Journal of Autism and Developmental Disorders (2018) 48:770–783
https://doi.org/10.1007/s10803-017-3438-0

ORIGINAL PAPER

Psychometric Properties of the Spanish Version of the Broad Autism Phenotype Questionnaire: Strengths, Weaknesses, and Future Improvements

Marta Godoy-Giménez[1] · Antonio González-Rodríguez[1] · Fernando Cañadas[2] · Angeles F. Estévez[2] · Pablo Sayans-Jiménez[1]

Published online: 27 December 2017

Abstract
The Broad autism phenotype (BAP) refers to a set of subclinical behavioural characteristics qualitatively similar to those presented in Autism spectrum disorders (ASDs). The BAP questionnaire (BAPQ) has been widely used to assess the BAP both in relatives of ASD people and within the general population. The current study presents the first Spanish version of the BAPQ (BAPQ-SP) and analyses its psychometric properties, including validity evidences based on the BAPQ scores relationship with other variables. Our results only support the use of the Aloof and Rigid sub-scales to assess this phenotype, whereas Pragmatic Language sub-scale seems to be the main source of misfit. This research represents a first step in the study of the BAP features in the Spanish population.

Keywords Broad autism phenotype (BAP) · Spanish version of the broad autism phenotype questionnaire (BAPQ-SP) · Confirmatory factor analysis (CFA) · Validity evidences

Introduction

Autism spectrum disorder (ASD) is a neurodevelopmental disorder characterized by atypical and persistent impairments in social interaction and communication, as well as a pattern of repetitive, restricted behaviour and interests (American Psychiatric Association [APA] 2013). ASD is a highly heterogeneous set of disorders with wide variations in symptom severity, intellectual level, and functional disability (Geschwind 2009), caused by multiple possible etiological underpinnings, which remains poorly understood (Betancur 2011; Bruining et al. 2010; Jeste and Geschwind 2014; Lenroot and Yeung 2013). As a consequence, the importance of accurate diagnoses has never been greater, particularly given also the growing prevalence (Fombonne 2009), and the considerable family and societal costs (Ganz 2007). Furthermore, undertaking the autism as a spectrum disorder remarks that the autistic traits could extend to levels of non-clinical significance within the general population (Landry and Chouinard 2016). In fact, if we understood autistic patterns as dimensional diagnostic criteria, they would obviously vary from subclinical to clinical levels of expression, throughout the general population (Constantino and Todd 2003; Ronald et al. 2006; Ruzich et al. 2015).

Some studies have employed the concept of broad autism phenotype (BAP), to refer to a set of subclinical characteristics qualitatively similar to those presented in ASD, but milder in their expression. The BAP was first observed by Leo Kanner and Hans Asperger in 1940s; these authors reported that parents showed behavioural features similar to those typical in their autistic offspring. More specifically, Kanner (1943) observed subtle manifestations of the autistic traits in some of his patient's parents such as an obsession with details, social awkwardness, and rigid behaviours. Later on, several studies have revealed social-communication,

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s10803-017-3438-0) contains supplementary material, which is available to authorized users.

✉ Angeles F. Estévez
mafernan@ual.es

✉ Pablo Sayans-Jiménez
psj242@ual.es

[1] Department of Psychology, University of Almeria, 04120 Almería, Spain

[2] CERNEP Research Center, University of Almeria, 04120 Almería, Spain

personality, and cognitive characteristics similar to those presented in ASD in parents, siblings, and other extended relatives of people with an ASD diagnosis (i.e., enhanced aloof, rigid, and hypersensitive personality traits; increased rates of internalizing psychiatric conditions; and, deficits in pragmatic language skills as well as in social reciprocity; for reviews, see Bailey et al. 1998; Pisula and Ziegart-Sadowska 2015; Sucksmith et al. 2011).

Most of the research on the BAP has deeply focused on genetic liability for autism by examining autistic traits in a subgroup of first and second degree relatives of ASD individuals (for a review, see Bailey et al. 1998). Recently, some authors have suggested that since ASD is diagnosed based on behavioural methods, the study of those people who exhibit subclinical autistic traits could be a useful tool for understanding the relationship between behaviours and traits which are similar to those of ASD (D'Cruz et al. 2013; Landry and Chouinard 2016). Thus, behavioural, neuropsychological and personality profiles have been explored in parents and siblings of ASD patients (Klusek et al. 2014; Losh et al. 2008, 2009; Ruzich et al. 2015).

