Paidéia
2021, Vol. 31, e3125.doi:https://doi.org/10.1590/1982-4327e3125
ISSN 1982-4327 (online version)

Psychological Evaluation

Validity Evidence of the Autistic Spectrum Disorder Behavior Scale

Camila Costa e Silva[1]

Daniela Sacramento Zanini[1]

Abstract: Screening instruments to detect symptoms of Autistic Spectrum Disorder (ASD) in the school setting are still needed. The study aimed to develop the Autistic Spectrum Disorder Behavior Scale (ASD-BS) according to the criteria of the Diagnostic and Statistical Manual of Mental Disorders (DSM-V). The scale consists of 31 items, divided into communication; social interactions; restrictive, repetitive and ritual behaviors; and other indicators. The content validity included five expert judges and a pilot sample with 29 children between six and 12 years old. The results were shown to be adequate, all items reached agreement rates above 80% and *Kappa* of 0.84. Exploratory statistical analyzes showed *Cronbach*'s alpha equal to 0.78, sensitivity values ranging from 86 to 93% and specificity from 34 to 100%. ASD-BS is of quick implementation, uses the teacher as an informant and contributes to psychological assessment.

Keywords: autism spectrum disorder, psychological assessment, children

Evidências de Validade da Escala de Comportamentos do Transtorno do Espectro Autista

Resumo: Instrumentos de rastreio para detectar sintomas do Transtorno do Espectro Autista (TEA) no contexto escolar ainda são necessários. O estudo teve por objetivo a construção da Escala de Comportamentos do Transtorno do Espectro Autista (EC-TEA) de acordo com os critérios do Manual Diagnóstico e Estatístico de Transtornos Mentais (DSM-V). A escala é composta por 31 itens, divididos em comunicação; interações sociais; comportamentos restritivos, repetitivos e rituais; e outros indicadores. A validade de conteúdo contou com cinco juízes especialistas e amostra piloto com 29 crianças entre seis e 12 anos. Os resultados se mostraram adequados, todos os itens alcançaram índices de concordâncias acima de 80% e *Kappa* de 0,84. As análises estatísticas exploratórias demonstraram Alfa de *Cronbach* igual a 0,78, valores de sensibilidade variando de 86 a 93% e especificidade de 34 a 100%. EC-TEA é de rápida aplicação, utiliza o professor como informante e contribui com a avaliação psicológica.

Palavras-chave: transtorno do espectro autista, avaliação psicológica, crianças

Validez de la Escala de Comportamiento del Trastorno del Espectro Autista

Resumen: Aún se necesitan instrumentos de cribado para detectar los síntomas del trastorno del espectro autista (TEA) en el contexto escolar. Este estudio tuvo como objetivo construir la Escala de Comportamiento del Trastorno del Espectro Autista (EC-TEA) según los criterios del Manual Diagnóstico y Estadístico de los Trastornos Mentales (DSM-V). La escala consta de 31 ítems, divididos en comunicación; interacciones sociales; conductas restrictivas, repetitivas y rituales; y otros indicadores. La validez de contenido incluyó a cinco jueces expertos y la muestra piloto con 29 niños de entre 6 y 12 años de edad. Los resultados demostraron ser adecuados, todos los ítems alcanzaron tasas de acuerdo superiores al 80% y el *Kappa* de 0,84. Los análisis estadísticos exploratorios mostraron un alfa de Cronbach igual a 0,78, valores de sensibilidad que oscilan entre el 86% y el 93% y una especificidad del 34% al 100%. EC-TEA se aplica rápidamente, utiliza al profesor como informante y contribuye a la evaluación psicológica.

Palabras clave: transtorno del espectro autista, evaluación psicológica, niños

[1]*Pontifícia Universidade Católica de Goiás, Goiânia-GO, Brazil*

Article derived from the master's thesis of the first author under the supervision of the latter, defended in 2019, in the *Stricto Sensu* Postgraduate Program in Psychology at Pontifícia Universidade Católica de Goiás.

Correspondence address: Camila Costa e Silva. Pontifícia Universidade Católica de Goiás. Praça Universitária, 1440 – Setor Leste Universitário, Goiânia-GO, Brasil. CEP. 74.605-010. E-mail: camilacosta1802@gmail.com

The history of Autism is marked by extensive reformulations and, at first, the condition was associated with childhood schizophrenia or mental retardation. The definition of autism has changed over time as the characterization of autism and other disorders has been reorganized (Sella & Ribeiro, 2018). The disorder encompasses the so-called early childhood autism, childhood autism, Kanner's autism, high-functioning

autism, atypical autism, pervasive developmental disorder not otherwise specified, childhood disintegrative disorder, and Asperger's syndrome. Although multiple definitions are found, there is a continuity in the characterization of autism since Kanner, which is related to social, linguistic, and unusual behavior characteristics (Lins & Borsa, 2017).

The definition used in this study is the description published in 2014 by the American Psychiatric Association [APA] in the Diagnostic and Statistical Manual of Mental Disorders, fifth edition (DSM-V), which brings Autistic Spectrum Disorder (ASD) as characterized by persistent impairments in reciprocal social communication, social interaction, and restricted, repetitive patterns of behavior, interests or activities. Symptoms are present from early childhood and limit or impair the individual's functioning. The Autistic Spectrum Disorder (ASD) is often associated with intellectual impairment and language disorder. About 70% of people with ASD have a comorbid mental disorder and 40% may have two or more comorbid mental disorders, therefore, the relevant specifiers must be registered in the evaluation (APA, 2014; Tisser, 2018).

