Agreement Between a Brief Autism Observational Instrument and Established ASD Measures

Samantha L. Ward,[1,2] Karen A. Sullivan[1,3] and Linda Gilmore[4,5]

[1]*School of Psychology and Counselling, Queensland University of Technology, Australia*
[2]*Private Practice, Stepping Stones for Life Psychology QLD Pty Ltd Brisbane, Queensland, Australia*
[3]*Institute of Health and Biomedical Innovation, Queensland University of Technology, Australia*
[4]*School of Cultural and Professional Learning, Queensland University of Technology, Australia*
[5]*Children and Youth Research Centre, Queensland University of Technology, Australia*

Objective: Limited time and resources necessitate the availability of accurate, inexpensive and rapid diagnostic aids for Autism Spectrum Disorder (ASD). The Autistic Behavioural Indicators Instrument (ABII) was developed for this purpose, but its psychometric properties have not yet been fully established. **Method:** The clinician-rated ABII, the Autism Diagnostic Observation Schedule (ADOS), the Childhood Autism Rating Scale – Second Edition, Standard Version (CARS2-ST), and Diagnostic and Statistical Manual of Mental Disorders (*DSM-5*) diagnostic criteria were individually administered to children with an independent paediatrician *DSM-IV-TR* or *DSM-5* autism spectrum diagnosis, aged 2-6 years ($n = 51$, $M_{child\ age} = 3.6$ years). The agreement between each of the measures on autism diagnostic classification was calculated and compared, and the intercorrelation between the instruments examined. **Results:** There was significant moderate agreement for the classification of autism between the ABII and the *DSM-5*, and significant fair agreement between the ABII and ADOS and ABII and CARS2-ST. True positive diagnostic classifications were similar across the ABII ($n = 47$, 92.2%) and ADOS ($n = 45$, 88.2%), and significantly higher than the CARS2-ST ($n = 30$, 58.8%). The ABII total scale score was strongly positively correlated with both the ADOS and CARS2-ST total scores. **Conclusion**: The ABII's test characteristics were comparable to those of established measures, and the intercorrelations between selected measures support its convergent validity. The ABII could be added to the clinician's toolbox as a screening test.

■ **Keywords:** autism spectrum disorder, autism screening, ABII, ADOS, CARS2-ST, *DSM-5*

Although the detection and diagnosis of Autism Spectrum Disorder (ASD) is complex, it is generally agreed that the defining core features of ASD emerge within the first two to three years of life for most children and can be detected by trained professionals following comprehensive diagnostic assessment at two years of age (Centers for Disease Control and Prevention [CDC], 2014; Guthrie, Swineford, Nottke & Wetherby, 2014).

Received 10 December 2015; Accepted 29 March 2016; First published online 20 May 2016

Address for Correspondence: Samantha Ward, School of Psychology and Counselling, Queensland University of Technology, Australia. Email: samantha@steppingstonesqld.com.

Volume 33 | Issue 2 | 2016 | pp. 127–138 | © Australian Psychological Society Ltd 2016 | doi: 10.1017/edp.2016.1
https://doi.org/10.1017/edp.2016.1 Published online by Cambridge University Press

Yet, a large proportion of children are not seen for specialist diagnostic evaluation for ASD until after the age of three (Shattuck et al., 2009; Ward, Sullivan & Gilmore, in press). Optimising the timely and accurate referral of children with ASD for specialist evaluation is therefore critical to facilitate diagnosis. General practitioners play a pivotal role in the identification of children in need of referral for specialist ASD evaluation (Carbone, Farley & Davis, 2010). They are often the initial respondents to parental concerns (Fillipek, 1999) and can therefore initiate early referral of children for specialist ASD assessment. However, during brief healthcare visits it is difficult to elicit ASD specific information from parents (Zuckerman, Lindy & Sinche, 2015) or quantify the symptoms to assess risk (Gabrielsen et al., 2015). This can reduce the accuracy of referrals for specialist review, which itself could have several negative flow on effects, including delayed diagnosis (Reichow, Barton, Boyd & Hume, 2012). Allied health clinicians, such as psychologists, also play a pivotal role in the multidisciplinary assessment of children with ASD (Ward et al., in press). They are often involved in completing comprehensive diagnostic assessment, which can be an expensive and timely process (Charman & Gotham, 2014). Brief and inexpensive methods of quantifying ASD risk could help to inform psychologists in their decisions regarding initiation of in-depth evaluation.

The use of ASD specific screening instruments in primary health and allied health care settings can contribute to optimising timely and accurate referral of children with ASD (Charak & Stella, 2001–2002). However, existing instruments have yielded high misclassification errors and produced lower than desirable sensitivity and specificity values, limiting their utility in clinical settings (Al-Qabandi, Gorter & Rosenbaum, 2011; Barbaro & Dissanyake, 2012; Osterling et al., 2010). The accurate differential detection of children with ASD from children with other developmental disorders and delays is important to guide referrals, to reduce costly diagnostic assessment, and to reduce undue parental distress associated with false positive screens on instruments (Bölte et al, 2013; Lipkin & Hyman, 2011). One of the most common non-specific "red flags" of ASD is speech and language delay (Johnson & Myers, 2007). Children with speech and language delay are over classified as having ASD on a number of existing screening instruments (Whitehouse, Barry & Bishop, 2007). The inclusion of non-specific symptomology, such as speech and language delay, to ascertain ASD risk, complicates and delays the detection of ASD. It can produce over inclusive referrals of children and bases detection on the absence of a skill that is not evident until later on in the child's development (Downey et al., 2002).

