Research in Autism Spectrum Disorders 62 (2019) 1–9

Contents lists available at ScienceDirect

Research in Autism Spectrum Disorders

journal homepage: www.elsevier.com/locate/rasd

Psychometric characteristics of the aberrant behavior checklist in a well-defined sample of youth with autism Spectrum disorder

Megan Norris[a,*], Michael G. Aman[b], Micah Osborne Mazurek[c], Jessica F. Scherr[a], Eric M. Butter[a]

[a] Nationwide Children's Hospital, Columbus, OH, United States
[b] The Ohio State University, Nisonger Center, Columbus, OH, United States
[c] University of Missouri, Thompson Center for Autism and Neurodevelopmental Disorders, Columbia, MO, United States

ARTICLE INFO

Keywords:
Aberrant behavior checklist
Autism spectrum disorder
Factor analysis
Validity

ABSTRACT

Background: Behavior and emotional difficulties often occur in children and adolescents with Autism Spectrum Disorder (ASD), yet there are few instruments available to assess such problems in this population. The Aberrant Behavior Checklist (ABC), one option for this, is widely used and has substantial psychometric support. Despite this, only two studies to date have examined its structural validity in samples of individuals diagnosed exclusively with Autism Spectrum Disorder (ASD). This study sought to further examine the ABC's validity for use with children and adolescents with ASD.

Method: Data from 470 individuals aged 2–14 years were submitted to confirmatory factor analysis (CFA). Correlations with other measures were examined. MANOVA was conducted to examine effects of subject characteristics on subscale scores of the ABC and assist in developing norms.

Results: Results supported the original factor structure of the ABC. Convergent and divergent validity analyses indicated correspondence with analogous measures and lack of relationship for dissimilar constructs. Results of normative analyses were very comparable to the only previously published norms for youth with ASD.

Conclusions: Overall, results provide further psychometric support for use of the ABC in individuals with ASD and better understanding the ABC's clinical usefulness, particularly given the similarity between current and previous normative data. In a broader context, these results lend further support to the ABC as an instrument with wide applications.

1. Introduction

Assessment of behavioral and emotional difficulties in individuals with Autism Spectrum Disorder (ASD) is an important area of study due to the increased risk for such difficulties in this population. Recent data suggest a prevalence of ASD of 6 per 1000 children, 30% of whom function within the range of intellectual disability (ID) (Elsabbagh et al., 2012). The prevalence of behavioral and emotional difficulties in those with ASD is much higher than in individuals without an ASD diagnosis (e.g., Buck et al., 2014; Leyfer et al., 2006; Siminoff et al., 2008). While many instruments exist to assess for ASD symptoms (e.g., the Modified-Checklist for Autism

* Corresponding author at: Child Development Center, Nationwide Children's Hospital, 187 West Schrock Rd, Westerville, OH, 43081, United States.
E-mail address: Megan.norris@nationwidechildrens.org (M. Norris).

https://doi.org/10.1016/j.rasd.2019.02.001
Received 23 April 2018; Received in revised form 7 February 2019; Accepted 12 February 2019
Available online 26 February 2019
1750-9467/

in Toddlers, the Social Responsiveness Scale) relatively few instruments have been developed to assess behavioral and emotional difficulties in those with ASD. In the absence of such instruments, researchers and clinicians often use instruments developed for assessing emotional and behavior problems in people with other developmental disabilities. One such instrument that has demonstrated promise is the Aberrant Behavior Checklist (ABC; Aman & Singh, 2017).

The ABC is a widely used instrument originally designed to assess treatment effects in individuals with developmental disabilities. It has been used in hundreds of studies ranging from medication efficacy research to behavioral phenotyping of specific genetic conditions, such as Fragile X, Prader-Willi, Cri-du-chat, Angleman's syndrome, and Smith Megenis Syndrome (e.g., Aman & Singh, 2017; Baumgardner, Reiss, Freund, & Abrams, 1995; Clarke, Boer, Chung, Sturmey, & Webb, 1996; Clarke & Dykens, 1997; Summers & Feldman, 1999). It has been translated from English into at least 35 other languages (Aman & Singh, 2017). It was originally developed for people living in residential facilities, but has been adapted for use in the community (Brown, Aman, & Havercamp, 2002). It is well-studied and has demonstrated strong psychometric properties across a range of samples, including adults, adolescents, and children with developmental disabilities (e.g., Hill, Powlitch, & Furniss, 2008; Marshburn & Aman, 1992). For instance, in a meta-analysis of 14 replication studies, Aman and Singh (2017) found that 85% of ABC items, across studies, loaded on the originally reported factors (Aman & Singh, 1985). However, to date only two studies (Brinkley et al., 2007; Kaat, Lecavalier, & Aman, 2014) have examined the ABC's psychometric properties within samples comprising ASD alone. This is a valuable area of inquiry given the increasing prevalence rates of ASD, high risk of behavioral and emotional problems in those with ASD, and the stability of ABC psychometric properties across other populations.

