviors do not occur and are therefore not measurable in people with
disabilities. For example, eye contact is absent in blind individuals but these people can still be aware of and interested in other people; one only has to measure it in another way (Hoevenaars-Van Den Boom et al., 2009; Kraijer & De Bildt, 2005; Livesley & Jackson, 1992).

Since screening and diagnostic instruments are based on criteria for ASD and ASD-typical behaviors, behavioral overlap with disabilities can cause a decreased usability of ASD instruments in persons with sensory and intellectual disabilities. Despite the ample availability of screening and diagnostic instruments for people with intellectual disabilities alone (Matson & Williams, 2014), there is a lack of instruments that are adjusted to the behaviors of persons with the combination of intellectual and sensory disabilities. Many test instruments, not only for ASD but also for other pathologies, assume that the person under study is able to see and hear (Tobin & Hill, 2011). As was noted by (Bodsworth, Clare, Simblett, & Deafblind UK, 2011), this leads to a lack of suitable instruments to assess people with multiple disabilities. This is why many unsuitable instruments are still used in clinical practice.

As screening and diagnostic instruments play such an important part in the process of diagnosing ASD, the current critical review focuses on the question how valid existing instruments are for the assessment of ASD in people with sensory and intellectual disabilities. This review took into account not only commonly used screening and diagnostic instruments, but also instruments with another focus that partly assess ASD or behaviors that are typical for ASD (see American Psychiatric Association, 2013), as these are often used in ASD assessments. We looked at the quality of the assessment material, ease of use, the psychometric properties (reliability and content validity), the presence of norms for people with sensory and intellectual disabilities, and the applicability of all individual items. This review will provide insights for practitioners and researchers into which instruments are suitable for the diagnosis and assessments of ASD in people who have combined sensory and intellectual disabilities.

**Method**

**Materials**

Screening and diagnostic instruments for ASD were selected, as well as assessment instruments related to characteristics of ASD. Possible instruments were gathered based on a review of existing instruments for the detection and assessment of ASD (O’Brien, Pearson, Berney, & Barnard, 2001), and a literature search using the following keywords: diagnosis, assessment, instruments, screening, autism, autism spectrum disorders, intellectual disabilities, visual impairments, and multiple disabilities. Instruments
were only included if they were available through purchase or online download by the authors in 2014 and 2015 and were available in English or Dutch. Only instruments for which psychometric data were available were included. Of the 14 instruments described in the aforementioned review, 10 were included in the current review in the same or different version. Our search led to a selection of 13 screening and diagnostic instruments and 7 instruments assessing ASD-typical characteristics.

**Procedure**

The characteristics of the instruments were assessed according to the guidelines of the Dutch committee on tests and testing (COTAN) (Evers, Sijtsma, Lucassen, & Meijer, 2010) and the BUROs center for testing (see: www.buros.org). The BUROs center is a large testing review center that is part of the University of Nebraska-Lincoln that evaluates tests on general characteristics, development and psychometric properties (BUROs, 2015). The COTAN evaluates tests on their theoretical soundness, quality of materials and manual, norms, and psychometric properties (COTAN, 2015). COTAN reviews of tests as well as reviews of the BUROs center for testing were added if they were available for our selected instruments. For every instrument the following characteristics were assessed: (1) availability and quality of the manual; and (2) the scientific foundations of the development, reliability, and validity of the instrument.

For the current study, the following information was collected from manuals, scientific literature, and judgments by COTAN or Buros: (1) name and abbreviation of the instrument, authors, and year of publication; (2) purpose of the instrument: screening, diagnosis, evaluation, or which characteristic it evaluates; (3) number of items; (4) duration of the assessment; (5) whether a manual is available or whether the test can be downloaded online; (6) training requirements to use this instrument and which professionals can use it; (7) which method the instrument uses: checklist, interview, or observation; (8) in case of the screening and diagnostic instruments, the source of the ASD criteria, for example DSM-IV, DSM-5, ICD-10 and/or scientific literature; (9) for which target group the instrument was developed; (10) languages in which this instrument is available (however, we are aware that it is possible that more translations may exist that we don’t know of); (11) reliability, and (12) test validity (for reliability and validity multiple sources were used, and in cases of contradictory results all results were reported along with references to the original source; (13) the availability of norms for people with both sensory and intellectual disabilities; and (14) the number of inappropriate items for people with both sensory and intellectual disabilities.
Item appropriateness for people with both sensory and intellectual disabilities was rated by the first author, in collaboration with the other authors, who all have expertise and clinical experience in the field of sensory and intellectual disabilities. These are individuals who besides an intellectual disability (as defined by the American Association on Intellectual and Developmental Disabilities, 2013) also have a visual impairment or deafblindness.