The detection of ASD based on primary unique ASD indicators, that is, those indicators that have an early emergence and that are specific to ASD (Clifford, Young & Williamson, 2007) improves referral accuracy (Bölte et al., 2013). Formal diagnostic evaluations for ASD are associated with high demands in terms of cost and time for already constrained systems, and possible increased stress for families (Horlin et al., 2014; Shattuck & Grosse, 2007). With early detection of children deemed to be of increasing importance to facilitate earlier diagnosis (Bölte et al., 2013), there is scope for continued development of screening tools to improve the accuracy of referrals. To this end, a number of parent questionnaires have been developed to help to identify children in need of further ASD diagnostic evaluation. While parent interview serves as an important informant source in the symptomatic description and quantification of ASD behaviours (Sacrey et al., 2015), direct behavioural observation is also important, particularly when guided by standardised instruments specifically designed to elicit

https://doi.org/10.1017/edp.2016.1 Published online by Cambridge University Press

key ASD behaviours (American Academy of Pediatrics [AAP], 2006; Charak & Stella, 2001–2002; Hampton & Strand, 2015; Norris & Lecavalier, 2010).

A recently developed semi-structured brief observational instrument, the Autistic Behavioural Indicators Instrument (ABII; Ward & Gilmore, 2010), has shown preliminary utility in classifying children aged two to six years with and without Autistic Disorder (AD). In an initial examination of instrument classification efficiency, the ABII correctly classified all children with AD ($n = 20$) and produced no misclassification errors of children with typical development ($n = 20$) or speech and language impairment ($n = 20$). The ABII is a unique instrument in that it is currently the only brief non-verbal observational instrument designed to quantify ASD risk based on the increased presence, rather than absence, of unique indicators of ASD. Conceptually, this feature may help to reduce misclassification errors (Wetherby et al., 2004). In addition, the ABII has previously shown very good discrimination of children with speech and language impairment. It does not require children to have developed the use of language or to understand spoken language and is designed to elicit key unique indicators of ASD for rapid quantification to ascertain ASD risk. The ABII was designed to optimise referral accuracy and is not intended to replace "gold standard' diagnostic evaluation procedures that involve comprehensive assessment methods.

The current study sought to further establish the psychometric properties of the ABII by exploring 1) the classification accuracy of children across the full range of diagnostic subcategories as operationalised in the Diagnostic and Statistical Manual of Mental Disorders, fourth edition, text revision (*DSM-IV-TR*, APA, 2000), including Autistic Disorder (AD), Asperger's Syndrome (AS), and Pervasive Developmental Disorder-Not Otherwise Specified (PDD-NOS); 2) the classification accuracy of children across the full range of current diagnostic subcategories as operationalised in the Diagnostic and Statistical Manual of Mental Disorders, fifth edition (*DSM-5*, APA, 2013), including Autistic Spectrum Disorder (ASD), severity level 1, severity level 2 and severity level 3; and 3) examining agreement on *DSM-IV-TR* and *DSM-5* autism diagnostic classification between the ABII and existing autism instruments. These measures included the Autism Diagnostic Observation Schedule (ADOS; Lord, Rutter, DiLavore, & Risi, 1999), a "gold standard" semi-structured observation instrument, and the Childhood Autism Rating Scale – Second Edition, Standard Version (CARS2-ST; Schopler, Van Bourgondien, Wellman & Love, 2010), a clinician observation instrument of unstructured activities. The ADOS and the CARS2-ST were specifically chosen for comparative analysis as both are standardised instruments requiring clinician evaluation. Similar to the ABII, the ADOS and CARS2-ST provide a method of quantifying direct clinician observations of ASD behaviours, thus enabling the comparison of the ABII with similar clinician observation tools.

Method

Participants

Participants included children with an independent paediatrician verified diagnosis of autism using the criteria outlined in the *DSM-IV-TR* and *DMS-5*, aged between 2 and 6 years ($n = 51$, $M_{child\ age} = 3.6$ years). A set of revised criteria and a change to the diagnostic subcategory were released in the *DSM-5*. Children who were diagnosed

https://doi.org/10.1017/edp.2016.1 Published online by Cambridge University Press