One of the studies examining the validity of the ABC exclusively in a sample of those with ASD conducted exploratory and confirmatory factor analysis (EFA and CFA, respectively) with 275 individuals aged three to 21 years (Brinkley et al., 2007). Results overall suggested the ABC factor structure was robust in this sample, though some discrepancies from the original factor structure were noted. In EFA, both four- and five- factor solutions emerged. In the five-factor solution, the majority of items (over 70%) loaded onto original factors and 76% of the variability was accounted for by the solution. The Irritability factor did not emerge, and instead many items from the original Irritability factor loaded with items on the original Hyperactivity/Noncompliance subscale, creating a "general disruptive behavior" factor. Additionally, a few items formed a new factor representing self-injury, a finding which has sometimes occurred in other samples, such as those with Fragile X (Sansone et al., 2012). CFA indices of fit for this model suggested moderate fit. The four-factor solution accounted for 71% of the variance in the data. This solution was compared to the four-factor solution obtained with teacher-completed data in a sample of special education students (Marshburn & Aman, 1992) and 78% of the items were found to load onto the anticipated factor. Again an Irritability factor failed to emerge and a Self-injury factor (with three items) did emerge.

More recently, Kaat et al. (2014) submitted ABC data from 1893 youth with ASD aged two to 18 years to EFA and CFA; the large overall sample size permitted for both a calibration and a validation sample. Overall, Kaat et al. found that the original factor structure performed well in exploratory and confirmatory analyses, and convergent and divergent validity was supported. In the five-factor solution which emerged from EFA, 90% of the items loaded onto the original factors. Five different models (including both the 4- and 5-factor models reported by Brinkley et al., 2007 and others) were submitted to CFA with marginal-to-poor fits indicated for all. Possible effects of participant characteristics on subscale scores were examined. Age, sex, and IQ were found to be largely unrelated to scores, though to a mild extent increased age was associated with decreased hyperactivity and irritability, and lower IQ was associated with increased stereotypic behavior. Autism symptom severity (as measured by the ADOS comparison score) was generally unrelated to the ABC subscale scores, and scores of the Child Behavior Checklist (CBCL; Achenbach & Rescorla, 2000) were generally correlated with analogous ABC subscales, as anticipated. Normative data also were presented based on age group and IQ, as analyses did not support gender-related differences; these norms are the only published set for individuals with ASD to date.

Of the two previous studies examining the ABC in ASD-specific samples, Kaat et al. drew participants from a large registry (the Autism Treatment Network) which permitted for a large sample size. The other study conducted by Brinkley et al. had a well-characterized though smaller sample, which did not permit for splitting participants into separate subsamples for the EFA and CFA; this can be problematic as it may capitalize on sample-specific error. Both studies provided support for use of the ABC in ASD samples, but there also were differences in results across these studies which call for further investigation. This study aimed to provide such examination, an important activity given relative lack of instruments validated for measuring behavioral and emotional problems in those with ASD (Hanratty et al., 2015). The aims of this study were to: (1) examine the ABC factor structure in a well-characterized sample of youth with ASD; (2) examine convergent and divergent validity of the ABC; and (3) provide normative data for the ABC as a basis for comparison to the only set of norms in this population that has been described.

2. Method

2.1. Participants

The sample included 470 youth ranging in age from 2 through 14 years inclusive, and their guardians/caregivers. Additional inclusion criteria were that children have a clinical diagnosis of ASD (confirmed by an ADOS-2 score above the Autism Spectrum cut-off, plus meeting diagnostic criteria by DSM-IV-TR/5 per documentation), and that participants be receiving one of the several types of intervention relevant to the larger study's goals (as noted below). Eight-two percent of the sample was male (n = 386) and the mean age at baseline was 7.36 years (SD = 3.83). Eighty-one percent of the sample identified as White, 14% as Hispanic, 9% as Black/African-American, 9% as Asian, 3% as American Indian/Alaskan Native, 2% as Native Hawaiian or Other Pacific Islander; the total for these exceeds 100% as participants could endorse more than one category.

The vast majority of informants were biological mothers (n = 368, 78%), though biological fathers (n = 53, 11%) also were represented. Other informants included adoptive parents, biological grandparents, 1 foster parent, and 1 great aunt.

2.2. Procedure

This study involved secondary analysis of data from a large, multi-site effort designed to examine the validity of the Autism Impact Measure (AIM; Kanne et al., 2014). Three sites participated: the Thompson Center for Autism in Missouri, Rady Children's Hospital in San Diego, and Nationwide Children's Hospital in Ohio. Institutional Review Boards at all sites approved this NIH-funded study. Informed consent was obtained from all adult informants, with assent obtained from all children over the age of 9 years with the cognitive capacity to assent. Families received modest compensation for participation. Participants were recruited from several different treatment groups relevant to the primary study (e.g., medication management; early intensive behavioral intervention or EIBI). A comprehensive assessment battery was completed at baseline, and a subset of measures was re-administered at 6-week intervals across treatment. For the purpose of the current study, only baseline data were examined.