Topographically the same behaviors can have different causes or functions for people with sensory impairments or intellectual disabilities and people with ASD. Instruments for ASD assessment normally check only for symptoms and not the underlying cause. As a result, test items can sometimes be invalid for assessing ASD in children with sensory impairments and intellectual disabilities. Items were rated as inappropriate or sensory biased if at least one of the five following criteria applied to them. Criteria were based upon generally known behaviors of children with sensory impairments (e.g. Knoors & Marschark, 2014; Pérez-Pereira & Conti-Ramsden, 2005; Pring, 2005; Warren, 1994). First, the absence of a behavior is an obvious direct consequence of the sensory impairment. For example, gaze following is impossible for blind people and showing a reaction to speech or sound for deaf people. Second, the behavior is caused by indirect or long-term effects of the sensory impairment, such as language impairments in deaf people or odd or clumsy motor behaviors (e.g., to prevent collisions) in blind people. Third, the behavior is a characteristic that develops differently or more slowly in people with sensory impairments. An example is “theory of mind” (ToM), which develops later in children with blindness or deafness. Fourth, the behavior is more likely to be adaptive for people with sensory impairments than typical for ASD. Examples are odd body postures to hear someone better or to focus vision, and head nodding to counteract eye movements caused by nystagmus. Fifth, the behavior is used for compensatory purposes, an example is echolalia in children who are blind. Echolalia, for a blind child, is an expression of practicing language by repeating over and over pieces of speech or to check for the presence of an unseen conversation partner (Pérez-Pereira & Conti-Ramsden, 1999).

Results

Screening and Diagnostic Instruments for ASD

Of the reviewed screening and diagnostic instruments (Table 1), nine were designed for screening purposes and three for diagnostic purposes. Instruments were created (or revised) between 1978 and 2009. Number of items ranged from 12 to 206, with the screening instruments having the lowest number of items. No training is required for screening instruments.
<table>
<thead>
<tr>
<th>Name of Instrument, Authors, and Sources</th>
<th>Purpose</th>
<th>Number of Items</th>
<th>Administration (in minutes)</th>
<th>Manual</th>
<th>Training/ expertise</th>
<th>Method</th>
<th>Criteria Based On</th>
<th>Target Population</th>
<th>Languages</th>
<th>Reliability</th>
<th>Validity</th>
<th>Norms for people with sensory and intellectual disabilities</th>
<th>Non-applicable item</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ABC</strong></td>
<td>Autism Behavior Checklist</td>
<td>Screening</td>
<td>57</td>
<td>10–20</td>
<td>Checklist available online + part of larger assessment tool (ASIEP)</td>
<td>Experience or using training tapes is recommended</td>
<td>Checklist</td>
<td>Checklists and characteristics of people with ASD.</td>
<td>18 months–adulthood</td>
<td>English</td>
<td>Interrater reliability—Good (Olmi, 1998); Internal consistency—Good on total score, poor on subscales (O’Brien et al., 2001)</td>
<td>Content validity, Concurrent validity &amp; Criterion validity—Well established (Olmi, 1998).</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>ADI-R</strong></td>
<td>Autism Diagnostic Interview-Revised</td>
<td>Diagnosis</td>
<td>93</td>
<td>90–150 (including scoring)</td>
<td>Available</td>
<td>Experience with ASD and interviewing. Need to study manual, no additional training required.</td>
<td>Interview</td>
<td>ICD-10 &amp; DSM-IV</td>
<td>Early child-adulthood, mental age above 2 years.</td>
<td>Over 15 languages including English, French, Spanish and Dutch</td>
<td>Interrater reliability, test-retest reliability, internal consistency—High (O’Brien et al., 2001); Interrater-and-test-retest reliability—Good (Rutter et al., 2003)</td>
<td>Construct validity with ICD-10 &amp; Convergent validity—Good (O’Brien et al., 2001); Discriminative validity—Good (Rutter et al., 2003)</td>
<td>No</td>
</tr>
<tr>
<td><strong>ADOS</strong></td>
<td>Autism Diagnostic Observation Schedule, module 1</td>
<td>Diagnosis</td>
<td>37</td>
<td>30–60</td>
<td>Available</td>
<td>Training is necessary, at least 2 days of training + additional practice</td>
<td>Assessment + Observation</td>
<td>DSM-IV &amp; ICD-10</td>
<td>Children and adults; 15 months and older (module 1 = preverbal)</td>
<td>Primary school children</td>
<td>Interrater reliability, test-retest reliability—High (O’Brien et al., 2001)</td>
<td>Construct validity with ICD-10 &amp; Convergent validity—Good (O’Brien et al., 2001)</td>
<td>No</td>
</tr>
<tr>
<td><strong>ASAS</strong></td>
<td>Australian Scale for Asperger Syndrome</td>
<td>Screening</td>
<td>25</td>
<td>10–15</td>
<td>Available</td>
<td>Not reported</td>
<td>Questionnaire</td>
<td>Not reported</td>
<td>Primary school children</td>
<td>English, Dutch, German</td>
<td>Not reported (O’Brien et al., 2001)</td>
<td>Not reported (O’Brien et al., 2001)</td>
<td>No</td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Name of Instrument, Authors, and Sources</th>
<th>Purpose</th>
<th>Number of Items</th>
<th>Administration (in minutes)</th>
<th>Manual</th>
<th>Training/ expertise</th>
<th>Method</th>
<th>Criteria Based On</th>
<th>Target Population</th>
<th>Languages</th>
<th>Reliability</th>
<th>Validity</th>
<th>Norms for people with sensory and intellectual disabilities</th>
<th>Non-applicable item†</th>
</tr>
</thead>
<tbody>
<tr>
<td>AUTI-R: Van Berckelaer-Onnes and Hoekman (1991); Hoekman (1992); Evers et al. (2010)</td>
<td>Screening</td>
<td>51</td>
<td>–</td>
<td>Available</td>
<td>–</td>
<td>Interview/Questionnaire</td>
<td>Literature, construct of “early childhood autism”</td>
<td>Verbal and non-verbal children from 10–155 months.</td>
<td>Dutch</td>
<td>Internal consistency—Good (Evers et al., 2010); Good to very good (Hoekman, 1992); Test-retest reliability—Good &amp; Inter-rater reliability—Good (Evers et al., 2010; Van Berckelaer-Onnes &amp; Hoekman, 1991) Content validity &amp; Criterion validity—good (O’Brien et al., 2001)</td>
<td>No</td>
<td>1/2</td>
<td></td>
</tr>
<tr>
<td>CARS: Childhood Autism Rating Scales Schopler, Reichler &amp; Renner (1988); Schopler, Reichler, DeVellis, and Daly (1980); Malcolm (2014); O’Brien et al. (2001); McLellan (2014)</td>
<td>Screening</td>
<td>15</td>
<td>30-60</td>
<td>Available</td>
<td>For psychologist or therapists, experience in assessments required</td>
<td>Observational assessment</td>
<td>ASD criteria described in literature</td>
<td>Children two ears and older.</td>
<td>English, Dutch, Swedish, Japanese and more</td>
<td>Internal consistency—Good (Malcolm, 2014; O’Brien et al., 2001); Acceptable (McLellan, 2014); Inter-rater reliability—Acceptable (McLellan, 2014); Adequate to High (O’Brien et al., 2001)</td>
<td>Construct validity—Moderate to Good (Malcolm, 2014); Construct and Convergent validity—established O’Brien et al., 2001)</td>
<td>No</td>
<td>2/3</td>
</tr>
</tbody>
</table>