However, innovative studies have highlighted that the BAP also exists within the general population (Baron-Cohen et al. 2001; Constantino and Todd 2003, 2005), suggesting its relevance as a research target. Actually, recent studies have already provided meaningful information for a range of characteristics related to autism which are continuously distributed within the general population (Wainer et al. 2011), including restricted and repetitive interests and hobbies (Baron-Cohen et al. 2001), atypical visuospatial and cognitive performance (Grinter et al. 2009; Stewart and Austin 2009), reduced gaze reciprocity (Chen and Yoon 2011), and abnormal speech perception (Stewart and Ota 2008) as well as recognition from faces and body language (Ingersoll 2010). It is important to note that the studies focused on this population are still scarce but highly needed given that their results could help to gain a better understanding of the developmental trajectories of autistic tendencies, as it has been suggested in Landry and Chouinard (2016) research, which confirmed the presence of traits associated with ASD in the typically developing population. But before conducting these studies, it is necessary first to develop specific tools for measuring the BAP in the general population.

Measuring BAP

Initially, the BAP was assessed through clinical interviews with the purpose of studying behavioural and genetic variability among the first-degree family members of an individual with ASD in an attempt to locate the underlying genes of the disorder (Bolton et al. 1994; Dawson et al. 2007; Piven et al. 1994). Later on, three primary self-report questionnaires have emerged for measuring BAP characteristics—the Autism Quotient (AQ), the Social Responsiveness Scale (SRS), and the Broad Autism Phenotype Questionnaire (BAPQ)—none of them have been translated or validated in Spanish-speaking populations. The *AQ* is a self-report questionnaire originally developed to assess ASD in adults with average intelligence and it is comprised of five subscales: social skills, attention switching, attention to detail, communication, and imagination (Baron-Cohen et al. 2001; Ingersoll et al. 2011). Likewise, the *SRS* is a questionnaire which assesses ASD as a single trait—social reciprocity—that is made up of five subscales: social awareness, social cognition, social communication, social motivation, and repetitive and restrictive behaviour (Constantino and Gruber 2007). Finally, the *BAPQ* was designed to apprehend the key features of the BAP in adults comprising three dimensions: Aloof, Rigid, and Pragmatic Language (Hurley et al. 2007).

Although many studies have found that the AQ, the SRS, and the BAPQ are suitable for measuring this phenotype in relatives of people with an ASD diagnosis, the BAPQ has been suggested as the best instrument to identify BAP individuals since it was the only questionnaire specifically developed for this purpose (Hurley et al. 2007; Wainer et al. 2011). Its items were derived from the modified personality assessment schedule—revised (MPAS-R) and the pragmatic rating scale (PRS) which had previously been used in studies exploring the BAP (Hurley et al. 2007; Piven et al. 1997a, b). The questionnaire is divided into three subscales referred to each of the three dimensions. According to Hurley et al. (2007, p. 1681) aloof personality refers to a limited interest in or enjoyment of social interaction; rigid personality relates to resistance, and/or difficulty adapting to change; and, pragmatic language problems comprises deficits in the social use of language that lead to difficulties with effective communication and/or conversational reciprocity.

Overall, the BAPQ scores have shown good reliability and validity (see Hurley et al. 2007; Ingersoll et al. 2011; Sasson et al. 2013a). Several studies have provided validity evidences supporting the inferences made on BAPQ scores mainly in terms of differential scores among groups (e.g., Hurley et al. 2007; Meera et al. 2015; Sasson et al. 2013a; Shi et al. 2015). BAPQ scores have also supportive validity evidences based on their relations with other variables. In fact, this questionnaire is well-correlated with the AQ and the SRS (Ingersoll et al. 2011; Sasson et al. 2013b), and additional criterion evidences have been found both in a non-clinical sample of college student (Ingersoll et al. 2011; Wainer et al. 2011) and within the general population (Sasson et al. 2013b). These evidences showed, for example, that subthreshold autism-social characteristics within the general population, measured by using the BAPQ, were connected to poorer social skill and social-cognitive ability (Sasson et al. 2013b). Likewise, these characteristics were related to a poor performance on an

aggregate of social-cognitive aspects such as face recognition, affect processing, and theory of mind abilities.