2.3. Instruments

2.3.1. Aberrant behavior checklist-community (ABC; Aman & Singh, 1985)

The ABC is a 58-item rating scale assessing behavior difficulties in individuals with developmental delays. Items are rated from 0 (not at all a problem) to 3 (the problem is severe in degree). Scores are generated on five subscales: Irritability (15 items), Lethargy/Social Withdrawal (16 items), Stereotypic Behavior (7 items), Hyperactivity/Noncompliance (16 items), and Inappropriate Speech (4 items); an overall score is not generated and is strongly discouraged (Aman & Singh, 2017).

2.3.2. Tests of intellectual ability

Overall intellectual ability (i.e., Full Scale IQ) was assessed using one of two measures (described below), depending on the child's age and/or expressive language level.

The Mullen Scales of Early Learning (Mullen; Mullen, 1995) was administered to 16% of the sample. The Mullen is an individually administered test of general development designed for children aged birth to 5 years, 8 months. It assesses development across domains of gross motor, fine motor, visual reception, receptive language, and expressive language. The Early Learning Composite Standard score was used as a measure of Full Scale IQ.

The Differential Ability Scales, Second Edition (DAS-II; Elliott, 2007) was administered to 80% of the sample. The DAS-2 is an individually-administered standardized test of cognitive abilities, including verbal, nonverbal, and spatial information processing in youth ages 2 years, 6 months to 17 years, 11 months. It assesses verbal and visual working memory, immediate and delayed recall, visual recognition and matching, processing and naming speed, phonological processing, and understanding of basic number concepts. The General Cognitive Ability standard score was used as a measure of Full Scale IQ.

2.3.3. Autism diagnostic observation schedule, second edition (ADOS-2; Lord et al., 2012)

The ADOS-2 is a semi-structured, play- and conversation-based assessment designed to measure symptoms of ASD. It is administered by trained clinicians who lead individuals through a variety of activities designed to elicit specific social-communication and play behaviors. The ADOS-2 consists of 5 different modules depending on the individual's age and language level (from the Toddler module for children between 12–30 months not yet demonstrating phrase speech, to Module 4 for older adolescents/adults with fluent language). Module 4 was not used for this study given the ages of participants. After administration, the ADOS-2 is coded following a detailed rubric and a portion of items are then summed to yield a total score. To make comparison across modules easier, Module 1, 2, and 3 were revised to include a comparison score ranging from 1 to 10, with higher scores indicating a greater severity of ASD symptoms (Gotham, Risi, Pickles, & Lord, 2007).

2.3.4. Vineland adaptive behavior scales, second edition, parent interview form (VABS-II; sparrow et al., 2005)

The VABS-II is a semi-structured interview conducted by a trained clinician with an individual's caregiver. It assesses the individual's adaptive behavior skills in the areas of Communication, Daily Living, Social, and Motor. An overall Adaptive Behavior Composite score also is generated.

2.4. Ohio autism clinical impairment scale severity (OACIS- severity; Butter & Mulick, 2006)

The OACIS is a clinician-completed rating scale which has scores modeled on the Clinical Global Impression (CGI) scale. It consists of 10 items rated on a 7-point scale ranging from 1 (normal) to 7 (among the most severe/symptoms). The OACIS was completed based on direct observations and interactions with the child and information provided via interview from caregivers.

2.4.1. Social responsiveness scale, second edition (SRS-2; Constantino & Gruber, 2012)

The SRS-2 is a parent-completed scale that measures the severity and type of social impairments that are characteristic of autism spectrum conditions in children and adolescents. It is meant for individuals aged 2.5 through 18 years. Sixty-five items are rated on a 4-point scale ranging from 1 (not true) to 4 (almost always true). An overall score is generated as well as five treatment subscales: Social Awareness, Social Cognition, Social Communication, Social Motivation, and Restricted Interests and Repetitive Behavior.

2.4.2. The repetitive behavior scale- revised (RBS-R; Bodfish, Symons, & Lewis, 1999)

The RBS-R is a 44-item questionnaire used to measure the full range of repetitive behaviors in individuals with ASD. It consists of six subscales: Stereotyped Behavior, Self-injurious Behavior, Compulsive Behavior, Routine Behavior, Sameness Behavior, and Restricted Behavior. It asks raters to consider the previous month, and rate items on a 4-point scale ranging from 0 (Behavior does not occur) to 3 (Behavior occurs and is a severe problem).