(Continued)
<table>
<thead>
<tr>
<th>Name of Instrument, Authors, and Sources</th>
<th>Purpose</th>
<th>Number of Items</th>
<th>Administration (in minutes)</th>
<th>Manual</th>
<th>Training/ expertise</th>
<th>Method</th>
<th>Criteria Based On</th>
<th>Target Population</th>
<th>Languages</th>
<th>Reliability</th>
<th>Validity</th>
<th>Norms for people with sensory and intellectual disabilities</th>
<th>Non-applicable item</th>
</tr>
</thead>
<tbody>
<tr>
<td>DISCO</td>
<td>Diagnosis</td>
<td>105 items assessing history</td>
<td>3 hours</td>
<td>—</td>
<td>Needs adequate clinical experience, training is possible</td>
<td>Interview</td>
<td>Earlier instruments and ASD criteria</td>
<td>Children and adults</td>
<td>English, Dutch</td>
<td>Not reported (O’Brien et al., 2001)</td>
<td>Not reported (O’Brien et al., 2001)</td>
<td>No</td>
<td>History: 1/5 Current: 1/3</td>
</tr>
<tr>
<td>ESAT</td>
<td>Screening</td>
<td>14</td>
<td>Available</td>
<td>Experience with ASD screening needed. Additional training in manual</td>
<td>Interview/checklist</td>
<td>Literature on predictors of ASD</td>
<td>Children under 20 months</td>
<td>Dutch, English</td>
<td>Test-retest reliability—Insufficient (Evers et al., 2010), Very good (Butelaar et al., 2009)</td>
<td>Content validity—Insufficient (Evers et al., 2010); Criterion validity—Insufficient (Evers et al., 2010); Good predictive validity (Butelaar et al., 2009)</td>
<td>No</td>
<td>1/3</td>
<td></td>
</tr>
<tr>
<td>M-CHAT-R/F</td>
<td>Screening</td>
<td>20</td>
<td>Available online</td>
<td>Little to no training necessary for health care professionals</td>
<td>Checklist</td>
<td>Literature, clinical instruments, clinical experience</td>
<td>Young children</td>
<td>Over 15 languages including English, French, Spanish and Dutch</td>
<td>Test-retest reliability—Good (O’Brien et al., 2001)</td>
<td>Sensitivity and Specificity—Good (O’Brien et al., 2001)</td>
<td>No</td>
<td>3/4</td>
<td></td>
</tr>
<tr>
<td>Name of Instrument, Authors, and Sources</td>
<td>Purpose</td>
<td>Number of Items</td>
<td>Administration (in minutes)</td>
<td>Manual</td>
<td>Training/ expertise</td>
<td>Method</td>
<td>Criteria Based On</td>
<td>Target Population</td>
<td>Languages</td>
<td>Reliability</td>
<td>Validity</td>
<td>Norms for people with sensory and intellectual disabilities</td>
<td>Non-applicable items</td>
</tr>
<tr>
<td>----------------------------------------</td>
<td>---------</td>
<td>----------------</td>
<td>-----------------------------</td>
<td>--------</td>
<td>---------------------</td>
<td>--------</td>
<td>-------------------</td>
<td>-------------------</td>
<td>-----------</td>
<td>-------------</td>
<td>---------</td>
<td>---------------------------------------------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>PDD-MRS (AVZ-R)</td>
<td>Screening</td>
<td>12</td>
<td>10–30</td>
<td>Available</td>
<td>Experience with ASD, no additional training.</td>
<td>Checklist</td>
<td>DSM-III, other screening tools, literature</td>
<td>Children and adults with intellectual disabilities; aged 2–70 years</td>
<td>Dutch, English, German, Italian</td>
<td>Interrater reliability, internal consistency &amp; test-retest reliability—Good (Evers et al., 2010; Meadows, 2007)</td>
<td>Content validity—Good &amp; Criterion validity—sufficient (Evers et al., 2010); Construct validity—Satisfactory (Meadows, 2007); Sensitivity and specificity—Good (Kraijer, 1999; O’Brien et al., 2001)</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>SRS</td>
<td>Screening</td>
<td>65</td>
<td>15–20</td>
<td>Available</td>
<td>For psychologist or therapists, experience in assessments required</td>
<td>Checklist</td>
<td>DSM-5</td>
<td>Children 4-17</td>
<td>English, Dutch, German</td>
<td>Internal consistency—Excellent (Hoff &amp; Doepke, 2014) Good (Evers et al., 2010); Test-retest reliability—Moderate to strong (Hoff &amp; Doepke, 2014)</td>
<td>Criterion validity—Good (Evers et al., 2010); Sensitivity &amp; Specificity—Excellent &amp; Convergent validity—Low to moderate (Hoff &amp; Doepke, 2014)</td>
<td>No</td>
<td></td>
</tr>
</tbody>
</table>