prior to the release of the revised criteria ($n = 24$, 47.1%) had been independently diagnosed using the *DSM-IV-TR* diagnostic criteria (Autistic Disorder [AD], $n = 13$ (21%); Asperger's Syndrome [AS], $n = 7$, 13.7%; Pervasive Developmental Disorder – Not otherwise specified [PDD = NOS], $n = 4$, 7.8%). Children who were independently diagnosed after the release of the *DSM-5* ($n = 27$, 52.9%) had been diagnosed using the revised diagnostic criteria (Autism Spectrum Disorder, severity level 1, $n = 11$ (21.6%); severity level 2, $n = 7$ (13.7%); severity level 3, $n = 9$ (17.6%). There was no significant difference in the distribution of children based on level of symptom severity, as outlined in the *DSM-IV-TR* and *DSM-5* diagnostic subcategories, with both more severe (*DSM-IV-TR* diagnosis of AD: $n = 13$, 21% or *DSM-5* diagnosis of ASD, severity level 3, $n = 9$, 17.6%), and less severe (*DSM-IV-TR* diagnosis of AS, $n = 7$, 13.7%, PDD = NOS, $n = 4$, or a *DSM-5* diagnosis of ASD, severity level 1, $n = 11$, 21.6% or severity level 2, $n = 7$, 13.7%) presentations represented in the sample, $\chi^2(1) = .961$, $p = .327$. Children with more severe presentations were significantly younger than those children with less severe presentations, $t(49) = -5.35$, $p = <.000$, CI [-1.68, -.76]. $\eta^2 = .37$. The final cohort was predominately male (male, n=40, female, $n = 11$), $\chi^2(1) = 16.49$, $p = <.001$; however, there was no significant difference in gender distribution between more severe presentations (male = 14, female = 6) and less severe presentations (male = 24, female = 5), *fisher exact*, $p = .50$, $\varphi_c = .12$.

Measures

The Autistic Behavioural Indicators Instrument (ABII; Ward & Gilmore, 2010). The ABII is intended to be administered by registered general practitioners and allied health clinicians as a level 1 screening instrument to identify children in need of specialist diagnostic evaluation for ASD (National Initiative for Autism: Screening and Assessment [NIASA], 2003). The ABII assesses the presence of behavioural markers with an early emergence within key ASD domains. The ABII is an 18 item non-verbal screening instrument that includes a fixed sequence of standardised and structured tasks that elicit specific target behaviours across social attention, sensory and behavioural domains (see Ward and Gilmore, 2010 for a review of instrument development). The social attention subscale (SAS) comprises tasks that measure social orienting (e.g., preferential gaze to social or non social stimuli), and joint attention behaviours (e.g., preferences for shared engagement with a caregiver or solitary play), and displays of affect across social and non-social stimuli. The sensory subscale (SS) comprises tasks that measure visual, tactile, and oral sensory seeking behaviours (e.g., duration of time engaged in sensory exploration), and the presence of hypo- or hyper—responsiveness. The behavioural subscale (BS) comprises naturalistic observations of children when demands or denials are placed on the child (e.g. frequency and duration of behavioural protests). On all of the ABII items, a score of 0 represents typical behavioural responses and a score of 1 represents the presence of autistic behavioural indicators. Scores are added within each domain to calculate a subscale score and the aggregate of subscales is calculated to provide a total ABII scale score. Higher subscale and total ABII scores represent a greater presence of autistic behavioural indicators. A total ABII score of 11 or above is indicative of a "positive" screen. In an initial examination of the utility of the ABII in classifying children aged 2 to 6 years with a *DSM-IV-TR* diagnosis of Autistic Disorder, a speech and language impairment, and neurotypcial development (Ward & Gilmore, 2010), a cut-off score of 11 produced maximum sensitivity and

https://doi.org/10.1017/edp.2016.1 Published online by Cambridge University Press

specificity. Administration of the structured tasks on the ABII generally takes five to ten minutes, depending on the level of engagement and ability of the child.

The Autism Diagnostic Observation Schedule (ADOS; Lord et al., 1999). The ADOS is a standardised semi-structured play based observational tool considered a "gold standard" instrument to guide diagnosis of ASD in children 2+ years of age. The instrument has a series of structured and semi-structured presses for interaction, accompanied by coding of specific target behaviours and general ratings of the quality of behaviours. The ADOS consists of four modules, one of which is selected for administration based upon the child's expressive language ability. Modules 1 to 3 were used in the current study (Module 1: $n = 13$; Module 2: $n = 13$; Module 3: $n = 36$). The ADOS algorithm provides diagnostic cut-offs for autistic disorder, ASD, and non-ASD. Higher scores on the ADOS are indicative of greater impairment in core areas of deficit. The ADOS requires specialist training and takes approximately 40–60 minutes to administer. The measure has demonstrated strong psychometric properties (Lord et al., 2008). A revised version of the ADOS was released in 2012 (ADOS-2; Lord et al. 2012), but was not available when the data were collected for the current study.

The Childhood Autism Rating Scale – Second Edition, Standard Version (CARS2 –ST; Schopler et al., 2010). The CARS2-ST is a behavioural rating scale used to identify the presence and severity of symptoms of ASD in children 2+ years of age and to distinguish them from children with other developmental disorders. Items on the CARS2-ST are drawn from five prominent systems for diagnosing ASD. The CARS2-ST is completed after direct observation of the child during unstructured activities. Ratings are based on a 4-point response scale that measures the intensity, abnormality and duration of behaviours. Based on combined ratings from the 15 items, the child's score can be classified as indicative of mild, moderate, or severe autism, or no autism. Higher CARS2-ST scores indicate more severe forms of ASD. The CARS2-ST takes approximately 5 to 10 minutes to administer and can be used by a range of professionals (e.g. physicians, special educators and psychologists). The CARS2-ST has demonstrated sound psychometric properties (Schopler et al., 2010).