In the diagnosis of autism, methods are used to obtain information, such as direct observations and interviews, applied to the child and/or their parents, teachers, and possible caregivers who are in contact with them. The observational approach fits structured or semi-structured situations, on the other hand, there are instruments that use interviews or checklists (Bosa & Teixeira, 2017). Studies carried out in Brazil use instruments from the international literature considered the gold standard for the diagnosis of ASD (Autism Diagnostic Interview-Revised – ADI-R and Autism Diagnostic Observation Schedule-Generic – ADOS) and screening instruments, such as: Autistic Traits of Evaluation Scale (ATA), Autism Behavior Checklist (ABC), Childhood Autism Rating Scale (CARS), Autism Screening Questionnaire (ASQ) and Modified Checklist for Autism in Toddlers (M-CHAT). At last, there is the Behavioral Assessment Protocol for Children referred for a possible Autism Spectrum Disorder (PROTEA-R) for non-verbal children aged from 24 and 60 months, created by Brazilian researchers (Bosa & Salles, 2018; Bosa, Zanon, & Backes, 2016).

In a systematic review of the literature of the last 10 years, Silva and Elias (2020) characterized the articles that referred to the validation of instruments for screening and/or diagnosis of ASD in Brazil. Thus, it was found that there are no instruments available in the Psychological Testing Assessment System (SATEPSI) for ASD assessment, in other words, the instruments mentioned in the literature do not present analysis reports by SATEPSI and there are no instruments that use teachers as informants.

In these review findings, the instruments presented sensitivity and specificity data, namely: Behavior Problems Inventory (BPI-01), ranging from 0.65 to 0.91; Social Responsiveness Scale-2 (SRS-2) of 0.952; Autism Diagnostic Interview – Revised (ADI-R) of 0.976; Children's Communication Checklist-2 (CCC-2) of 0.93; and the Clinical Risk Indicators for Child Development (IRDI) instrument of 0.89; as well as Autism Screening Questionnaire (ASQ) of 0.895. Regarding the *ROC* curve values, the studies described the following data: BPI-01 had sensitivity between 50 and 80% and specificity between 3 and 10%; SRS-2, 96.8% sensitivity and 100% specificity; ADI-R, 100% sensitivity and specificity; and ASQ, with a sensitivity of 92.5% and specificity of 95.5% (Baraldi, Rojahn, Seabra, Carreiro, & Teixeira, 2013; Barbosa et al., 2015; Becker et al., 2012; Machado, Palladino, & Cunha, 2014; Sato et al., 2009).

In the international setting, we located the DSM-V teacher screening questionnaire for autistic spectrum disorder and social communication disorder (EDUTEA), a 4-point *Likert* scale questionnaire with a total of 11 questions and a sample of more than 291 children. The EDUTEA scale showed high internal reliability: social communication impairments (= 0.95) and restricted behavior patterns (= 0.93). The *ROC* curve provided high values for sensitivity (87%), specificity (91.2%) and positive predictive value (0.87) (Morales-Hidalgo, Hernández-Martínez, Voltas, & Canals, 2017). Although the populations are evidently different, the EDUTEA researchers used gold standard instruments in the evaluation, Autism Diagnostic Interview Revised (ADI-R) and the Autism Diagnostic Observation Schedule (ADOS-2) and diagnostic criteria from the DSM-V (Morales-Hidalgo et al., 2017).

Thinking of the teacher as a contributor to the assessment and a specialist in education, it is important to include them in the gathering of information. In a studied sample, 37.7% of interviewed teachers reported that they have an important role in helping early diagnosis. Many said they had difficulties due to lack of diagnosis. Behavioral changes are noted, but there are no referrals for proper assessment. Another obtained data is that 62% of teachers described knowing the main characteristics of autism, but 91.4% felt unprepared to work with children with ASD. In this sample, according to the teachers, education and work experience did not influence the aspects considered important for working (Castro & Giffoni, 2017).

Stereotyped obsessive behaviors, introversion, self-stimulation, and incomprehensible speech are the most significant obstacles teachers have considered in the educational integration of students with autism. Among the surveyed teachers, 28.1% had relevant training, experience and education in autism, but only 20.5% felt able to face the daily problems of students with this disorder. Overall, the results showed that teachers are positive for the integration of students with autism (Cassimos, Polychronopoulou, Tripsianis, & Syriopoulou-Delli, 2015). Furthermore, agreement between informants, parents, and teachers is generally greater when the disorder is more severe (Azad, Reisinger, Xie, & Mandell, 2016). However, they do not agree with the social deficits of less affected children and the variability of reports may be correlated with differences in environmental demands or in the experiences of informants.