However, not all the validity evidences clearly support the inferences made on BAPQ scores. Validity evidences based on the internal structure are probably one of the BAPQ features that need more attention in future research. When factor analysis is performed the main goal is to achieve a simple structure (i.e., factors should be defined by indicators loading highly on them and cross-loadings or factor overlapping should be avoided; see Thurstone 1947) that enhances the interpretation of the scores, even more, when coarse factor scores (averaging or summing the unweighted raw of indicators with salient loadings on each factor) are used to determine Aloof, Rigid, and Pragmatic Language. Nevertheless, BAPQ factor structure has shown comparative fit index (CFI) indicators far from acceptable limits (e.g., Broderick et al. 2015). Results have also shown no evidence of a simple factor structure, sometimes because there are several items with cross-loadings or without loadings at least in one factor (e.g., Sasson et al. 2013a) and, other times, because factor loadings are not reported (e.g., Broderick et al. 2015; Ingersoll et al. 2011; Wainer et al. 2011).

Regardless of these internal structure problems, the BAPQ appears to be the best questionnaire to explore the BAP in the general population, as mentioned above (Ingersoll et al. 2011). For this reason and given that none of the three questionnaires is available for Spanish speakers, the purpose of this study is to present a new Spanish self-report version of the BAPQ. To do this, we have translated and adapted the BAPQ to Spanish speakers and its psychometric properties have been tested in a non-clinical sample of college students. In this study, we specifically emphasize the information related to the internal structure. If the reported lack of fit related to the internal structure is confirmed here, we will provide information regarding its causes and possible solutions with the aim of finding a measurement tool with empirically supported inferences based on the BAPQ specific dimensions.

In the present study we will perform confirmatory factor analysis (CFA) to test BAPQ simplest factor structure (three factors with only primary loadings). We will also analyse and describe global and local fit, with special attention to modification indices in order to provide future directions to improve this questionnaire. It is important to note that Sasson et al. (2013a) found that when the self-report version of the BAPQ was applied in a sample of parents of people diagnosed with ASD, five out of seven salient cross-loadings were caused by the items of the Pragmatic Language dimension (BAPQ items 2, 7 11, 29, and 34). Based on this study, our main hypothesis is that one of the possible reasons underlying the lack of fit of BAPQ internal structure might be due to the malfunctioning of Pragmatic Language items.

On the other hand, once we have identified an acceptable simple factor structure, validity evidences based on the BAPQ scores relationship with other variables will be tested. Specifically, the *NEO Five-Factor Inventory* (NEO-FFI; Costa and McCrae 1992) dimensions of extroversion, openness, and agreeableness and the three subscales of the Pragmatic Awareness Questionnaire (PAQ; Rodríguez-Muñoz 2012) will be regressed on BAPQ dimensions. NEO-FFI dimensions (NEO Five-Factor Inventory is the short version of the NEO-PI-R) were selected based on previous studies with parents with ASD children that found a negative correlation between Aloof (or social features measured by the MPAS or the MPAS-R) and the Extraversion factor measured by the NEO-PI-R, and between Rigidity (or rigid personality) and the Openness and Agreeableness dimensions (Costa and McCrae 1995, as cited in Amaral et al. 2011, p. 463; Piven et al. 1997a, b). Later on, two studies exploring the relationship between autistic traits and personality in university students by using the AQ and the Personality Mini-Markers (Austin 2005) or the NEO-PI-R (Wakabayashi et al. 2006), have also found that AQ score was negatively correlated with Extraversion and Agreeableness. Taking into account these findings as well as the definition of the three dimensions assessed by the BAPQ (Aloof personality, Rigid personality and Pragmatic language problems; Hurley et al. 2007, p. 1681) we hypothesize that Aloof personality will have a high negative relationship with Extraversion and a moderate negative relationship with Agreeableness, whereas Rigid personality will be weakly negatively related to Openness. The PAQ (Rodríguez-Muñoz 2012) will be used to test the convergent validity of the Pragmatic Language scores. In this case, moderate/high negative relationships among the PAQ sub-scales and the Pragmatic Language dimension are expected whereas Aloof and Rigid dimensions should show low negative relations with the PAQ sub-scales.

Method

Participants

Three hundred and forty-nine undergraduates participated in this study. Gender composition was 26.1% of men, ranging in age from 18 to 52 years, M (SD) = 21.56 (4.63), and 73.9% of women ranging in age from 18 to 47 years, M (SD) = 20.88 (4.01). Written informed consent was obtained from all participants. They all were informed with respect to the purposes of the study and none of them had previous experience with any of the questionnaires applied (BAPQ, PAQ, and NEO-FFI). Incidental and snowball sampling was performed. The inclusion criteria were to have no history of autism or any related developmental disorders in first-degree relatives. Those participants who did not meet the inclusion

criteria, have severe medical or genetic conditions, or history of psychiatric disorders (e.g., depression, schizophrenia…) were excluded. This study received ethics approval by the local Human Research Committee and was conducted in accordance with the approved guidelines and the Declaration of Helsinki.