Note: † Proportion of non-applicable items for people with both sensory and intellectual disabilities as judged by the authors according to the criteria stated in the method section.
according to their manuals, but at least some training or experience with ASD is required for diagnostic instruments. The screening instruments are either checklists or interviews. Of the diagnostic instruments, the ADI-R (Rutter, Le Couteur, & Lord, 2003) and DISCO (Wing, 2003) are interviews and the ADOS is an observation (Lord, Rutter, DiLavore, & Risi, 1999). When it comes to the ASD criteria on which these instruments are based, the screening instruments were all based on scientific literature, while the diagnostic instruments took into account the ASD criteria from the DSM-IV (American Psychiatric Association, 2000) and ICD-10 (World Health Organization, 1992).

The reliability of the instruments ranged from poor to excellent. The diagnostic instruments showed moderate to excellent reliability, whereas reliability was poor to excellent for the screening instruments. In case of the ABC, (Krug, Arick, & Almond, 1978) the reliability alone ranged from poor to excellent. This wide range of quality is the result of the fact that different sources reported different types of reliability. Their first research paper (Krug, Arick, & Almond, 1980) reported excellent intrarater reliability and good interrater reliability, but later research found low split-half reliability on scales of language and social or self-help (Volkmar et al., 1988). The other instruments showed more straightforward results when it came to validity and reliability. The results ranged from good and very good for the ADOS (Lord et al., 1999) to insufficient for the ESAT (Buitelaar et al., 2009).

Concerning their applicability for people with both sensory and intellectual disabilities, none of the instruments has norms for this target group except the ABC (Krug et al., 1978). However, though the ABC has norms for people with deafblindness, these norms are relatively old because they stem from 1978 and the sensory disabilities were not taken into account during the development of the instrument. The latter is also seen in the large number of inappropriate items for people with both sensory and intellectual disabilities. Of the instruments that do not have norms for people with sensory and intellectual disabilities, at least one third of the items of were not applicable for use in people with sensory and intellectual disabilities according to one or more of the five validity criteria described in the method section.