Generally, teachers notice more deficits in social behaviors than parents. Prosocial behavior refers to positive interactions with other people, when they help, share, cooperate, and comfort. These actions can be better evaluated in a group setting at school than in a one-on-one situation at home (Cassimos et al., 2013). Regarding children who were already diagnosed with autism, parents achieved higher scores on the assessed scale compared to those who were in the outpatient clinic without a diagnosis. This may be related to the fact that parents of diagnosed children already know about the disorder and what they should observe

and assess in their child. Parents of undiagnosed children reflect the behavior of parents in clinical settings and are generally unaware of the ASD when the child is referred to the outpatient clinic (Mattila et al., 2009).

From this perspective, the construction of the scale is important to equip the psychologist in the professional and research environment. Instruments in psychological assessment are considered qualifying elements for psychological practice. However, for the variable measurement be reliable, the instrument needs to measure what it was designed to measure (Lins & Borsa, 2017). For behaviors represent traits, psychometrics analyzes operational definitions through theoretical and empirical or statistical analysis. Theoretical analysis is the process carried out by judges, whose objective is the semantic analysis of the items and the construct analysis. In the semantic analysis of items, there are two concerns: (a) check if the items are comprised by the lowest stratum of the population; (b) the sample should also contain a more sophisticated stratum to ensure the apparent validity of the test. In content analysis, it is intended to judge whether the items form a representation of the trait in question or not. The empirical analysis consists of a series of parameters, such as unidimensionality, difficulty, discrimination, biases, response bias, validity, and accuracy. The test must be valid and reliable. These procedures are ensured in the construction or adaptation of the test (Baptista et al., 2019).

The study aimed to develop the Autistic Spectrum Disorder Behavior Scale (ASD-BS) according to the criteria of the Diagnostic and Statistical Manual of Mental Disorders – DSM-V.

Method

Participants

Five female judges, aged between 24 and 48 years, psychologists, three of them graduate students and two professors participated in the evaluation of the Autistic Spectrum Disorder Behavior Scale. Regarding educational level, one of the judges had a doctoral degree, two had a master's degree and two were specialists.

For content validation in other groups, representatives of possible target samples were invited: children's parents, teachers, and caregivers. There were four representatives from each group, two males and the rest females, aged between 20 and 41 years and educational level between secondary and higher education.

The pilot sample consisted of 29 children aged from six and 12 years, divided into two groups: diagnosed with ASD (N = 17) and undiagnosed with developmental disorders (N = 12). The teachers of the respective children – all women, aged from 29 to 54 years – were the respondents of the ASD-BS scale.

Instruments

The *Autistic Spectrum Disorder Behavior Scale* (ASD-BS), developed by the authors of this study and submitted for validation, was an instrument with the objective of evaluating the characteristic behaviors of ASD, organized into four domains (31 items): communication (9 items); social interactions (9 items); restrictive, repetitive, and ritual behaviors (6 items); and other indicators (7 items). In each domain, selected according to the criteria set out in the DSM-V, behaviors that reflected the category and that could be observed were chosen. The teacher should read each sentence and answer, on a five-point *Likert* scale, how much the situation relates to the child, namely: (1) I totally disagree; (2) I partially disagree; (3) I agree; (4) I partially agree; and (5) I completely agree. Some sentences are inverted, that is, they present adapted behaviors, in this case, the score is also inverted on the *Likert* scale. The description of the items that compose the ASD-BS scale is presented in Table 1.

Table 1
Description of items covered in the ASD-BS scale

Domain	Number of Items	Description
Communication	9	Understanding or expression of speech; Use of gestural communication; Repetition of words or phrases immediately or previously heard; Confusion in the use of pronouns; Invention of words or phrases.
Social interactions	9	Interest and initiative in social interactions; Spontaneous smile; Variation in facial expression; Preference for individual activities; Pointing behavior to show objects or events; Directing the gaze to where people point; Comments to share interests.
Restrictive, repetitive and ritual behaviors	6	Repetitive movements of the face and body with no apparent purpose; Fixed and rigid sequence for carrying out activities; Significant and exaggerated interest in things involving a single theme; Attachment to unusual objects compared to peers; Restricted or very selective food intake.
Other indicators	7	Imitation; Very intense and discrepant fears in relation to peers; Hyposensitivity or hypersensitivity to sensory stimuli; Loss of skills previously acquired in childhood; Delay in physical or motor development; Make-believe game.

Procedures

The items were selected from the literature, according to the diagnostic criteria of the DSM-V and grouped into four dimensions: (1) communication; (2) social interactions;

(3) restrictive, repetitive and ritual behaviors; and (4) other indicators. The scale consists of 31 items, namely: nine items related to communication; nine items to social interactions; six items to restrictive, repetitive, and ritual behaviors; and seven items to other indicators. Pasquali (2010) suggests that around 20 items are a reasonable number to represent a construct. The items were elaborated considering the school environment and the behaviors subject to the teacher's observation.

The ASD-BS scale was evaluated by expert judges regarding the item adequacy to the ASD diagnosis proposed by the DSM-V. Thus, the instrument was sent by email to the judges along with the description of the disorder and they were asked to read each of the 31 items, in which they should mark one of the following options: bad, regular and good. In each item, a space was given to suggest changes. With the judges' answers, it was possible to calculate the index of agreement between them and to perform the *Kappa* calculation between the two most disagreeing judges.

The groups representing the possible target samples (children's parents, teachers, and caregivers) were asked to read the scale, assuming that they would be evaluating the child they are responsible for. After reading, it was asked if there was any doubt regarding the general information, indicated instructions, marking, and content displayed on the scale. Everyone was able to understand the items and there was no suggestion for changes.