2.5. Statistical analyses

EFA is employed when there is limited information with which to hypothesize the underlying structure of data, while CFA is used when sufficient data to support factor structure are available. The current study employed CFA due to the large number of previous studies supporting the ABC's factor validity, which supplied several reasonable models for examination. Tested models included: the original model; the four- and five-factor models found by Brinkley et al. (2007); a four-factor model detailed by Brown et al. (2002); the 5-factor model which emerged from EFA conducted by Kaat et al. (2014); and a four-factor model proposed by Marshburn and Aman (1992). Missing data were addressed via pairwise deletion (also called pairwise present analysis); this method considers all available data for each pair of variables when estimating the sample statistics to which the model is fit. Thus, observations with some missing data were used when they could be, and this sample had only a small number of missing observations. CFA was conducted using Mplus version 7.31 with robust weighted least squares estimation (WLSMV) on the polychoric correlation matrix. Several different measures of fit were considered, as is generally recommended when conducting CFA (Brown, 2015) including Root Mean Square Error of Approximation (RMSEA), Closeness of Fit (CFI), and Tucker-Lewis Index (TLI). The RMSEA adjusts for complexity in the model; established guidelines for interpretation indicate values less than 0.05 indicate good fit, values between 0.05 and 0.08 indicate reasonable fit, values between 0.08 and 0.10 indicate marginal fit, and values greater than 0.10 indicate poor fit (Browne & Cudeck, 1992). For the TLI and CFI, values near 0.95 or greater are indicative of good fit (Brown, 2015). Cronbach's alpha also was calculated to assess internal consistency of the optimal model's factors. Per guidelines, $\alpha \geq .90$ are excellent, between 0.80 and 0.90 are good, between 0.70 and 0.80 are acceptable, and $\leq .70$ are questionable to unacceptable.

To assess convergent and divergent validity, Spearman-rank correlations between ABC subscale scores and variables measuring theoretically similar constructs (e.g., ABC Irritability and Aberrant/ Abnormal behavior item of the OACIS) as well as theoretically dissimilar constructs (e.g., VABS domain scores and ABC subscale scores) were calculated; Spearman is preferable to Pearson correlations if normality is violated or if data are ordinal (as the ABC data are). Correlations were classified per guidelines proposed by Cohen (1992): $r < .10$ suggests negligible correlations; small correlations range from $0.10 \leq r < .30$; moderate correlations range from $.30 \leq r < .50$; and $r \geq .50$ suggests large correlations.

As only one set of normative data have been published for the ABC with an ASD sample, similar methods were used in this sample in order to compare across studies. Preliminary analyses revealed that the ABC subscale scores were not normally distributed; therefore, a square root transformation was performed on the data. A MANOVA was conducted to examine possible effects of gender (male/female), age (a preschool group for those aged 2–6 years, a school-aged group for those 7–12 years, and a teen group for 13–14 year olds), and functional level (as measured by either the DAS General Cognitive Ability standard score/GCA or the Mullen Early Learning Composite/ELC) on ABC subscale scores. Scheffe post-hoc comparisons were employed to examine pairwise group differences for significant findings. Cut-points between groups for these analyses were selected in order to maximize the ability to compare current results to the only published set of norms for those with ASD, although this resulted in some cells with a small n (i.e., the subsample of teens is very small, though still included for the above-stated reason).

3. Results

3.1. Factor analysis

Results of CFAs are comparable to previous studies and lend additional support to the ABC's factor validity. Results suggest similar fits for the original 5-factor structure and both the 4- and 5-factor structures reported by Brinkley et al. (fit estimates found in Table 1). The 4-factor models found by Brown et al. and Marshburn et al. had poorer fits than these other models. The CFI and GFI did not indicate acceptable fit for any model. The model which emerged from EFA in the Kaat et al. study, which was rather similar to the

Table 1
Indices of Fit for CFA.

	RMSEA				CFI	TLI
Model	Estimate	95% CI		Prob. ≤ .05		
1 Original	0.076	0.073	0.078	0.0000	0.875	0.870
2 Brinkley	0.076	0.074	0.078	0.0000	0.874	0.868
3 Brinkley	0.077	0.075	0.079	0.0000	0.870	0.865
4 Marshburn	0.083	0.081	0.086	0.0000	0.848	0.841
5 Brown	0.086	0.084	0.088	0.0000	0.839	0.832

Note. CI = Confidence Interval, RMSEA = Root Mean Square Error Approximation; CFI = Closeness of Fit; TLI = Tucker-Lewis Index.

Table 2
Factor Loadings for Original/Retained Model Specification.