With the 12 items per BAPQ factor this sample size meets the requirements suggested by MacCallum et al. (1999) to obtain stable factor solutions. Additionally, a priori power analysis was performed to determine what would be the minimum sample size to detect very small effect sizes ($R^2 = .05$) in a linear multiple regression with three predictors using Gpower software (Faul et al. 2009). With an alpha = .05 and power = .80 the required sample size to detect the mentioned effect size was n = 253. Nevertheless, we recruited a bigger sample because this study is framed in a larger research.

Instruments/Measures

Broad Autism Phenotype Questionnaire, Spanish Version (BAPQ-SP)

The BAPQ (Hurley et al. 2007) is a screening instrument for the assessment of the personality and language characteristics identified as encompassing the BAP in parents of individuals with ASD (Ingersoll et al. 2011). The BAPQ includes 36 items grouped into three subscales (aloof, rigid, and pragmatic language) and offers both a score for each subscale and an overall score, which reflects the average of all items (Hurley et al. 2007). Each of the subscales is comprised of 12 of the total 36 items and it is related to each of the BAP traits proposed by Hurley et al.: aloof personality, rigid personality, and pragmatic language difficulties. The BAPQ items demand participants to rate how frequently each statement (e.g., "I like being around other people") applies to them on a scale from one (statement applies very rarely) to six (statement applies very often). This scale (1–6) does not offer neutral option as a response (Hurley et al. 2007). Fifteen of the total 36 items become reverse scored items (1, 3, 7, 9, 12, 15, 16, 19, 21, 23, 25, 28, 30, 34, and 36). There are two versions of the BAPQ (Hurley et al. 2007); in this study we only used the self-report version.

BAPQ adaptation to Spanish language was performed following the guidelines proposed by Muñiz et al. (2013) and the International Test Commission Guidelines on Adapting a Test (http://www.intestcom.org). The English version of the BAPQ was translated into Spanish by an official translator. Simultaneously, the BAPQ was also translated into Spanish by one of the researchers of the present study. Then, the concordance of each item from the two translations was analysed. In this comparison, we discussed the equivalence of the cultural aspects mentioned in the BAPQ to the Spanish population, their grammatical concordance, and the quality of the wording of the items of the questionnaire. Subsequently, a second researcher translated the Spanish version back to English. Both researchers involved in the translation of the BAPQ were Spanish native speakers. After comparing the back-translation version to the original text and discussing discrepancies in the retranslation with the BAPQ English version, a final version was established. As soon as the final version was obtained, a mixed panel conformed by three experts (one for each of the domains in the questionnaire—namely, Aloof personality, Rigid personality, and Pragmatic Language—being one of them also an expert in psychometrics applied to transcultural contexts, and trained in Educational settings), valued the adjustment of each of the items to their corresponding domains and the sense of the relation, that is, if they have a positive or a negative relation to the domain. The administered BAPQ-SP can be seen in Online Appendix 1.

Pragmatic Awareness Questionnaire (PAQ; Rodríguez-Muñoz 2012)

This questionnaire is a self-report measure of pragmatic language in Spanish population. It is composed of 26 items that are answered using ordinal scores following a 5-point Likert scale where: *1 is very bad, 2 is bad, 3 is regular, 4 is good and 5 is very good*. For its elaboration, the PAQ (Rodríguez-Muñoz 2012) was based mainly on the Prutting and Kittchner (1987) pragmatic protocol. PAQ items are organized into three pragmatic factors: enunciative (Cronbach α in this study .64), textual (Cronbach α in this study .81), and interactive (Cronbach α in this study .88); a score is given for each factor and an overall composite score can also be obtained. None of the items requires reversed scoring. The higher the score of each of the three factors, the better the pragmatic language is.

The *NEO Five-Factor Inventory* (NEO-FFI, Spanish Version; Costa and McCrae 1999)

This inventory is a short well-established and widely used measure of the big five model of personality proposed by Costa and McCrae (1989). It includes 60 items that yield an overall score composed of five scales: neuroticism, extraversion, openness, agreeableness, and conscientiousness. Each of the scales is comprised of 12 of the total 60 items (24 are reverse scored). Participants have to rate each item or statement according to their character on a nominal scale from A (totally disagree) to E (totally agree). Only the extraversion (Cronbach α in this study .90), agreeableness (Cronbach α in this study .74), and openness (Cronbach α in this study .77) dimensions were considered in the present study due to their theoretical relation to the aloof and rigid BAPQ factors.