Data collection. After the content validation steps, the collection was carried out in public schools and the children were randomly selected. Those responsible for them signed an informed consent form, as well as the teacher, authorizing the collection and participation in the research. For the application of the ASD-BS scale, the teacher was asked to answer the instrument according to the behavior of the assessed child, comparing it with children of the same age when necessary. Data were collected from 17 children diagnosed with ASD and from 12 children without any diagnosis.

Data analysis. Descriptive and exploratory analysis were performed to investigate the content validity of the ASD-BS scale. The analysis of agreement used, as a calculation, the formula that divides agreement by the number of agreements plus disagreement, and multiplies it by 100 (Fagundes, 2015). As interpretation parameters, the values with discrete agreement (0.00 to 0.20), regular agreement (0.20 to 0.40), moderate agreement (0.40 to 0.60), substantial agreement (0.60 to 0.80) and almost perfect agreement (0.80 to 1.00) were used. Regarding the *Kappa* calculation, the following recommendations were adopted as parameters: (a) above 0.75, adequate agreement, (b) between 0.40 and 0.75, satisfactory agreement, and (c) below 0.40, unsatisfactory agreement (Cleophas & Zwinderman, 2016). For the item to be considered adequate and maintained on the scale, it was defined that the judges should have an agreement above 80%, as recommended in the literature (Baptista et al., 2019).

Then, the data obtained from the pilot sample were entered into the *IBM SPSS Statistics* program, version 22, for the statistical validation of the instrument. Notably, the exposed analysis is exploratory, since the pilot sample is composed of 29 individuals. When the sample size is small, less than 30 pairs of measurements, *Spearman's* ρ coefficient, or rank-order correlation coefficient, is adopted. The value of *Spearman's* ρ coefficient varies between −1 and 1 and, thus, the closer it is to the extremes, the greater will be the association between the variables; the negative sign indicates that the categories vary in the opposite direction (Damásio & Borsa, 2018). For the analysis of internal consistency, *Cronbach's* alpha coefficient was used, whose value ranges from 0 to 1, 0 being the total absence of internal consistency of the items and 1 the presence of 100% consistency (Pasquali, 2010). To demonstrate the relationship between sensitivity and specificity, Receptor Operating Characteristic Curves (*ROC* Curves – Receiver Operating Characteristic) were used. The curve is a graph of sensitivity, or true positive rate versus false positive rate (Martinez, Louzada-Neto, & Pereira, 2003).

Ethical Considerations

This research was sent to the Research Ethics Committee for analysis of research projects of the Pontifícia Universidade Católica de Goiás (PUC-Goiás) and approved under the Certificate of Presentation for Ethical Consideration (CAAE nº 98127118.3.0000.0037). All study participants signed an informed form.

Results

The results will be presented in the sequence described in the procedural part of this study. Regarding the judges' analysis, the results indicate ideal agreement in all items of the ASD-BS scale and no item received a score equal to zero, that is, referring to poor quality, therefore, all items were selected. The average of the agreement indexes of the items evaluation by the judges, divided by criteria, was: communication (82%); social interactions (97%); restrictive, repetitive and ritual behaviors (93%); and other indicators (85%). Some agreement indexes were lower than others, thus, the two judges (one and four) who agreed the least in the assessment were selected for the analysis of the *Kappa* coefficient. Table 2 brings the cross tabulation of the results obtained and expected results and the statistical data achieved with the *Kappa* coefficient.

Table 2
Cross-tabulation of data, obtained and expected score and Kappa Statistics for the evaluation of selected judges

			Judge 4		
			Regular	Good	Total
Judge 1	Regular	Score	0	9	9
		Expected Score	1.5	7.5	9.0
	Good	Score	5	17	22
		Expected Score	3.5	18.5	22.0
	Total	Score	5	26	31
		Expected Score	5.0	26.0	31.0
			Number of valid cases		31
			Kappa agreement mean		0.84

Regarding the *Kappa* calculation, the result obtained of 0.84 indicates adequate agreement (Damásio & Borsa, 2018). The data show that all items on the ASD-BS scale present ideal agreement. From this perspective, there is no need to reformulate items and the pilot instrument will be able to proceed to the next steps of validation. In Table 3, the mean in the communication and interactions domains did not show significant differences between the groups, on the other hand, the behaviors and other indicators domains show differences in the mean between the groups and reject the null hypothesis.

Table 3

Comparison of mean between groups (without and with ADS) and Non-parametric test

Variable	Diagnosis		*U Mann Whitney P*
	Without ADS	With ADS	
	Mean *(SD)*	Mean *(SD)*	
Communication	21.00 (6.46)	24.94 (4.58)	0.16
Social interactions	23.33 (9.28)	24.52 (6.87)	0.91
Restrictive, repetitive and ritual behaviors	6.08 (0.28)	15.00 (4.86)	0.00*
Other indicators	12.00 (3.33)	18.70 (5.80)	0.00*

Note. *Reject the null hypothesis.