**Instruments Assessing Characteristics of ASD**

The instruments assessing characteristics of ASD, shown in Table 2, assess a variety of characteristics, such as: communication, social behavior, repetitive behavior, social functioning, ToM, and adaptive skills. Their purposes are mostly to assess the skill level or the severity level of the target behaviors. Only the CCC-2 is a screening instrument, screening for language impairment (Bishop & Geurts, 2007), and the ComFor, which not only assess level of understanding
<table>
<thead>
<tr>
<th>Name of Instrument, Authors, and Sources</th>
<th>Purpose</th>
<th>Number of Items</th>
<th>Duration (in minutes)</th>
<th>Manual Training/expertise Method</th>
<th>Target population</th>
<th>Languages</th>
<th>Reliability</th>
<th>Validity</th>
<th>Norms for people with sensory and intellectual disabilities</th>
<th>Non applicable item</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CCC-2</strong> Children’s Communication Checklist—2 (NL, Dutch translation) Bishop and Geurts (2007); Evers et al. (2010); McCauley (2010)</td>
<td>Screen for language impairments, pragmatic problems &amp; assist in ASD assessment</td>
<td>70</td>
<td>5–10 Available</td>
<td>General knowledge about tests and familiarity with the CCC-2 is required</td>
<td>Checklist</td>
<td>Children between 4–15 years, speak in full sentences, normal hearing</td>
<td>English, Dutch</td>
<td>Internal consistency—Sufficient (Bishop &amp; Geurts, 2007; Evers et al., 2010), Strong (McCauley, 2010), Test-retest reliability—Good to Excellent (McCauley, 2010), sufficient to Good (Bishop &amp; Geurts, 2007); Content validity—Good (McCauley, 2010)</td>
<td>Construct validity &amp; predictive validity—Insufficient (Evers et al., 2010); Convergent validity—Sufficient &amp; Divergent validity—good (Bishop &amp; Geurts, 2007); Content validity—Good (McCauley, 2010)</td>
<td>No 3/4</td>
</tr>
<tr>
<td><strong>ComFor</strong> Foresrunners in Communication Verpoorten et al. (2004); Noens et al. (2006); Evers et al. (2010)</td>
<td>Assess the most suitable form of augmentative communication and assess level of sense-making</td>
<td>36 45 Available</td>
<td>Psychologist, speech therapists with assessment qualification. Additional ComFor training is required. Assessment and observation</td>
<td>People with ID and ASD without or with few verbal communication skills</td>
<td>Checklist</td>
<td>Children 4–18 years; children with intellectual disabilities 4–18 years, children with PDD-NOS, ADHD, high-functioning ASD</td>
<td>English, French, Italian, Dutch</td>
<td>Internal consistency—Good, Inter-rater reliability—very high, Test-retest reliability—good (Noens et al., 2006), Overall reliability—Sufficient (Evers et al., 2010)</td>
<td>Construct validity—Good (Noens et al., 2006); Content validity—Insufficient (Evers et al., 2010)</td>
<td>No All</td>
</tr>
<tr>
<td><strong>CSBQ (VISK)</strong> Children’s Social Behavior Questionnaire Hartman et al. (2007); Evers et al. (2010); Luteijn, Luteijn, Jackson, Volkmar, and Minderaa (2000); De Bildt et al. (2009)</td>
<td>Assess social behavior in child with a pervasive developmental disorder</td>
<td>49</td>
<td>10 Available</td>
<td>Interpretation needs to be done by trained psychologist, pedagogue or psychiatrist</td>
<td>Checklist</td>
<td>Children 4–18 years; children with intellectual disabilities 4–18 years, children with PDD-NOS, ADHD, high-functioning ASD</td>
<td>English, Dutch</td>
<td>Internal consistency—Established (Hartman et al., 2007; Luteijn et al., 2000); Content validity—Sufficient (Evers et al., 2010)</td>
<td>Construct validity—Established (Hartman et al., 2007; Luteijn et al., 2000); Content validity—Sufficient (Evers et al., 2010)</td>
<td>No 2/5</td>
</tr>
<tr>
<td><strong>RBQ</strong> Repetitive Behaviors Questionnaire Honey et al. (2012); Van Kempen, De Vaan, and Versloot (2013)</td>
<td>Assessment of repetitive or stereotyped behaviors</td>
<td>33 — Questionnaire only, available online</td>
<td>— Questionnaire</td>
<td></td>
<td>Children with autism, 4–16 years</td>
<td>English, Dutch, Hebrew</td>
<td>Internal consistency—Good (Honey et al., 2012)</td>
<td>Construct validity—Insufficient &amp; Concurrent validity—Good (Honey et al., 2012)</td>
<td>No 1/4</td>
<td></td>
</tr>
</tbody>
</table>

(Continued)
Table 2. (Continued).