The Diagnostic and Statistical Manual of Mental Disorders, fifth edition, (DSM-5) diagnostic criteria for ASD. The Diagnostic and Statistical Manual of Mental Disorders, fifth edition, (*DSM-5*; APA, 2013), provides the diagnostic criteria used by clinicians to diagnose ASD. Revised criteria for ASD were released in the fifth edition. A single broader diagnostic category of autism spectrum disorder has replaced previously sub-categorised labels of AS, AD, and PDD-NOS, as outlined in the previous edition (*DSM-IV-TR*; APA, 2000). To meet diagnostic classification for ASD requires impairment and persistent deficits across criteria in two domains, namely social communication and restricted, repetitive patterns of behaviour, interests or activities. A specifier of severity can be applied ranging from 'level 1: requiring support' to 'level 3: requiring substantial support'.

Procedure

Parents were made aware of the study through a research flyer distributed at medical appointments, allied health intervention sessions, special education schools, and

https://doi.org/10.1017/edp.2016.1 Published online by Cambridge University Press

TABLE 1
Means, standard deviations and t-test statistics based on *DSM-IV-TR* or *DSM-5* symptom severity

	Group 1		Group 2					
	M	*SD*	*M*	*SD*	*df*	*t*	*CI [LL,UL]***	η^2
ABII	18.08	2.69	14.19	3.92	49	4.11*	1.99, 5.79	.26
CARS2-ST	40.70	7.67	27.06	3.86	49	7.98*	10.17, 17.11	.56
ADOS	24.40	4.87	15.85	4.38	49	6.60	5.95, 11.16	.47

Group 1, n = 25 (Autistic Disorder, n = 13, and Autism Spectrum Disorder severity level 3 n = 9). Group 2, n = 26 (Asperger's Syndrome, n = 7, Pervasive Developmental Disorder – Not otherwise specified, n = 4, Autism Spectrum Disorder, Severity level 1, n = 11, and severity level 2, n = 7).
* *p = <.001*
** 95% Confidence Interval (CI) Lower Limit and Upper Limit [LL,UL]

autism organisations. All parents provided informed consent and ethical approval was obtained from Human Research Ethics Committee of Queensland University of Technology to conducting the study. All children commenced testing sessions by engaging in an initial unstructured play session of approximately five minutes duration. During this time the child was permitted to play freely with toys either alone, with the primary caregiver, or with the examiner. The administration order of the ABII and the ADOS was randomly assigned and counterbalanced. After these tests were administered the examiner completed the CARS2-ST and the *DSM-5* diagnostic criteria for ASD. All measures were completed in the same test session for each participant and administered by a registered psychologist trained in the administration and scoring of each instrument. Total testing time varied from 45 minutes to 90 minutes depending on the child's level of engagement and ability.

Results

Correlations and group differences for the total instrument scores

The ABII total scale score was strongly correlated with both the ADOS ($r = 0.56$, $p = <.01$) and CARS2-ST ($r = 0.71$, $p = <.01$). To investigate any differences in instrument total scores based on ASD symptom severity, the sample was divided into two groups. Group 1 comprised children with more severe presentations (*DSM-IV-TR* diagnosis of AD and *DSM-5* diagnosis of ASD level 3). Group 2 included children with less severe presentations (*DSM-IV-TR* diagnosis of AS and PDD-NOS and *DSM-5* diagnosis of ASD level 1 and level 2). There were significant differences between the two groups in total instrument score with a large effect. Group 1 had significantly higher total scale scores on each of the instruments (see table 1 for means, standard deviations and *t*-test statistics).

Given that child age can influence diagnosis and symptom emergence, the pooled data were split into two age groups and the effect of age on these variables was examined. Group 1 was younger than 4 years of age ($n = 29$, $M_{child\ age} = 2.85$ years, $SD = .40$) and group 2 were 4 years or older $n = 22$, $M_{child\ age} = 4.61$ years, $SD = .59$). There was no significant age group difference in the proportion of children who were correctly classified on either the ABII, *fisher exact*, $p = .303$, $\varphi_c = .19$, or ADOS, *fisher exact*, $p = .073$, $\varphi_c = .28$. Older children were significantly more likely to be classified on the CARS2-ST, *fisher exact*, $p = <.001$, $\varphi_c = .64$.

https://doi.org/10.1017/edp.2016.1 Published online by Cambridge University Press

TABLE 2
True positive (TP) and false negative (FN) ASD classification of children and z statistic

	ABII and ADOS				
	TP		FP		
	ABII	ADOS	ABII	ADOS	*z*
All ASD	47 (92.2)	45 (88.2)	4 (7.8)	6 (11.8)	0.68
AD	13 (100)	13 (100)	-	-	0.00
AS	5 (71.4)	4 (57.1)	2 (28.6)	3 (42.9)	1.51
PDD-NOS	3 (75)	3 (75)	1(25)	1(25)	0.00
ASD 1	11 (100)	9 (81.8)	-	2 (18.2)	3.20*
ASD 2	6 (85.7)	7 (100)	1 (14.3)	-	-2.80*
ASD 3	9 (100)	9 (100)	-	-	0.00
	ABII and CARS2-ST				
	TP		FP		
	ABII	CARS2-ST	ABII	CARS2-ST	*z*
All ASD	47 (92.2)	30 (58.8)	4 (7.8)	21 (41.2)	3.92*
AD	13 (100)	12 (92.3)	-	1 (7.7)	2.02*
AS	5 (71.4)	2 (28.6)	2 (28.6)	5 (71.4)	4.32*
PDD-NOS	3 (75)	2 (50)	1(25)	2 (50)	2.61*
ASD 1	11 (100)	2 (18.2)	-	9 (81.8)	8.40*
ASD 2	6 (85.7)	3 (42.9)	1 (14.3)	4 (57.1)	4.51*
ASD 3	9 (100)	9 (100)	-	-	0.00