As can be seen in Table 4, the mean scores for each domain of the scale range from 11.31 (behaviors) to 24.59 (interactions). This situation allows for some dispersion of results, evidenced in the values of standard deviations, whose values ranged between 4.63 (other indicators) and 9.17 (interactions). Regarding the Internal Consistency of the four domains of the ASD-BS, it is verified that the *Cronbach'*s Alpha values surpass the threshold of 0.78, having oscillated between 0.30 (other indicators) and 0.82 (behaviors). *Spearman'*s correlation indicates that the correlations of domain compared to the general scale are significant at the 0.01 level. The behaviors and other indicators domains had a significant correlation at the 0.01 level with 0.61; and the communication and interactions domains showed a significant correlation at the 0.05 level, with 0.41.

The restrictive, repetitive, and ritual behaviors domain, analyzed based on the area under the ROC curve, demonstrated that, if randomly chosen, 96% of clinical cases will have higher scores than non-clinical cases in the ASD-BS ($AUC = 0.96$; $EP = 0.04$; $p < 0.01$; 95% $CI = 0.88 – 1.00$). In turn, the other indicators domain showed an 85% chance, being chosen randomly ($AUC = 0.85$; $EP = 0.07$; $p < 0.01$; 95% $CI = 0.70 – 1.00$). At last, communication had a chance of 72% ($AUC = 0.72$; $EP = 0.10$; $p = 0.05$; 95% $CI = 0.52 – 0.92$) and social interactions of 48% ($AUC = 0.48$; $EP = 0.12$; $p > 0.05$; 95% $CI = 0.24 – 0.72$). The differentiation of the graphics is shown in Figure 1.

The cutoff point that maximized sensitivity and specificity is equivalent to: restrictive, repetitive and ritual behaviors in 1/2 (i.e., scores up to 1 and from 2), with a sensitivity of 0.93 and a specificity of 1; other indicators in 6/7, with a sensitivity of 0.86 and a specificity of 0.75; communication on 10/11, with a sensitivity of 0.86 and a specificity of 0.59; and social interactions in 7/9, with a sensitivity of 0.86 and a specificity of 0.34.

Table 4

Mean, Standard Deviation, Cronbach's Alpha and Spearman's Correlation related to the domains of the ASD-BS scale

			α	r_{it}	α if item is deleted	*Spearman*'s Correlation			
	M	*SD*				Communication	Social interactions	Restrictive, repetitive and ritual behaviors	Other indicators
Communication	22.97	5.69	0.46	0.00 a 0.69	0.31 a 0.51	-			
Social interactions	24.59	9.17	0.79	0.03 a 0.89	0.73 a 0.87	0.41*	-		
Restrictive, repetitive and ritual behaviors	11.31	5.78	0.82	0.08 a 0.86	0.74 a 0.84	0.19	-0.01	-	
Other indicators	15.24	4.63	0.30	0.04 a 0.84	0.02 a 0.56	0.21	-0.02	0.61**	-
Scale Total	74.10	17.03	0.78	0.00 a 0.89	0.76 a 0.80	0.69**	0.56**	0.63**	0.64**

*Note.** The correlation is significant at the 0.05 level (bilateral);** The correlation is significant at the 0.01 level (bilateral);*r it* = correlation between items.

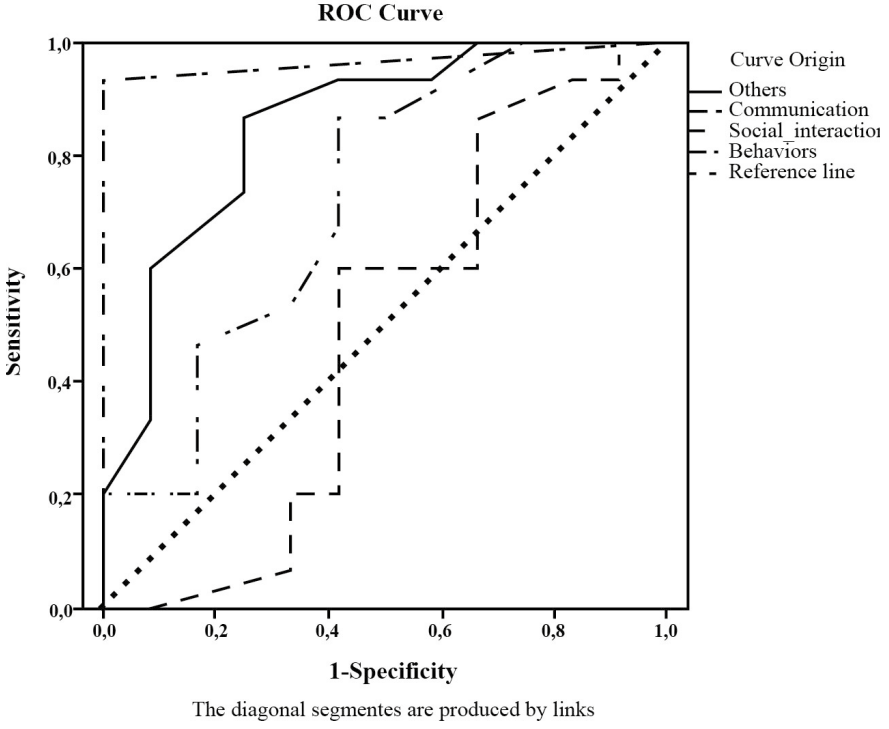

Figure 1. Analysis of the sensitivity and specificity of the domains of communication, social interactions, restrictive, repetitive and ritual behaviors, and other indicators.