<table>
<thead>
<tr>
<th>Name of Instrument, Authors, and Sources</th>
<th>Purpose</th>
<th>Number of Items</th>
<th>Duration (in minutes)</th>
<th>Manual Training/expertise</th>
<th>Method</th>
<th>Target population</th>
<th>Languages</th>
<th>Reliability</th>
<th>Validity</th>
<th>Norms for people with sensory and intellectual disabilities</th>
<th>Non applicable item</th>
</tr>
</thead>
<tbody>
<tr>
<td>SRZ/SRZ-i: Social Functioning Scale for the Mentally Retarded (-interview) Kraijer et al. (2004); Evers et al. (2010)</td>
<td>Assess social functioning in individuals with intellectual disabilities</td>
<td>31</td>
<td>10-25</td>
<td>Available</td>
<td>Questionnaire/interview</td>
<td>Children and adults with intellectual disabilities, 4 years and older</td>
<td>Dutch</td>
<td>Internal consistency—Good, interrater reliability—good, Test-retest reliability—good (Evers et al., 2010; Kraijer et al., 2004)</td>
<td>Content validity—&amp; criterion validity— Good (Evers et al., 2010); construct validity—Good (Kraijer et al., 2004)</td>
<td>No</td>
<td>1/4</td>
</tr>
<tr>
<td>TOM-test-R: Theory of Mind test Revised, 2009 Steerneman and Meesters (2009); Evers et al. (2010)</td>
<td>Assess Theory of Mind and precursors</td>
<td>14</td>
<td>20</td>
<td>Available</td>
<td>Needs to be assessed by clinicians</td>
<td>Interview</td>
<td>Children 4–12 years old</td>
<td>Dutch</td>
<td>Overall reliability insufficiently studied (Evers et al., 2010); internal consistency—Good (Steerneman &amp; Meesters, 2009), acceptable (Evers et al., 2010); test-retest reliability—Satisfactory (Steerneman &amp; Meesters, 2009), good (Evers et al., 2010); interrater reliability—Good (Steerneman &amp; Meesters, 2009)</td>
<td>Content validity (Steerneman &amp; Meesters, 2009); content validity—Insufficient (Evers et al., 2010) criterion validity—not studied (Evers et al., 2010)</td>
<td>No</td>
</tr>
<tr>
<td>Vineland-Z: Vineland-Z, for children and youth with an intellectual disability (Dutch manual) De Bildt et al. (2003); Evers et al. (2010)</td>
<td>Assess level of adaptive functioning</td>
<td>225</td>
<td>20-60</td>
<td>Available</td>
<td>Necessary</td>
<td>Interview</td>
<td>Children and youth with an intellectual disability, 4–18 years</td>
<td>English, Dutch, Spanish</td>
<td>Internal consistency—Good (Evers et al., 2010)</td>
<td>Content validity—Good; criterion validity—Good; construct validity—Good (Evers et al., 2010)</td>
<td>No</td>
</tr>
</tbody>
</table>

Note. † Proportion of non-applicable items for people with both sensory and intellectual disabilities as judged by the authors according to the criteria stated in the method section.
but also what kind of augmentative communication a person requires (Verpoorten, Noens, & Van Berckelaer-Onnes, 2004). The number of items ranges from 14 to 225, and test administration varies between 10 and 60 minutes. The CCC-2, CSBQ (Hartman, Luteijn, Moorlag, De Bildt, & Minderaa, 2007) and ToM-Test-R (Steerneman & Meesters, 2009) require an experienced clinician for the assessment and interpretation of the results, but additional training for administrators is only necessary for the ComFor and the Vineland-Z (De Bildt, Kraijer, Sparrow, Balla, & Cicchetti, 2003). The methods of administration are: assessment (1x), checklists (2x), interviews (2x), questionnaire (1x), and combined-interview questionnaire (1x). The instruments are typically designed for (young) children, though the SRZ (Kraijer, Kema, & De Bildt, 2004), the ComFor, and Vineland-Z are also designed for people with intellectual disabilities.

Similarly to the screening and diagnostic instruments, the psychometric properties range from insufficient to good. However, none of the instruments in Table 2 have norms for people with both intellectual and sensory disabilities. Similarly to the screening and diagnostic instruments, at least a quarter the items are inappropriate for people with sensory and intellectual disabilities.

Discussion

All instruments show at least adequate psychometric properties on some aspects, though mixed results have been found for the ABC (Krug et al., 1978), ESAT (Buitelaar et al., 2009), CARS (Scholper, Reichler, & Renner, 1988), PDD-MRS (Kraijer & De Bildt, 2005), PDDST-II (Siegel, 2004), SRS (Constantino & Gruber, 2005), CCC-2 (Bishop & Geurts, 2007), ComFor (Verpoorten et al., 2004), CSBQ (Hartman et al., 2007), and the ToM-test-R (Steerneman & Meesters, 2009), and no psychometric properties were reported for the ASAS (Garnett & Attwood, 1993) and DISCO (Wing, 2003). However, the remaining instruments appear to be suitable for the intended purpose and target group. The only instrument with norms for people with deafblindness is the ABC (Krug et al., 1978); however, these norms are the result of their 1978 study and are likely to be outdated. None of the other instruments have norms for people with both sensory and intellectual disabilities.