All ASD, n = 51 (DSM-IV-TR diagnosis of Autistic Disorder (AD), n = 13, Asperger's Syndrome (AS), n = 7, Pervasive Developmental Disorder – Not otherwise specified (PDD-NOS), n = 4, DSM-5 diagnosis of Autism Spectrum Disorder (ASD) severity level 1 (ASD 1), n = 11, severity level 2 (ASD 2), n = 7, severity level 3 (ASD 3), n = 9.
*$p = <.05$

Classification efficiency of the ABII, ADOS and CARS2-ST

There were no significant differences in the proportion of children diagnosed with more or less severe presentations who screened positive on the ABII, *Fisher exact*, $p = .110$, $\varphi_c = .29$. However, children with less severe presentations were significantly more likely to screen a false positive on the ADOS, *Fisher exact*, $p = .023$, $\varphi_c = .36$ and the CARS2-ST, *Fisher exact*, $p = <.001$, $\varphi_c = .741$. Table 2 presents the true positive (TP) and false negative rates (FN) and *z* statistics for the classification rates of the ABII, ADOS and CARS2-ST by the whole sample and by *DSM-IV-TR* and *DSM-5* diagnostic subcategory. For the classification of all children with a *DSM-IV-TR* and *DSM-5* diagnosis, the ABII (TP: $n = 47$, 92.2%) demonstrated similar diagnostic classification efficiency with no significant differences when compared to the ADOS (TP: $n = 45$, 88.2%). Compared to the CARS2-ST (TP: $n = 30$, 58.8%), the ABII correctly classified significantly more children as having ASD, $z = 3.92$, $p = <.05$.

Correct classification rates were similar across the ABII and ADOS for children diagnosed with AD, AS, PDD-NOS, and ASD severity level 3. The ABII correctly classified more children with ASD level 1 compared to the ADOS, $z = 3.20$, $p = <.05$, however, classified fewer children with ASD severity level 2 compared to the ADOS $z = -2.80$, $p = <.05$. Compared to the CARS2-ST, the ABII correctly classified more children with all *DSM-IV-TR* and *DSM-5* diagnoses, with the exception of children with a *DSM-5* ASD diagnosis severity level 3, where there were no significant differences. For the cohort, there was a significant moderate agreement on ASD classification

https://doi.org/10.1017/edp.2016.1 Published online by Cambridge University Press

between the ABII and *DSM-5* diagnostic criteria, $\kappa = 0.50$, $p = .001$, and significant though fair agreement between the ABII and ADOS, $\kappa = 0.34$, $p = .013$, and ABII and CARS2-ST, $\kappa = 0.22$, $p = .013$.

Discussion

Current rates of referral for specialist diagnostic evaluation are occurring later than desirable, delaying formal ASD diagnosis for many children (CDC, 2014). ASD specific screening instruments could improve ASD detection during primary health care visits, guiding physician referral for specialist assessment. ASD screening instruments can also contribute to informing decisions from allied health clinicians whether to proceed with costly diagnostic instruments. Given that diagnostic evaluation procedures are time consuming and expensive, the accuracy of diagnostic evaluation referrals and decisions to initiate comprehensive assessment is important. To increase clinical utility, instruments need to strike a balance between detection accuracy, and time and cost efficiency. The ABII was designed to address this requirement, providing an inexpensive screening instrument to elicit ASD behaviours for rapid quantification. The ABII is not designed to make a diagnosis or to replace specialist clinician diagnostic consultation. Rather, the ABII is designed to inform the accuracy of referrals to facilitate earlier detection of children in need of further specialist diagnostic evaluation. This study provides further information on the psychometric properties of the ABII (Ward & Gilmore, 2010) by examining its classification efficiency of children diagnosed with *DSM-IV-TR* or *DSM-5* diagnostic criteria for autism, establishing its correlation with existing instruments (ADOS and CARS2-ST) and comparing its agreement with diagnostic criteria (*DSM-5*), ADOS and the CARS2-ST in the classification of children with autism.