Discussion

In order to contribute to the area of psychological assessment by proposing the construction of the Autistic Spectrum Disorder Behavior Scale (ASD-BS) according to the diagnostic criteria in force in the Diagnostic and Statistical Manual of Mental Disorders – DSM-V, the study obtained results that were adequate to the criteria that the scale intends to assess.

Considering the content validation of the ASD-BS scale, all items were classified as adequate by the judges, reaching agreement indexes above 80%, classified as almost perfect agreement, and obtained a *Kappa* index of 0.84, indicating adequate agreement (Cleophas & Zwinderman, 2016). The ADI-R validation study showed a mean *Kappa* among observers of 0.824 (Becker et al., 2012), considered a gold standard instrument in the assessment of ASD and comparing the results, having been verified that the ASD-BS scale exhibits good indicators of content validation.

Results subsequent to the application of the ASD-BS test in the pilot sample are exploratory due to the low number of participants, but they signal initial evidence of the scale's domains. *Spearman'*s correlation showed significant domain correlations compared to the general scale at the $p \leq 0.01$ level; the behavior and other indicators domains had a correlation equal to $r = 0.61$ at the $p \leq 0.01$ level; and the communication and interactions domains with correlation equal to $r = 0.41$ at the ≤ 0.05 level. The correlations presented for the domains are classified as moderate (between 0.40 − 0.69) (Cavallo, 2020).

Another point to be considered is that the larger the sample size, the greater the confidence that the result is different from the expectation given by the null hypothesis (Damásio & Borsa, 2018). This shows that the results considered consistent in the analysis may come to be confirmed and the oscillations between domains may change.

The null hypothesis considers that there is equality of population means, while the research hypothesis states that population means are not equal. The null hypothesis is rejected when there is an exceptionally low probability of a random error occurring in the obtained results (probability equal to or less than 0.05) and thus we infer that the research hypothesis is correct by rejecting the null hypothesis (Cleophas & Zwinderman, 2016). This means that the data obtained in the *U Mann Whitney* analysis indicate that the behavior and other indicators domains are scored differently in the groups (with and without diagnosis).

Regarding the *Cronbach'*s alpha, the instruments that assess the ASD with Brazilian samples (in the last ten years) had reliability indices ranging from 0.65 to 0.97. In this study, the general scale ASD-BS obtained a *Cronbach'*s Alpha equal to 0.78, with oscillations between 0.30 (other indicators) and 0.82 (behaviors). Compared to the instruments applied in the Brazilian samples, the pilot study brought adequate reliability indices. However, it is necessary to increase the number of the sample to better assess the Alpha and verify whether the domains that received the lowest scores will suffer variations.

Some studies in the literature described the *ROC* curve values with sensitivity ranging from 50 to 100% and specificity

between 3 and 100%. The ASD-BS scale produced sensitivity values ranging from 86 to 93% and specificity between 34 and 100%. The *ROC* curve was significant in the domains of restrictive, repetitive and ritual behaviors and other indicators, with results ranging from good to adequate (above 0.85). This variation is considered acceptable in communication (= 0.72), but in social interactions it had an unsatisfactory area under the curve (= 0.48). ASD-BS showed greater ability to correctly classify those who did not have ASD (100% of cases) when compared to those who had ASD (93% of cases) in the domain of restrictive, repetitive and ritual behaviors.

Using validity studies in Brazil as a parameter, no instrument for the assessment or screening of the ASD is directed to the school setting, therefore, we do not have an instrument with the same proposal as the ASD-BS validated in the country. In the international setting, we located the DSM-V teacher screening questionnaire for autistic spectrum disorder and social communication disorder (EDUTEA), high internal reliability: social communication impairments (= 0.95) and restricted behavior patterns (= 0.93). The *ROC* curve provided high values for sensitivity (87%) and specificity (91.2%). Despite the limitations of the sample, these indications demonstrate that ASD-BS presents adequate results when compared to EDUTEA.

The teacher's reports on child psychopathology are favorable as they allow comparison of children's development within their normative peer group. Overall, teachers notice more deficits in social behavior than parents (Azad et al., 2016; Castro & Giffoni, 2017). The ASD-BS scale is designed to serve the respondent teacher, as this brings a perspective of comparison between peers in the school environment. Compared to parents, teachers identify more behaviors related to the disorder when it is of mild severity (Cassimos et al., 2013). The ASD-BS scale also validated its content for other groups, such as parents and caregivers, so the applicator could compare informants.

The research effort in preparing the instrument is justified by the need to have an instrument validated by the Federal Council of Psychology for the use of professional psychologists. There is an ethical clash with the instruments available in the literature for the assessment of ASD and their use in the professional environment should consider copyright, commercialization or availability by the authors. In addition, validation studies may not be released for use in the clinical setting (Backes, Mônego, Bosa, & Bandeira, 2014).

The professional psychologist does not have any instrument approved by SATEPSI, that is, the instruments have not passed through the Consultative Committee on Psychological Assessment of the Federal Council of Psychology [CFP] (2018), which assesses the technical-scientific quality of psychological tests and regulates the assessment in the professional context (Resolution No. 9, of 2018). Considering the diagnostic criteria of the DSM, we found that some instruments with validation studies in Brazil were aimed at evaluating or screening ASD, but only one used the criteria of the fifth edition (Barbosa et al., 2015), the others used previous versions of the manual.