Factor 1		Factor 2		Factor 3		Factor 4		Factor 5	
Items		Items		Items		Items		Items	
2	0.937	3	0.443	6	0.864	1	0.810	9	0.641
4	0.695	5	0.813	11	0.828	7	0.749	22	0.888
8	0.850	12	0.526	17	0.959	13	0.717	33	0.762
10	0.814	16	0.840	27	0.676	15	0.770	46	0.854
14	0.716	20	0.701	35	0.884	18	0.885		
19	0.861	23	0.482	45	0.780	21	0.777		
25	0.519	26	0.760	49	0.696	24	0.812		
29	0.802	30	0.904			28	0.696		
34	0.536	32	0.639			31	0.744		
36	0.763	37	0.759			38	0.726		
41	0.829	40	0.837			39	0.710		
47	0.753	42	0.850			44	0.647		
50	0.979	43	0.537			48	0.705		
52	0.909	53	0.592			51	0.645		
57	0.877	55	0.779			54	0.863		
		58	0.742			56	0.764		
Mean loading	.79		.70		.81		.75		.79

original factor structure, was not identified in this sample (i.e., the model is mis-specified) and as such further interpretation was not conducted. Considering current data, the widespread use of the original factor structure, and its retention in the largest study with ASD individuals performed to date, there are multiple reasons to prefer the original factor structure to either of the Brinkley models despite similar fit indices in this sample. Table 2 presents factor loadings and Table 3 presents correlations between factors.

Internal consistency for subscales of the original structure were similar to those found in previous studies and ranged from acceptable to excellent (Irritability = .92, Lethargy/Social Withdrawal = .88, Stereotypic Behavior = .86, Hyperactivity/Noncompliance = .93, Inappropriate Speech = .73).

3.2. Convergent and divergent validity

Results of these analyses are found in Table 4, with expected correlations bolded. Overall, anticipated relationships were found, supporting validity. All ABC subscales except the Inappropriate Speech subscale were negatively associated with all VABS domain scores; Inappropriate Speech was negatively correlated at a significant level with only the VABS Socialization score. Correlations were generally small to medium, on order of .2–.3, which is similar to previous studies examining divergent validity. Correlations between ABC subscales and the SRS total score also were significant, suggesting moderate relationship between difficulties measured by the ABC and parent-reported ASD symptomology. This also held true for clinician-assessed ASD symptoms per the OACIS (except with the ABC Inappropriate Speech scale), though correlations were smaller. Selected ABC and OACIS subscales correlated as anticipated, as did selected ABC and RBS-R subscales; the largest correlation was found between the ABC Stereotypy and RBS-R Stereotypic Behavior subscales (.73).

3.3. Normative data

MANOVA results revealed significant multivariate effects for age on the combined ABC subscale scores, F (10, 724) = 3.174, p = .001, Wilks' Λ = 0.918, partial η^2 = .042. No significant differences for gender ($p >$ .05, partial η^2 = .007) and functional level ($p >$ .05, partial η^2 = .020) on the combined ABC subscale scores were found. Multivariate results indicate a significant interaction between age and functional level, F (10, 724) = 2.485, p = .006, Wilks' Λ = 0.935, partial η^2 = .033. Given these results, and to allow for comparison with norms published by Kaat et al, norms are presented in Table 5 by age group and IQ. More detailed tables, showing the effects of the interactions, are available on request from the corresponding author.

Univariate analyses indicate significant effects of age on the following ABC subscales: Irritability (F (2, 366) = 3.969, p = .020,

Table 3
Factor Correlations.

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Factor 1	1.00	0.467	0.438	0.713	0.409
Factor 2	0.467	1.00	0.524	0.496	0.338
Factor 3	0.438	0.524	1.00	0.514	0.458
Factor 4	0.713	0.496	0.514	1.00	1.00
Factor 5	0.409	0.338	0.458	0.486	0.486

Table 4
Correlations between ABC and other Instruments.

	Irritability	Lethargy/Social Withdrawal	Stereotypic Behavior	Hyperactivity	Inappropriate Speech
VABS[a]: Socialization	−.30**	**−.38****	−.32**	−.31**	−.16**
VABS: Daily Living Skills	−.25**	−.33**	−.34**	−.29**	−0.03
VABS: Communication	−.16**	−.37**	−.30**	−.23**	0.04
SRS[b] Total	.45**	**.63****	.53**	.49**	.39**
RBS-R[c]: Stereotyped Behavior	.44**	.42**	**.73****	.51**	.34**
RBS-R: Self-injurious Behavior	**.51****	.28**	.34**	.36**	.25**
RBS-R: Total	.67**	.53**	.54**	.58**	.51**
OACIS[d]: Social Interaction	0.06	**.35****	.29**	.11*	0.05
OACIS: Aberrant and Abnormal	**.38****	.20**	.18**	.29**	.11*
OACIS: Repetitive and Ritualistic	.18**	.26**	**.37****	.20**	.13**
OACIS: Verbal Communication	0.04	.27**	.30**	0.09	**−0.06**
OACIS: Hyperactivity and Inattention	.10*	0.08	.13**	**.30****	0.08
OACIS: Total	.13**	.36**	.35**	.18**	0.06

Note. Anticipated correlations are bolded. **Significant at the .01 level, * Significant at the .05 level.
VABS = Vineland Adaptive Behavior Scales, SRS = Social Responsiveness Scale, RBS-R = Repetitive Behavior Scale-Revised, OACIS = Ohio Autism Clinical Impairment Scale Severity.