More importantly than the lack of norms, however, is the sensory bias in all of the instruments. In cases of people with sensory impairments, test items either cannot be assessed or cannot be interpreted as signs of ASD. In all instruments, at least one quarter of items show this bias and are therefore inappropriate for assessing ASD or ASD behaviors in people with sensory and intellectual disabilities. The manual of the AVZ-R, an instrument designed for people with intellectual disabilities, acknowledges that some
items are inappropriate for people with additional sensory impairments, but claims that this does not affect the total score (Kraijer, 1999). This conclusion seems illogical; taking into account the number of inappropriate items we think it is impossible to make a valid diagnosis. For several instruments, especially the screening instruments such as the ESAT (Buitelaar et al., 2009) and M-CHAT (Robins, Fein, & Barton, 2009), the cutoff score for ASD is rather low. Without taking into account possible invalid items, a person with sensory and intellectual disabilities would easily score within the clinical range of ASD on these instruments regardless of the actual presence of ASD. The large number of false positive scores reduces test specificity, but at the same time proper ASD symptoms in people with both sensory and intellectual disabilities are also missed. The sensory bias reduces specificity of the reviewed instruments, but also their sensitivity for true ASD symptoms.

If one looks at the five criteria for inappropriate items, the first conclusion to be drawn is that they cannot be used because the behaviors result not from ASD but are consequences of sensory impairments. This problem was found in all of the instruments. Items such as “Does the person make eye contact?” or “Does the person respond to calling their name?” measure abilities that cannot be measured in all people with sensory impairments. These examples are very straightforward, but this problem also occurs more subtly in items that measure gaze following, pointing, showing, and making conversation. Not only should the direct consequences of sensory impairments be taken into account; the indirect consequences of sensory impairments are also important. Many people with congenital deafblindness do not speak but communicate with sign language or gestures (Dalby et al., 2009), pictures or objects (Noens, Berckelaer-Onnes, Verpoorten, & Van Duijn, 2006). Especially when there is an additional intellectual disability, the typical language impairments of ASD such as echolalia (Lin, 2014; Roberts, 2014) do not occur in this target group, simply because most of them do not use speech to communicate. Not taking this into account would lead to diagnostic underrepresentation of ASD (Kraijer & De Bildt, 2005; Livesley & Jackson, 1992). When a behavior cannot occur one should not attempt to measure this for diagnostic purposes, and subsequently norms or cutoff points should be adjusted.

A construct such as ToM can be present in people with visual impairments, but cannot be tested in the same way as in people without visual impairments. The ToM test that was included in this review (Steerneman & Meesters, 2009) and other ToM tasks, such as false belief tasks such as the Sally Ann task (Baron-Cohen, Leslie, & Frith, 1985), all use visual stimuli to assess the presence of ToM. These stimuli cannot be properly perceived by visually impaired and blind people. The problem for deaf people is that test instructions are verbal and require good speech and language skills. Deaf children often fail ToM tasks not because they do not have ToM skills but
because of inappropriate language skills (Peterson, Wellman, & Liu, 2005). Also, when it comes to communication, people with hearing impairments often use sign language and people with deafblindness communicate through tactile signing (Miles, 2003), pictures, or objects (Noens et al., 2006), communication forms not taken into account by the reviewed instruments which largely rely on oral language to assess communication.

An important issue to consider is the level of intellectual disability and the developmental age of the person assessed in combination with the nature and severity of their sensory impairments. Some behaviors, such as pretend play or ToM, depend on cognitive skills that typically develop after a certain developmental stage is reached. The performance of persons functioning below this level cannot be interpreted in the same way as one would do for someone who does function at or above this developmental level (De Jonge & De Bildt, 2003). With respect to ToM, we now know that blind children are capable of having a ToM, but the development of ToM in children who are blind without intellectual disabilities takes about two years extra compared to sighted children (Brambring & Asbrock, 2010). With additional intellectual disabilities, this delay will surely be longer or ToM may not even develop at all. Developmental delays are also seen for play behavior. Blind children engage more in solitary play than children without visual impairments (Tröster & Brambring, 1994) and show less symbolic play at the same ages as sighted children (Hughes, Dote-Kwan, & Dolendo, 1998), and again, intellectual disabilities will increase these delays.