The ASD-BS is an instrument of quick implementation and uses the teacher as an informant of the characteristic behaviors of ASD. The instruments can be useful in research and clinical practice to discriminate children with ASD from typical children. The results of the present study indicate that all items reached agreement indexes above 80% and *Kappa* of 0.84. Exploratory statistical analysis demonstrated adequate consistency, with *Cronbach's* alpha equal to 0.78, sensitivity values ranging from 86 to 93%, and specificity from 34 to 100%. Although the initial findings point to positive aspects, this study was carried out with a small sample, obtained in a restricted area of the country, therefore, future studies should seek larger samples, which include regional, cultural and social variations, besides the characterization of the cognitive and language profile, using control groups with and without comorbidities in the Autistic Spectrum Disorder.

References

American Psychiatric Association. (2014). *Manual diagnóstico de transtornos mentais: DSM-5* [Diagnostic and statistical manual of mental disorders: DSM-5] (M. I. C. Nascimento, P. H. Machado, R. M. Garcez, R. Pizzato, & S. M. M. Rosa, Trans., 5th ed.). Porto Alegre, RS: Artmed.

Azad, G. F., Reisinger, E., Xie, M., & Mandell, D. S. (2016). Parent and teacher concordance on the social responsiveness scale for children with autism. *School Mental Health, 8*(3), 368-376. doi:10.1007/s12310-015-9168-6

Backes, B., Mônego, B. G., Bosa, C. A., & Bandeira, D. R. (2014). Psychometric properties of assessment instruments for autism spectrum disorder: A systematic review of Brazilian studies. *Jornal Brasileiro de Psiquiatria, 63*(2), 154-164. doi:10.1590/0047-2085000000020

Baptista, M. N., Muniz, M., Reppold, C. T., Nunes, C. H. S. S., Carvalho, L. F., Primi, R., … Pasquali, L. (2019). *Compêndio de avaliação psicológica* [Compendium of psychological assessment]. Petrópolis, RJ: Vozes.

Baraldi, G. S., Rojahn, J., Seabra, A. G., Carreiro, L. R. R., & Teixeira, M. C. T. V. (2013). Translation, adaptation, and preliminary validation of the Brazilian version of the Behavior Problems Inventory (BPI-01). *Trends in Psychiatry and Psychotherapy, 35*(3), 198-211. doi:10.1590/s2237-60892013000300007

Barbosa, I. G., Rodrigues, D. H., Rocha, N. P., Simões-e-Silva, A. C., Teixeira, A. L., & Kummer, A. (2015). Propriedades psicométricas da Escala de Responsividade Social-2 para Transtornos do Espectro Autista [Psychometric properties of the Social Responsiveness Scale-2 (SRS-2) for Autism Spectrum Disorder]. *Jornal Brasileiro de Psiquiatria, 64*(3), 230-237. doi:10.1590/0047-2085000000083

Becker, M. M., Wagner, M. B., Bosa, C. A., Schmidt, C., Longo, D., Papaleo, C., & Riesgo, R. S. (2012). Translation and validation of Autism Diagnostic Interview-Revised (ADI-R) for autism diagnosis in Brazil. *Arquivos de Neuro-Psiquiatria, 70*(3), 185-190. doi:10.1590/S0004-282X2012000300006

Bosa, C. A., & Salles, J. F. (2018). *Sistema PROTEA-R de avaliação da suspeita de Transtorno do Espectro Autista* [PROTEA-R system for the assessment of suspected Autism Spectrum Disorder]. São Paulo, SP: Vetor.

Bosa, C. A., & Teixeira, M. C. T. V. (Orgs.). (2017). *Autismo: Avaliação psicológica e neuropsicológica* [Autism: Psychological and Neuropsychological Assessment]. São Paulo, SP: Hogrefe.

Bosa, C. A., Zanon, R. B., & Backes, B. (2016). Autismo: Construção de um protocolo de avaliação do comportamento da criança – ProteaR [Autism: The construction of a behavioral assessment protocol for children - Protea-R]. *Psicologia: Teoria e Prática, 18*(1), 194-205. doi:10.15348/1980-6906/psicologia.v18n1p194-205

Cassimos, D. C., Polychronopoulou, S. A., Tripsianis, G. I., & Syriopoulou-Delli, C. K. (2015). Views and attitudes of teachers on the educational integration of students with autism spectrum disorders. *Developmental Neurorehabilitation, 18*(4), 241-251. doi:10.3109/17518423.2013.794870

Castro, A. C., & Giffoni, S. D. A. (2017). O conhecimento de docentes de educação infantil sobre o transtorno do espectro autístico [The knowledge of teachers in early childhood education on autism spectrum disorder]. *Revista Psicopedagogia, 34*(103), 98-106. Retrieved from http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S0103-84862017000100010&lng=pt&nrm=iso&tlng

Cavallo, B. (2020). Functional relations and Spearman correlation between consistency indices. *Journal of the Operational Research Society, 71*(2), 301-311. doi:10.1080/01605682.2018.1516178

Cleophas, T. J., & Zwinderman, A. H. (2016). *SPSS for starters and 2nd levelers*. Cham, Switzerland: Springer.