partial η^2 = .021), Lethargy/Social Withdrawal (F (2, 366) = 3580, p = .029, partial η^2 = .019), Stereotypic Behavior (F (2, 366) = 3.725, p = .025, partial η^2 = .020), and Inappropriate Speech (F (2, 366) = 8.982, p = .000, partial η^2 = .047). Scheffe post-hoc comparisons indicated that teens significantly differed from school-aged individuals on measures of Irritability (p = .001, Cohen's d = .498) and from preschool-aged children on measures of Irritability (p = .047, Cohen's d = .252) and Lethargy/Social Withdrawal (p = .006, Cohen's d = .410). School-aged children significantly differed from preschool children on measures of Inappropriate Speech (p = .001, Cohen's d = .362). Univariate testing also indicated the interaction between age and functional level to be significant on the ABC subscale Inappropriate Speech (F (2, 366) = 34.660, p = .010, partial η^2 = .025).

4. Discussion

Results of this study provide further support for use of the ABC in youth diagnosed with ASD. Regarding the first goal of examining factor validity, we chose to retain the original factor structure for theoretical and practical reasons. In this sample, fit indices were most supportive of this structure compared to other tested models. Fit indices for the four- and five-factor solutions proposed by Brinkley et al. were similar to indices for the original structure, but the original factor structure has been supported across 14 other studies with a variety of samples (Aman & Singh, 2017), suggesting it is generally robust and therefore preferable. Additionally, the largest previous study to examine the ABC's psychometric properties in an ASD sample (Kaat. et al.) also opted to retain the original factor structure as the model of choice, lending further support. The other models tested in this sample had poorer fits and less theoretical support for their retention. For instance, the 4-factor model proposed by Marshburn and Aman was developed using data provided by teachers who rated their students with developmental disabilities; and the 4-factor model proposed by Brown et al. was developed in a sample a community-based sample of children and adolescents receiving special education services, rated by their parents. Whenever CFA is employed, it is important to consider the possibility of "over-fitting" the model such that unique traits of a sample may suggest a solution which fits well for that particular sample but lacks generalizability. The goal of this study was not to specify a model that worked well for the current sample, but to examine multiple previously proposed models and determine which model(s) were supported within the context of other data. Considering results of the current and previous studies, the original factor structure appears to function well across a wide range of individuals, suggesting it is a preferable solution.

Alpha values also were similar to previous studies, and in the good-to-excellent range on all subscales except Inappropriate Speech. Across studies, the Inappropriate Speech subscale has behaved somewhat less consistently than other subscales in general, as it did not emerge as a factor in some previous FA studies (e.g., Brinkley et al., Brown et al., Marshburn et al.) and its alpha level is consistently lower than that for other subscales. It is likely that the small number of items on this subscale (only four) contributes to this. As all four of these items require presence of verbal language, it may be that samples that include very young, nonverbal, and/or minimally verbal individuals would not have high endorsement rates of these items, affecting how items behave when analyzed. (Indeed, 52% of our sample comprised preschoolers.) However, conceptually and clinically, this subscale is relevant for those with ASD and can help capture verbal behaviors that may be disruptive and warrant treatment.

Results of convergent and divergent validity analyses also lend support to the validity of the ABC in that anticipated correlations were found for predetermined variables in the expected directions. All domain scores on the VABS were found to be negatively correlated with all ABC scales except Inappropriate Speech (perhaps due to reasons discussed above), though the Inappropriate Speech factor was negatively associated with the VABS Socialization domain. It is possible that behaviors captured by this scale (e.g., repetitive speech) interfere with social interactions more than they do with self-care skills, safety awareness, following directions, or similar adaptive skills measured by other VABS domains.

Normative analyses with this sample yielded similar results to the only previously published norms for those with ASD (Kaat

Table 5
Normative Data.

	Sample Size			Irritability			Lethargy/Social Withdrawal			Stereotypic Behavior			Hyperactivity/noncompliance			Inappropriate Speech		
	All	IQ≤70	IQ > 70	All	IQ≤70	IQ > 70	All	IQ≤70	IQ > 70	All	IQ≤70	IQ > 70	All	IQ≤70	IQ > 70	All	IQ≤70	IQ > 70
pre-school	197	76	121	11.6 (8.9)	11.5 (8.6)	11.7 (9.2)	8.8 (7.1)	11.2 (7.6)	7.3 (6.3)	4.1 (4.4)	5.2 (4.8)	3.5 (4.0)	18.8 (11.3)	19.4 (10.4)	18.5 (11.9)	3.1 (2.9)	2.4 (2.5)	3.6 (3.0)
school-age	110	19	91	13.8 (9.1)	16.9 (7.9)	13.2 (9.2)	9.6 (6.6)	13.9 (5.0)	8.7 (6.6)	4.7 (4.3)	7.7 (4.2)	4.0 (4.1)	20.0 (10.2)	23.7 (8.7)	19.2 (10.4)	4.1 (2.7)	5.6 (2.6)	3.8 (2.6)
teens	71	8	63	9.4 (9.4)	11.1 (9.5)	9.1 (9.4)	12.1 (8.8)	13.1 (10.8)	12.0 (8.6)	3.6 (3.5)	3.1 (2.4)	3.7 (3.6)	14.2 (8.7)	17.3 (7.9)	13.8 (8.8)	3.2 (2.4)	4.1 (2.2)	3.1 (2.4)
Full sample	378	103	275	11.8 (9.2)	12.5 (8.7)	11.6 (9.3)	9.6 (7.4)	11.9 (7.5)	8.8 (7.2)	4.2 (4.2)	5.5 (4.7)	3.7 (3.9)	18.3 (10.7)	20.0 (10.1)	17.6 (10.9)	3.4 (2.8)	3.1 (2.8)	3.5 (2.8)