Adherence to routines is often seen as characteristic of ASD. However, for people with sensory impairments routines are important to get a grip on life, especially with limited options for communication. People who cannot see or hear need routines to understand where they are going, what they are doing, or what they can expect. Therefore adherence to routines cannot be a differentiating factor in itself if one does not check for the perseverance or ability to stop the routine or repetitive behavior (Gense & Gense, 2005). Finally, communication develops differently in people with sensory and intellectual disabilities. They use other modes of communication than spoken language. In addition, social skills such as showing empathy, expressing moral emotions, and supporting peers are shown less frequently in deaf children, likely because they have fewer opportunities to incidentally learn these behaviors (Ketelaar, Wiefferink, Frijns, Broekhof, & Rieffe, 2015; Netten et al., 2015). The social partner in communication also plays a key role in the quality of social interactions with someone who is deafblind and has an intellectual disability (Damen, 2015). When a person’s social partners do not adjust their way and mode of communication properly, test scores could easily underrate a person’s true ability.
Implications

The fact that many items are inappropriate and there are no norms does not mean an ASD instrument is completely worthless. In fact, with care, instruments could potentially still be valuable in the assessment of people with sensory and intellectual disabilities. Their use for screening diagnostic purposes is however limited. The ADI-R, for example, can be used in people with deafblindness to assess their clinical needs (Rutter et al., 2003). Some instruments, especially the ones assessing characteristics of ASD, could also be helpful to assess progress over time, and to evaluate skill levels or whether treatment goals are reached. Though a number of test items are definitely not suitable, many are. It can still be clinically relevant to see how someone’s social or adaptive skills have changed over time. As long as the main goal of the instrument is not to decide on a diagnosis of ASD it can still be used to assess someone’s level of functioning. Instruments that partly assess ASD, such as the CSBQ (Hartman et al., 2007), the RBQ (Honey, McConachie, Turner, & Rodgers, 2012), the SRZ (Kraijer et al., 2004), and the Vineland-Z (De Bildt et al., 2003) are applicable for use in people with intellectual and sensory disabilities for their intended use but not for diagnosing ASD in this group because the norms no not apply. In our opinion, the clinician involved in the assessment must in all cases have expertise on and experience with people with sensory and intellectual disabilities, whether or not instruments are used. Keep in mind, however, that the clinical opinion can be biased too, especially when instruments cannot give a clear answer (De Bruyn, 2006). We therefore recommend multidisciplinary assessments and the use of multiple instruments if possible (Risi et al., 2006; Rutter, 2006; Volkmar et al., 2014).

A solution to the validity problem might be to adjust test items or instruments. For example, the ComFor (Verpoorten et al., 2004) was not usable at all in people with sensory impairments, so the authors recently developed the ComFor-V, an adaptation suitable for people with both intellectual disabilities and visual impairments (KU Leuven, 2015). For other instruments, test items can be adjusted to make them more appropriate for people with sensory impairments and intellectual disabilities. One can replace spoken words by sign language to assess communication, for example, in people who are deaf. Items assessing joint attention can also be adjusted for people who are blind, for example by changing gaze following to a more appropriate form of inferring attention such as freezing or motor movements, a change in breathing, signs of concentrated listening, or tactile cues given by the blind person in tactile signing (Miles, 2003). The obvious downside of adjusting items is that norms are no longer valid and existing research regarding validity and reliability no longer applies to the adjusted version. Furthermore, as these constructs often develop differently in people
with sensory and intellectual disabilities, the interpretation of the results should always be done with care. Extra care is necessary since in the assessment of people with multiple disabilities, adjustments are often required on an individual level as not everyone has the same level of visual impairment, auditory impairment, or communication skills (Boers, Janssen, Minnaert, & Ruijsseenaars, 2013). In these cases, individual progress can be measured but one cannot compare between individuals.

**Conclusion**

Commonly used instruments that were designed to assess the presence of ASD or characteristics of ASD were reviewed for their use in people with sensory and intellectual disabilities. The validity and reliability of these instruments have in most cases been scientifically supported for people without disabilities; the ADOS (Lord et al., 1999) and ADI-R (Rutter et al., 2003) are considered to be the preferred instruments in ASD assessment (De Bildt et al., 2004; Reaven, Hepburn, & Ross, 2008). However, this does not make them applicable for use in people with sensory and intellectual disabilities. The instruments typically used to assess ASD or ASD characteristics are in general not valid for use in people with sensory impairments in addition to intellectual disabilities. For this specific population new instruments are urgently needed. When more information about certain behaviors is required, the reviewed assessment tools can be helpful, only it is important to keep their limitations in mind, to use multiple tools and a multidisciplinary team, and most of all, to take into account a person’s individual characteristics, limitations, and possibilities.

**Acknowledgements**

Thanks to Lisette Westland for her early work on this article and to Antoinette van der Pol and Titia van Raamsdonk for their help at the test library at Royal Dutch Kentalis. Our research was funded by ZonMW (grant number 60-00635-98-0081), Royal Dutch Kentalis, and the Behavioural Science Institute, Radboud University Nijmegen.

**References**