The ABII yielded acceptable correct classification efficiency, agreeing with an independent paediatrician diagnosis of ASD in close to ninety percent of cases. The ABII total scale score was strongly correlated with both the ADOS and the CARS2-ST, demonstrating good interrelationship between the ABII and these previously established and validated measures on the classification of children with ASD. The ABII was in moderate agreement for the classification of ASD with the *DSM-5* and and in fair agreement with the ADOS and CARS-2-ST. Classification efficiency on the ABII was similar to the ADOS and higher compared to the CARS2-ST. In addition, the accurate detection of children across the autism spectrum was high on the ABII, with both more and less severe presentations of ASD, as determined by *DSM-IV-TR* and *DSM-5* ASD diagnostic subcategory, screening positive on the ABII. In comparison, children with less severe presentations of autism were more likely to be misclassified on both the ADOS and the CARS2-ST. Although the ABII is an empirically derived instrument designed to measure the presence of specific behavioural markers of ASD with an early emergence, results from this examination suggest these specific indicators may be present up to the age of six years. Both younger and older children were correctly classified on the ABII, with no significant differences in classification accuracy based on age. This result warrants further investigation. ASD symptoms unfold over time (Estes et al., 2015), which may influence symptom manifestation and detection, potentially resulting in differences in symptom presentation between younger and older children. Theoretically, this evolving presentation of behavioural markers of ASD may result in the requirement for age specific screening instruments. A screening instrument that is highly sensitive at one age may not be as sensitive at another age, depending on when specific symptoms emerge. In this examination, the

https://doi.org/10.1017/edp.2016.1 Published online by Cambridge University Press

ABII correctly detected a very high proportion of children across all *DSM-IV-TR* and *DSM-5* diagnoses of autism between the ages of 2 years to 6 years, suggesting that it may be a highly sensitive measure for detecting ASD indicators that is not influenced by symptom severity and age.

The difficultly that practitioners report in detecting ASD during brief clinical observations (Gabrielsen et al., 2015), in combination with practical time and resource constraints during standard primary child health care visits, may increase the risk of either misidentification of children in need of further referral for ASD, or over identification of children with non specific developmental concerns and delays, producing inaccurate referrals. Further research is required to investigate how well the ABII can distinguish children with ASD from children with other developmental concerns and disabilities.

With direct clinician observation an important informant source in the measurement of ASD symptoms, and with structured behavioural observational instruments providing a potential source to improve detection of ASD behaviours (Charak & Stella, 2001–2002), continuing efforts are required to develop reliable and valid instruments. Few direct behavioural observation instruments to aid in the detection of children in need of further diagnostic evaluation are currently available for use. The ABII may help to address this resource gap, along with providing an instrument that possesses unique features compared to existing instruments. The ABII is the only brief non-verbal observational instrument designed to quantify ASD "risk" based on the increased presence rather than absence of unique indicators of ASD. These features may reduce misclassification errors.

Limitations and future directions

Although the present study aimed to further establish psychometric properties of the ABII in correctly classifying children across the full range of diagnostic subcategories of autism as operationalised under both *DSM-IV-TR* and *DSM-5* diagnostic criteria, inclusion of children without an ASD diagnosis would have enabled additional exploration of specificity values, along with possible alteration to ASD severity cut-off scores to establish level of ASD severity, in line with the diagnostic criteria and ADOS and CARS2-ST severity cut-off scores. Therefore, results from this investigation should be regarded as exploratory. A future prospective longitudinal investigation would be of benefit to establish sensitivity, specificity, positive predictive validity and negative predictive validity, and to compare the agreement between the ABII and the newly released revised ADOS-2 (Lord et al., 2012).

A further study limitation is that the examiner was not blind to child diagnosis. This may have influenced the observation and rating of ASD indicators. Although children were not assessed for IQ and speech and language skills in this study, unlike the CARS2-ST and ADOS, there was not a significant difference in the classification rates of children diagnosed with more or less severe presentations of autism, as outlined in the *DSM-IV-TR* and *DSM-5* diagnostic subcategories and severity levels. This result may provide preliminary evidence that the ABII is measuring behavioural markers within key ASD domains that are present in children across the full spectrum of autism. While there was no significant difference in classification accuracy of the ABII between older and younger children, future research could further explore and control for any potential effect of age on classification. Finally, although screening children for ASD is of potential benefit to inform the accuracy of referrals, any

https://doi.org/10.1017/edp.2016.1 Published online by Cambridge University Press

screening instrument can carry inherent risk due to misclassification errors. ASD is an heterogeneous disorder with symptoms that can unfold across the early years of a child's life, therefore screening should not be seen as a discrete process; rather, developmental surveillance methods need to be employed with possible screening at multiple stages, a practice that would necessitate inexpensive and rapid instruments with strong psychometric properties. Until such time that an instrument fulfils these requirements, screening for ASD needs to remain a cautionary practice.

Conclusions

This examination sought to expand upon previously reported psychometric properties of the ABII. Results indicate the ABII performs similarly to the ADOS and is superior to the CARS2-ST in the classification of children across the autism spectrum, and it is strongly correlated with both measures. In this study, the ABII was in fair agreement with each of the measures and moderate agreement with the *DSM-5* diagnostic criteria for ASD. Although not intended to replace comprehensive assessments or measures, the ABII could be a useful addition to clinicians' batteries by providing a rapid method of quantifying ASD indicators to inform referral accuracy of children in need of further diagnostic evaluation of ASD.

Acknowledgements

None

Financial Support

This research received no specific grant from any funding agency, commercial, or not-for-profit sectors.

Conflicts of Interest

None

Ethical standards

The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008."