Higher scores imply higher levels of extraversion, agreeableness, and openness.

Procedure

Trained staff administered, in groups of 30 participants, a booklet containing a set of six self-report questionnaires: the three just described in the instruments section (BAPQ, PAQ, and NEO-FFI), and three additional questionnaires out of the scope of this study. Participants answered the survey in a quiet room free from distraction. There was no time limit (the approximate time necessary to complete the questionnaires was 40–50 min) and each participant received course credits in return. The trained staff confirmed that all the participants were participating voluntarily and informed them that their answers would be handled only for scientific purposes. Participants were also notified they could stop their participation at any time and that their anonymity and confidentiality was guaranteed. Furthermore, all of them signed an informed consent.

Two booklet versions (model A and model B) were made by changing the order of two of the three additional questionnaires although the BAPQ was always the first in filling and the PAQ always preceded the NEO-FFI. Participants received randomly one of the two models. First, participants read a letter from the researchers endorsing the aims of the research and the informed consent. Following signing the consent, the instructions to complete the whole survey were explained. No questions were allowed to be asked to the researchers. After finishing the set of questionnaires, participants were asked to check that they had completed all the questions and that they had answered them appropriately and honestly. Finally, some socio-demographic variables (name, age, gender and educational level—Degree, Master degree or PhD-) were registered.

Data Analysis

The descriptive statistics of all BAPQ indicators as well as the descriptive statistics of PAQ and NEO-FFI total scores were analyzed. Outlier detection was done using Mahalanobis distances. According to Hair et al. (2009), an observation can be considered as an outlier if the coefficient D^2/df exceed the value of 3 or 4 in big samples. CFA and structural equations models were performed to test the BAPQ internal structure and its relationships with PAQ and NEO-FFI scales. Parameter estimations were performed using the maximum likelihood. Missing data was treated with Full Information Maximum Likelihood. We assigned the latent factors metric by fixing the mean to .00 and variance at 1.00 for all the latent variables. All the correlations between factors were freed and all the cross-loadings were fixed to 0.

Fit to the models was checked using the Chi square test, the Comparative Fit Index (CFI), and the Root Mean Square Error Approximation (RMSEA; the confidence interval was set at 90%). We consider good fit when RMSEA $\leq .05$ or CFI $\geq .97$ (Schermelleh-Engel et al. 2003); adequate fit when RMSEA is close to .06 or CFI $\geq .95$ (Hu and Bentler 1999). RMSEA values between .08 and .10 and CFI values between .95 and .90 can be considered as an acceptable fit. When RMSEA is higher than or equal to .10 and CFI values are lower than .90 the model should be discarded (Brown 2006). The 95% confidence intervals were estimated using MPlus (based on the application of the delta method; Raykov and Marcoulides 2004). Analyses were performed using SPSS v19.0 and MPlus v7.0.

Validity Evidences Based on the Internal Structure

We tested a first order factor model with three dimensions, each one composed by the Aloof, Rigid, and Pragmatic Language items respectively (BAPQ-Model; the data set contained 315 complete cases). Omega coefficients were estimated to check the score reliability on each dimension (McDonald 1999). Adequate omega values should be higher than .80 (Raykov and Marcoulides 2011). Additionally, due to the misfit of the tested model and based on Sasson et al.'s (2013a) results, we decided to test a new first order factor model including only the Aloof and Rigid dimensions (AlRig-Model; the data set contained 329 complete cases). With this model we try to test whether the only psychometric problem of the BAPQ questionnaire are the Pragmatic Language items and dimension or, on the contrary, Aloof and Rigid items and dimensions also show severe malfunctioning. Modification indices higher that ten were analysed in both models in order to provide general guidelines to improve BAPQ internal structure.

Validity Evidences Based on the Relationships with Measures of Other Variables

Two structural models were created to test the relationships between BAPQ dimensions and PAQ and NEO-FFI scales. Extraversion, agreeableness, openness, enunciative, textual, and interactive dimensions were regressed on BAPQ factors of BAPQ-Model and AlRig-Model (BAPQ-StrModel and AlRig-StrModel models respectively). The dataset in the BAPQ-StrModel contained 299 complete cases whereas in AlRig-StrModel the data set contained 311 complete cases.