Conselho Federal de Psicologia. (2018, 25 de abril). *Resolução No. 9, de 25 de abril de 2018*. Estabelece diretrizes para a realização de Avaliação Psicológica no exercício profissional da psicóloga e do psicólogo, regulamenta o Sistema de Avaliação de Testes Psicológicos - SATEPSI e revoga as Resoluções No. 002/2003, No. 006/2004 e No. 005/2012 e Notas Técnicas No. 01/2017 e 02/2017 [Resolution No. 9, of April 25, 2018. Establishes guidelines for conducting Psychological Assessment in the professional practice of the psychologist and psychologist, regulates the Psychological Testing Assessment System - SATEPSI and repeals Resolutions No. 002/2003, No 006/2004 and No. 005/2012 and Technical Notes No. 01/2017 and 02/2017]. Retrieved from http://crp11.org.br/upload/Resolu%C3%A7%C3%A3o-CFP-n%C2%BA-09-2018-com-anexo.pdf

Damásio, B. F., & Borsa, J. C. (2018). *Manual de desenvolvimento de instrumentos psicológicos* [Manual for the development of psychological instruments]. São Paulo, SP: Vetor.

Fagundes, A. J. F. M. (2015). *Descrição, definição e registro de comportamento* [Behavior description, definition and record] (17a ed.). São Paulo, SP: Edicon.

Lins, M. R. C., & Borsa, J. C. (2017). *Avaliação psicológica: Aspectos teóricos e práticos* [Psychological assessment: Theoretical and practical aspects]. Petrópolis, RJ: Vozes.

Machado, F. P., Palladino, R. R. R., & Cunha, M. C. (2014). Adaptation of the Child Development Clinical Risk Indicators instrument to retrospective parent report. *CoDAS, 26*(2), 138-147. doi:10.1590/2317-1782/2014001IN

Martinez, E. Z., Louzada-Neto, F., & Pereira, B. B. (2003). A curva ROC para testes diagnósticos [Analysis of diagnostic tests using ROC curves]. *Cadernos de Saúde Coletiva, 11*(1), 7-31.

Mattila, M. L., Jussila, K., Kuusikko, S., Kielinen, M., Linna, S. L., Ebeling, H., … Moilanen, I. (2009). When does the Autism Spectrum Screening Questionnaire (ASSQ) predict autism spectrum disorders in primary school-aged children? *European Child and Adolescent Psychiatry, 18*(8), 499-509. doi:10.1007/s00787-009-0044-5

Morales-Hidalgo, P., Hernández-Martínez, C., Voltas, N., & Canals, J. (2017). EDUTEA: A DSM-5 teacher screening questionnaire for autism spectrum disorder and social pragmatic communication disorder. *International Journal of Clinical and Health Psychology, 17*(3), 269-281. doi:10.1016/j.ijchp.2017.05.002

Pasquali, L. (2010). *Instrumentação psicológica: Fundamentos e práticas* [Psychological instrumentation: Fundamentals and practices]. Porto Alegre, RS: Artmed.

Sato, F. P., Paula, C. S., Lowenthal, R., Nakano, E. Y., Brunoni, D., Schwartzman, J. S., & Mercadante, M. T. (2009). Instrument to screen cases of pervasive developmental disorder: A preliminary indication of validity. *Revista Brasileira de Psiquiatria, 31*(1), 30-33. doi:10.1590/s1516-44462009000100008

Sella, A. C., & Ribeiro, D. M. (2018). *Análise do comportamento aplicada ao transtorno do espectro autista* [Behavior analysis applied to autism spectrum disorder]. Curitiba, PR: Appris.

Silva, C. C., & Elias, L. C. S. (2020). Instrumentos de avaliação no Transtorno do Espectro Autista: Uma revisão sistemática [Assessment tools for Autism Spectrum Disorder: A systematic review]. *Avaliação Psicológica, 19*(2), 189-197. doi:10.15689/ap.2020.1902.09

Tisser, L. (2018). *Transtornos psicopatológicos na infância e na adolescência* [Psychopathological disorders in childhood and adolescence]. Novo Hamburgo, RS: Sinopsys.

Camila Costa e Silva is a Master's degree graduate of the Postgraduate Program in Psychology at Pontifícia Universidade Católica de Goiás, Goiânia-GO, Brazil.

Daniela Sacramento Zanini is a Professor of the Postgraduate Program in Psychology at Pontifícia Universidade Católica de Goiás, Goiânia-GO, Brazil.

Authors' Contribution:
All authors made substantial contributions to the conception and design of this study, to data analysis and interpretation, and to the manuscript revision and approval of the final version. All authors assume public responsibility for content of the manuscript.

Associate Editor:
Vanessa Barbosa Romera Leme

Received: Jul. 07, 2020
1st Revision: Jan. 16, 2021
Approved: Mar. 05, 2021

How to cite this article:
Silva, C. C., & Zanini, D. S. (2021). Validity evidence of Autistic Spectrum Disorder Behavior Scale. *Paidéia (Ribeirão Preto), 31,* e3125. doi:https://doi.org/10.1590/1982-4327e3125