et al.). The pattern of scores across subtests was similar, with mean subscale scores for the whole sample falling within 1.0 point of each other. In general, means from the current sample were slightly lower than those found by Kaat et al., possibly due to this sample being somewhat higher functioning (e.g., 46% of the Kaat et al. sample had IQs ≤ 70, while only 27% of the current sample functioned within this range). Indeed, in all cases, mean scores were lower for the > 70 IQ group in our normative data (Table 5). While normative data from the previous study was developed with a larger sample size and greater heterogeneity, there may be a diagnostic advantage to considering current norms given this sample was comprised of individuals meeting DSM-5 diagnostic criteria. Regardless, the similarity of scores between studies gives confidence that they are robust and reliable.

4.1. Limitations

While this study contributes to the body of evidence on the ABC, several limitations are noteworthy. While this study involved a relatively large and well-characterized sample of individuals with ASD, there was limited heterogeneity in terms of level of functioning. In particular, most individuals in this sample had cognitive levels above the ID range. Though this is consistent with some recent data (e.g., Elsabbagh et al., 2012) it may limit generalizability of results to those with lower IQs. Additional limitations of this sample included relatively few teenagers and very few teenage females, which may have affected normative analyses and suggests results may be less reliable in this subgroup.

4.2. Implications

There is consensus that individuals with ASD are at greater risk for experiencing behavior and emotional difficulties, yet assessing these problems is hampered by a lack of instruments developed for and validated with this population. Identification of these difficulties is a necessary prerequisite to obtaining appropriate intervention. The ABC has been widely used and validated in many samples, making it an excellent candidate for use in those with ASD. The only two previous studies examining its factor validity in those with ASD yielded different results, and prior to this study, only one set of normative data for those with ASD had been published. In addition to adding to the larger literature supporting the psychometric properties of the ABC generally, this study provides an important contribution to the relatively small but growing body of work on the ABC's validity for youth with ASD. Results support its use not only because factor, convergent, and divergent validity were supported, but also because results indicate the ABC is useful for those across ages and functioning levels, which is especially important given the heterogeneity of presentation in ASD. Normative results also draw attention to the need to consider possible age-related differences in presentation, which may be a fruitful area for further inquiry.

In conclusion, the robust psychometric performance of the ABC in multiple studies, including this one, provides compelling data for professionals to consider when seeking an instrument to assess behavioral and emotional challenges in youth with ASD. This is especially exciting given the psychometric support for its use in other populations of individuals with developmental disabilities, suggesting the ABC had wide applications in both research and clinical contexts.

Conflict of interest

The authors have no conflicts of interest to disclose.

Acknowledgement

This work was supported by the National Institutes of Health (grant number R01 MH 097726).

References

Achenbach, T. M., & Rescorla, L. A. (2000). *CBCL/1.5-5 & TRF/1.5-5 profiles. Research center for children, youth and families.* Burlington: University of Vermont.
Aman, M. G., & Singh, N. (1985). *The aberrant behavior checklist.* East Aurora, NY: Slosson Educational Publications.
Aman, M. G., & Singh, N. N. (2017). *Aberrant behavior checklist manual* (second edition). East Aurora, NY: Slosson Educational Publications, Inc.
Baumgardner, T. L., Reiss, A. L., Freund, L. S., & Abrams, M. T. (1995). Specification of the neurobehavioral phenotype in males with Fragile X Syndrome. *Pediatrics, 95*(5), 744–752.
Bodfish, J. W., Symons, F. J., & Lewis, M. H. (1999). *). The repetitive behavior scale. Western carolina research reports.*
Brinkley, J., Nations, L., Abramson, R. K., Hall, A., Wright, H. H., Gabriels, R., et al. (2007). Factor analysis of the Aberrant Behavior Checklist in Individuals with Autism Spectrum Disorders 1949-1959(ASD). *Journal of Autism and Developmental Disorders, 37*, 1949–1959.
Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (2nd ed.). New York, NY: Guilford Press.
Brown, E. C., Aman, M. G., & Havercamp, S. M. (2002).). Factor analysis and norms for parent ratings on the Aberrant Behavior Checklist-Community for young people in special education. *Research in Developmental Disabilities, 23*, 45–60.
Browne, M. W., & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological Methods & Research, 21*(2), 230–258.
Buck, T. R., Viskochil, J., Farley, M., Coon, H., McMahon, W. M., Morgan, J., et al. (2014). Psychiatric comorbidity and medication use in adults with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 44*, 3063–3071.
Butter, E., & Mulick, J. (2006). *The Ohio Autism Clinical Impressions Scale (OACIS).* Columbus: OH: Children's Research Institute.
Clarke, D. J., & Dykens, E. M. (1997). Correlates of maladaptive behavior in individuals with 5p-(cri du chat) syndrome. *Developmental Medicine and Child Neurology, 39*, 752–756.
Clarke, D. J., Boer, H., Chung, M. C., Sturmey, P., & Webb, T. (1996). Maladaptive behavior in Prader-Willi syndrome in adult life. *Journal of Intellectual Disability Research, 40*, 159–165.
Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*, 155–159.