In order to discard any potential effect of the gender variable we have included measurement and structural invariance tests among genders for each measurement and structural model. Gender invariance will be assumed if CFI differences between models are lower than .01 (Cheung and Rensvold 2002).

Results

The descriptive statistics for all the variables can be seen in Table 1. The sample means of the present study were lower than those observed in the PAQ original sample, and similar to those found in the NEO-FFI. No outliers were identified according to Hair et al.'s proposed cut-off (2009). No variables showed extreme kurtosis or skewness.

Validity Evidences Based on the Internal Structure

Fit indicators for BAPQ-Model were: χ^2 (591) = 1460.22, p < .001, the RMSEA was .07 [.06, .07] and the CFI was .74. This measurement model showed extremely low CFI values, suggesting that it should be discarded. Measurement invariance among genders was confirmed for this model, ΔCFI = .002. Factor loadings and modification indices can be seen in Tables 2 and 3. Factor loadings in absolute values were moderate-high *M* = .62, range = [.31, .80] for Aloof dimension, moderate *M* = .50, range = [.29, .72] for the Rigid dimension, and moderate-low *M* = .31, range = [.03, .73] for the Pragmatic Language dimension. Four out of 12 items of Pragmatic Language showed non-significant loadings. Omega reliability estimations [95% CI] were .89 [.87, .90] for Aloof factor, .79 [.77, .82] for Rigid factor, and .54 [.48, .59] for Pragmatic Language. Cronbach's alfa reliability estimations were .85 for Aloof factor, .79 for Rigid factor, and .62 for Aloof factor. All correlations among factors were statistically significant (p < .001): − .44 for Aloof-Rigid, − .80 for Aloof-P. Language, and .47 for Rigid-P. Language. Given that the Aloof factor includes mostly reverse scored items and that none of the items were inverted in this study, the direction of the Aloof relationships is reversed (i.e., a high score in this dimension indicates lack of aloofness or high interest in or enjoyment of social interaction; this is applicable henceforth). Subscale intercorrelations were comparable to the BAPQ original intercorrelations found in the control group (Hurley et al. 2007) although Aloof-P. Language correlation was bigger than expected. All these results make inadvisable to use Pragmatic Language observed scores to represent or infer this construct.

Additionally, when local fit related to factor loadings was explored, six out of 12 items of Pragmatic Language showed modification indices higher than 10, whereas only two items of Aloof and Rigid, respectively, showed local misfit. Finally, modification indices related to potential correlations between item uniqueness pointed again to Pragmatic Language items as the main contributors to the global misfit of the BAPQ measurement model (i.e., BAPQ-Model).

Table 1 Descriptive statistics for all the BAPQ items and related variables

	M(SD)	SK	K
Item 1	5.01 (0.99)	−0.96	0.58
Item 2	2.73 (1.26)	0.75	−0.02
Item 3	3.05 (1.19)	0.40	−0.28
Item 4	2.00 (1.21)	1.58	2.18
Item 5	2.21 (1.17)	1.09	0.92
Item 6	2.43 (1.16)	0.80	0.44
Item 7	4.38 (1.07)	−0.40	−0.11
Item 8	2.04 (1.19)	1.25	1.10
Item 9	4.82 (1.13)	−0.94	0.68
Item 10	2.25 (1.22)	1.08	0.78
Item 11	2.08 (1.02)	1.03	1.20
Item 12	4.13 (1.29)	−0.42	−0.62
Item 13	2.08 (1.01)	0.95	0.78
Item 14	2.33 (1.13)	1.15	1.42
Item 15	4.02 (1.09)	−0.02	−0.73
Item 16	4.21 (1.34)	−0.34	−0.74
Item 17	2.75 (1.36)	0.79	−0.07
Item 18	2.26 (0.99)	0.79	0.75
Item 19	4.88 (1.09)	−0.65	−0.45
Item 20	3.23 (1.46)	0.25	−0.84
Item 21	4.54 (1.10)	−0.56	0.10
Item 22	2.40 (1.15)	1.08	1.18
Item 23	2.13 (1.00)	0.84	0.62
Item 24	2.44 (1.15)	0.74	0.19
Item 25	4.42 (1.21)	−0.47	−0.53
Item 26	2.27 (1.27)	1.19	1.07
Item 27	1.82 (1.05)	1.40	1.71
Item 28	4.70 (1.16)	−0.63	−0.27
Item 29	2.02 (0.94)	1.35	3.04
Item 30	3.10 (1.09)	0.48	0.20
Item 31	2.28 (1.17)	0.83	0.46
Item 32	2.96 (1.35)	0.43	−0.57
Item 33	3.38 (1.31)	0.32	−0.53
Item 34	4.16 (1.05)	−0.25	−0.45
Item 35	2.97 (1.14)	0.58	0.13
Item 36	4.81 (1.13)	−0.81	0.17
Extraversion	3.74 (0.72)	−0.66	0.01
Agreeableness	3.53 (0.52)	−0.43	0.22
Openness	3.64 (0.59)	−0.15	−0.52
Enunciative P.	4.05 (0.58)	−0.39	−0.17
Textual P.	4.02 (0.60)	−0.40	−0.21
Interactive P.	3.95 (0.51)	−0.14	0.14