References

Al-Qabandi, M., Gorter, J.W., & Rosenbaum, P. (2011). Early autism detection: are we ready for routine screening? *Pediatrics*, *128*, e211–e217. doi: 10.1542/peds.2010-1881

American Academy of Paediatrics (AAP). (2006). Identifying infants and young children with developmental disorders in the medical home: An algorithm for developmental surveillance and screening. *Pediatrics*, *118*, 405–420. *doi*: 10.1542/peds.2006-1231

American Psychiatric Association (APA). (2000). Diagnostic and statistical manual of mental disorders (4th ed., Text Revision). Washington, DC: Author.

American Psychiatric Association (APA). (2013). Diagnostic and statistical manual of mental disorders (5th ed.). Washington, DC: Author.

https://doi.org/10.1017/edp.2016.1 Published online by Cambridge University Press

Barbaro, J., & Dissanayake, C. (2012). Early markers of Autism Spectrum Disorders in infants and toddlers prospectively identified in the social attention and communication study. *Autism*, *17*, 64–86. doi: 10.1177/1362361312442597

Bölte, S., Marschik, P.B., Falck-Ytter, T., Charman, T., Roeyers, H., Elsabbagh, M. (2013). Infants at risk for autism: A European perspective on current status, challenges and opportunities. *European Child and Adolescent Psychiatry*, *22*, 341–348. doi: 10.1007/s00787-012-0368-4

Carbone, P.S., Farley, M., & Davis, T. (2010). Primary care for children with Autism. *American Family Physician, 15*, 453–460. Retrieved from www.aafp.org

Centers for Disease Control and Prevention (CDC). (2014). Prevalence of Autism Spectrum Disorder among children aged 8 years: Autism and developmental disabilities monitoring network, 11 sites, United States, 2010. *Morbidity and Mortality Weekly Report, 63*, 1–22. Retrieved December 6, 2015 from www.cdc.gov/mmwr/pdf/ss6302.pdf

Charak, D.A., & Stella, J.L. (2001-2002). Screening and diagnostic instruments for identification of Autism Spectrum Disorders in children, adolescents, and young adults: A selective review. *Assessment for Effective Intervention, 27*, 5–17. doi: 10.1177/073724770202700103

Charman, T., & Gotham, K. (2014). Measurement issues: Screening and diagnostic instruments for Autism Spectrum Disorders – lessons from research and practice. *Child and Adolescent Mental Health, 18*, 52–63. doi: 10.1111/j.1475-3588.2012.00664x

Clifford, S., Young, R., & Williamson, P. (2007). Assessing the early characteristics of Autistic Disorder using video analysis. *Journal of Autism and Developmental Disorders, 37*, 301–313. doi: 10.1007/s10803-006-0160-8

Downey, D., Mraz, R., Knott, J., Knutson, C., Holte, L., Van Dyke, D. (2002). Diagnosis and evaluation of children who are not talking. *Infants and Young Children, 15*, 38–48. doi: 10.1097/00001163-200210000-00006

Estes, A., Zwaigenbaum, L., Gu, H., John, T., Paterson, S., Elison, J.T., Hazlett, H., Botteron, K., Dager, S.R., Schultz, R.T., Kostopoulos, P., Evans, A., Dawson, G., Eliason, J., Alvarez, S., Piven, J., & IBIS Netwron, (2015). Behavioural, cogntivie, and adaptive development in infants with Autism Spectrum Disorder in the first 2 years of life. *Journal of Neurodevelopmental Disorder, 7*: 24. doi: 10.1186/s11689-015-9117-6

Filipek, P.A., Accardo, P.J., Baranek, G.T., Cook, E.H., Dawson, G., Gordon, B., Gravel, J.S., Johnson, C.P., Kallen, R.J., Levy, S.E., Minshew, N.J., Prizant, B.M., Rapin, I., Rogers, S.J., Stones, W.L., Teplin, S., Tuchman, R.F., & Volkmar, F.R. (1999). The screening and diagnosis of Autistic Spectrum Disorders. *Journal of Autism and Developmental Disorders, 29*, 439–484. doi: 10.1023/A:1021943802493

Gabrielsen, T.P., Farley, M., Speer, L., Villalobos, M., Baker, C.N., Miller, M. (2015). Identifying autism in a brief observation. *Pediatrics*, *135*, 330–338. doi: 10.1542/peds.2014-1428

Guthrie, W., Swineford, L.B., Nottke, C., & Wetherby, A.M. (2013). Early diagnosis of Autism Spectrum Disorder: Stability and change in clinical diagnosis and symptom presentation. *Journal of Child Psychology and Psychiatry, 54*, 582–590. doi: 10.1111/jcpp.12008

Hampton, J., & Strand, P.S. (2015). A review of level 2 parent-report instruments used to screen children aged 1.5-5 for autism: A meta-analytic update. *Journal of Autism and Developmental Disorders, 45*, 1–12.

Horlin, C., Falkmer, M., Parsons, R., Albrecht, M.A., & Falkmer, T. (2014). The cost of Autism Spectrum Disorders. *PLoS One, 9*, e106552. doi: 10.1371/journal.pone.0106552

Johnson, C.P., & Myers, S.M. (2007). Identification and evaluation of children with autism spectrum disorders. *Pediatrics*, *120*, 1183–1215. doi: 10.1542/peds.2007-2361

Lipkin, P.H., & Hyman, S.L. (2011). Should all children be screened for Autism Spectrum Disorders? Yes: Merging science, policy and practice. *American Family Physician, 15*, 361–378. Retrieved from www.aafp.org

Lord, C., Rutter, M., DiLavore, P., & Risi, S. (1999). *Autism Diagnostic Observation Schedule*. Los Angles, CA: Western Psychological Services.