Constantino, J. N., & Gruber, C. P. (2012). *Social responsiveness scale, (SRS-2)*. Torrance, CA: Western Psychological Services.
Elliott, C. D. (2007). *Differential ability scales* (2nd ed). San Antonio, TX: Harcourt Assessment.
Elsabbagh, M., Divan, G., Koh, Y. J., Kim, Y. S., Kauchali, S., Marcín, C., et al. (2012). Global prevalence of autism and other pervasive developmental disorders. *Autism Research, 5*, 160–179.
Gotham, K., Risi, S., Pickles, A., & Lord, C. (2007). The Autism Diagnostic Observation Schedule: Revised algorithms for improved diagnostic validity. *Journal of Autism and Developmental Disabilities, 37*.
Hanratty, J., Livingstone, N., Robalino, S., Terwee, C. B., Glod, M., Oono, I. P., et al. (2015). Systematic review of the measurement properties of tools used to measure behavior problems in young children with autism. *PloS One, 10*(12), e0144649. https://doi.org/10.1371/journal.pone.0144649.
Hill, J., Powlitch, S., & Furniss, F. (2008). Convergent validity of the Aberrant Behavior Checklist and Behavior Problems Inventory with people with complex needs. *Research in Developmental Disabilities, 29*, 45–60.
Kaat, A. J., Lecavalier, L., & Aman, M. G. (2014). Validity of the Aberrant Behavior Checklist in children with autism spectrum disorder. *Journal of Autism and Development Disorders, 44*, 1103–1116.
Kanne, S. M., Mazurek, M. O., Sikora, D., Bellando, J., Branum-Martin, L., Handen, B., et al. (2014). The Autism Impact Measure (AIM): Initial development of a new tool for treatment outcome measurement. *Journal of Autism and Developmental Disorders, 44*, 168–179.
Leyfer, O. T., Folstein, S. E., Bacalman, S., Davis, N. O., Dinh, E., Morgan, J., et al. (2006). Comorbid psychiatric disorders in children with autism: Interview development and rates of disorders. *Journal of Autism and Developmental Disorders, 36*, 849–861.
Lord, C., Rutter, M., DiLavore, P., Risi, S., Gotham, K., & Bishop, S. (2012). *Autism diagnostic observation schedule – 2nd edition (ADOS-2)*. Los Angeles, CA: Western Psychological Corporation.
Marshburn, E. C., & Aman, M. G. (1992). Factor validity and norms for the Aberrant Behavior Checklist in a community sample of children with mental retardation. *Journal of Autism and Developmental Disabilities, 22*, 357–373.
Mullen, E. M. (1995). *Mullen scales of early learning: AGS edition.* Circle Pines, MN: American Guidance Service.
Sansone, S. M., Widaman, K. F., Hall, S. S., Reis, A. L., Lightbody, A., Kaufmann, W. E., et al. (2012). Psychometric study of the Aberrant Behavior Checklist in Fragile X Syndrome and implications for targeted treatment. *Journal of Autism and Developmental Disorders, 42*, 1377–1392.
Siminoff, E., Pickles, A., Charman, T., Chandler, S., Loucas, T., & Baird, G. (2008). Psychiatric disorders in children with autism spectrum disorders: prevalence, comorbidity, and associated factors in a population-derived sample. *Journal of American Academy of Child and Adolescent Psychiatry, 47*, 921–929.
Sparrow, S. S., Balla, D. A., & Cicchetti, D. V. (2005). *Vineland adaptive behavior scales* (second edition). Circle Pines, MN: AGS Publishing.
Summers, J. A., & Feldman, M. A. (1999). Distinctive pattern of behavioral functioning in Angelman Syndrome. . *American Journal on Mental Retardation, 104*, 376–384.