M mean, *SD* standard deviation, *Sk* skewness, *K* kurtosis

Accordingly, we ran an alternative measurement model in which Pragmatic Language items were excluded. Fit indicators for AlRig-Model were: χ^2 (591) = 579.78, p < .001, the RMSEA was .06 [.06, .07] and the CFI was

Table 2 Factor loading, modification indices, and expected parameter change for factor loadings in BAPQ-Model

		λ	Modification indices	Standardised expected parameter change	Factor in which the item should load
Aloof	Item 1	.72	10.31	.31	P. Language
	Item 5	−.53			
	Item 9	.76			
	Item 12	.64			
	Item 16	.70	10.00	.31	P. Language
	Item 18	−.51			
	Item 23	−.31			
	Item 25	.69			
	Item 27	−.63			
	Item 28	.51			
	Item 31	−.60			
	Item 36	.80			
Rigid	Item 3	−.50			
	Item 6	.52			
	Item 8	.55	11.00	.21	P. Language
	Item 13	.59			
	Item 15	−.41			
	Item 19	−.58	10.39	.18	Aloof
	Item 22	.72			
	Item 24	.56			
	Item 26	.31			
	Item 30	−.58			
	Item 33	.29			
	Item 35	.35			
Pragmatic Language	Item 2	.50	10.24	.39	Aloof
	Item 4	.28			
	Item 7	−.49	45.53	.81	Aloof
	Item 10	*.13*			
	Item 11	.76			
	Item 14	.43	23.59	.58	Aloof
	Item 17	−.06	15.89	.49	Aloof
	Item 20	*.17*	14.05 13.22	.46 .25	Aloof Rigid
	Item 21	−*.07*			
	Item 29	.54			
	Item 32	*.03*			
	Item 34	−.29			

Non statistically significant factor loadings (λ) were italic, the rest were statistically significant ($p < .001$)

.87. Fit indicators for this model are significantly better than BAPQ-Model indicators although model fit indices cannot be compared directly. Measurement invariance among genders was confirmed also for this model, $\Delta CFI = .000$. On the other hand, the examination of AlRig-Model factor loadings and modification indices is the best way to test if the internal structure problems caused by Pragmatic Language have been now alleviated. Factor loadings and modification indices can be seen in Tables 4 and 5. Factor loadings in absolute values were moderate-high $M = .61$, range = [.30, .81] for Aloof factor and moderate $M = .50$, range = [.29, .72] for the Rigid factor. In the same way that factor loadings, Aloof and Rigid omega estimations remained stable (when compared to BAPQ-Model): Omega estimations [95% CI] were .89 [.87, .90] for Aloof dimension and .79 [.77, .82] for Rigid dimension. Correlation between Aloof and Rigid dimensions was −.44 ($p < .001$). These results support the use of Aloof and Rigid scores for the proposed goals; in this

Table 3 Modification indices pointing potential correlations between items uniqueness in BAPQ-Model