Lord, C., Rutter, M., DiLavore, P. C., & Risi, S. (2008). *Autism Diagnostic Observation Schedule Manual*. Los Angeles: Western Psychological Services.

Lord, C., Rutter, M., DiLavore, P., & Risi, S., Gotham, K., Bishop, S.L. (2012). *Autism Diagnostic Observation Schedule – second edition*. Los Angles, CA: Western Psychological Services.

https://doi.org/10.1017/edp.2016.1 Published online by Cambridge University Press

National Initiative for Autism: Screening and Assessment (NIASA). (2003). *National Initiative for Autism: Screening and Assessment (NIASA) (2003). National Autism Plan for Children (NAPC): Plan for the identification, assessment, diagnosis and access to early interventions for pre-school and primary school aged children with autism spectrum disorders (ASD)*. Great Britain: The National Autistic Society. Retrieved December 6, 2015, from iier.isciii.es/autism/pdf/aut_napc.pdf

Norris, M., & Lecavalier. (2010). Screening accuracy of level 2 Autism Spectrum Disorder rating scales. A review of selected instruments. *Autism, 14*, 263–283. doi: 10.1177/1362361309348071

Osterling, I.J., Wensing, M., Swinkles, S.H., Jan van der Gaag, R., Visser, J.C., Woudenberg, T., Minderaa, R., Steenhuis, M., & Buitlelaar, J.K. (2010). Advancing early detection of Autism Spectrum Disorder by applying an integrated two-stage screening approach. *Journal of Child Psychology and Psychiatry, 51*, 250–258. doi: 10.1111/j.1469-7610.2009.02150x

Reichow, B., Barton, E.E., Boyd, B.A., Hume, K. (2012). Early intensive behavioral intervention (EIBI) for young children with autism spectrum disorders (ASD). *The Cochrane Database of Systematic Reviews*. doi: 10.1002/14651858.CD009260.pub2.

Sacrey, L.R., Zwaigenbaum, L., Bryson, S., Brian, J., Smith, I.M., Roberts, W., Szatmari, P., Roncadin, C., Garon, N., Novak, C., Vaillancourt, T., McCormick, T., MacKinnon, B., Jilderda, S., & Armstrong, V. (2015). Can parents' concerns predict Autism Spectrum Disorder? A prospective study of high-risk siblings from 6 to 36 months of age. *Journal of the American Academy of Child & Adolescent Psychiatry, 54*, 470–478. doi: 10.1016/j.jaac.2015.03.014

Schopler, E., Van Bourgondien, M.E., Wellman, G.J., & Love, S.R. (2010). *The Childhood Autism Rating Scale 2nd Edition (C.A.R.S. 2)*. Los Angeles: Western Psychological Services.

Shattuck, P.T., Durkin, M, Maenner, M., Newschaffer, C., Mandell, D.S., Wiggins, L., Lee, L.C., Rice, C., Giarelli, E., Kirby, R., Baio, J., Pimyo-Martin, J., Cuniff, C. (2009). Timing of identification among children with an Autism Spectrum Disorder: Findings from a population-based surveillance study. *Journal of the American Academy of Child and Adolescent Psychiatry, 48*, 474–483. doi: 10.1097/CHI.0b013e31819b3848

Shattuck, P.T., & Grosse, S.D (2007). Issues related to the diagnosis and treatment of autism spectrum disorders. *Mental Retardation and Developmental Disabilities Research Reviews, 13*, 129–135. doi: 10.1002/mrdd.20143

Ward, S.L., Sullivan, K.A., & Gilmore, L. (in press). Practitioner perceptions of the assessment and diagnosis of Autism in Australia [Special issue]. *Australian Psychologist*.

Ward, S.L., & Gilmore, L. (2010). The Autistic Behavioural Indicators Instrument (ABII): development and instrument utility in discriminating Autistic Disorder from speech and language impairment and typical development. *Research in Autism Spectrum Disorders, 4*, 28–42. doi: 10.1016/j.rasd.2009.07.003

Wetherby, A.M., Woods, J., Allen, L., Cleary, J., Dickinson, H., & Lord, C. (2004). Early indicators of Autism Spectrum Disorders in the second year of life. *Journal of Autism and Developmental Disorders, 34*, 473–493. doi: 10.1007/s10803-004-2544-y

Whitehouse, A.J.O., Barry, J.G., & Bishop, D.V.M. (2007). The broader language phenotype of autism: A comparison with specific language impairment. *Journal of Child Psychology and Psychiatry, 48*, 822–830. doi: 10.1111/j.1469-7610.2007.01765.x

Zuckerman, K.E., Lindly, O.J., & Sinche, B.K. (2015). Parental Concerns, provider response, and timeliness of Autism Spectrum Disorder diagnosis. *The Journal of Pediatrics, 166*, 1431–1439. doi: 10.1016/j.jpeds.2015.03.007

https://doi.org/10.1017/edp.2016.1 Published online by Cambridge University Press