Pair of items	Modification indices	Standardised expected parameter change	BAPQ factors potentially affected
Item 34–Item 21	62.08	.43	P. Language
Item 32–Item 17	52.19	.39	P. Language
Item 20–Item 14	40.06	.35	P. Language
Item 20–Item 17	36.84	.33	P. Language
Item 17–Item 26	25.36	.27	P. Language and Rigid
Item 14–Item 2	23.73	.28	P. Language
Item 19–Item 16	22.47	.28	Rigid and Aloof
Item 32–Item 14	20.93	.25	P. Language
Item 32–Item 20	20.34	.24	P. Language
Item 17–Item 14	20.27	.25	P. Language
Item 32–Item 26	19.85	.24	P. Language and Rigid
Item 19–Item 6	17.83	−.25	Rigid
Item 9–Item 1	16.23	.26	Aloof
Item 13–Item 23	15.96	.23	Aloof and Rigid
Item 22–Item 19	13.48	.25	Rigid
Item 20–Item 7	12.98	.20	P. Language
Item 34–Item 33	12.82	.20	P. Language and Rigid
Item 29–Item 10	12.55	.20	P. Language
Item 33–Item 19	12.53	.20	Rigid
Item 14–Item 26	11.86	.19	P. Language and Rigid
Item 18–Item 9	11.59	.20	Aloof
Item 14–Item 33	11.56	−.19	P. Language and Rigid
Item 32–Item 35	11.11	.18	P. Language and Rigid
Item 24–Item 23	10.58	.18	Rigid and Aloof
Item 33–Item 24	10.14	.18	Rigid
Item 34–Item 28	10.10	.18	P. Language and Aloof

case, only one item (item 19 of Rigid dimension) showed potential cross-loadings.

Validity Evidences Based on the Relationships with Measures of Other Variables

Fit indicators for BAPQ-StrModel were: χ^2 (789) = 1925.20, $p < .001$, the RMSEA was .06 [.06, .07] and the CFI was .77. Structural invariance among genders was confirmed for this model, ΔCFI = .003. Since the measurement model included in this model is identical to BAPQ-Model, factor loadings and relationships between uniqueness are the same. Consequently, fit problems are replicated and this model should not even be considered as acceptable. Furthermore, correlations and regression coefficients show certain inconsistent validity evidences when supporting the interpretation of BAPQ factors (see Table 6). As expected, the Pragmatic Language dimension is highly and negatively related with the PAQ subscales, and not related in a statistically significant way with extraversion, agreeableness, and openness. Additionally, Aloof dimension is highly positively related to Extraversion and moderately positively related to Agreeableness, whereas Rigid dimension show low/moderate negative relation with Openness. However, on the other hand, the relationship between Aloof dimension (a high score in this dimension should be interpreted as lack of Aloof) and the PAQ subscales are moderate and negative in this model. These relationships could be taken as slight negative validity evidences.

On the other hand, AlRig-StrModel showed almost acceptable fit estimations χ^2 (789) = 854.18, $p < .001$, the RMSEA was .06 [.06, .07] and the CFI was .86 (model fit indices cannot be compared directly with the previous model). As in the measurement model, the structural model shows adequate RMSEA estimations but CFI estimations slightly below the recommended cut-off. Structural invariance among genders was also confirmed for this model, ΔCFI = .003. Correlations, regression coefficients, and R^2 estimations of AlRig-StrModel can be seen in Table 7. This model shows relationships among BAPQ factors and the rest of variables that match with the expected results.

Table 4 Factor loading, modification indices, and expected parameter change for factor loadings in AlRig-Model

		λ	Modification indices	Standardised expected parameter change	Factor in which the item should load
Aloof	Item 1	.73			
	Item 5	−.52			
	Item 9	.76			
	Item 12	.62			
	Item 16	.72			
	Item 18	−.51			
	Item 23	−.30			
	Item 25	.67			
	Item 27	−.63			
	Item 28	.51			
	Item 31	−.59			
	Item 36	.81			
Rigid	Item 3	−.50			
	Item 6	.52			
	Item 8	.54			
	Item 13	.59			
	Item 15	−.41			
	Item 19	−.58	11.53	.19	Aloof
	Item 22	.72			
	Item 24	.56			
	Item 26	.31			
	Item 30	−.59			
	Item 33	.29			
	Item 35	.34			

All factor loadings (λ) were statistically significant ($p < .001$)

Table 5 Modification indices pointing potential correlations between items uniqueness in AlRig-Model

Pair of items	Modification indices	Standardised expected parameter change	BAPQ factors potentially affected
Item 19–Item 16	20.74	.27	Rigid and Aloof
Item 19–Item 6	18.14	−.26	Rigid
Item 13–Item 23	16.08	.23	Rigid and Aloof
Item 9–Item 1	14.07	.24	Aloof
Item 22–Item 19	13.44	.25	Rigid
Item 33–Item 19	13.44	.21	Rigid
Item 27–Item 18	10.61	.19	Aloof
Item 18–Item 9	10.54	.20	Aloof
Item 24–Item 23	10.35	.18	Rigid and Aloof

Furthermore, in this model the relationships among Aloof and the PAQ subscales also coincide the expected results and the confidence intervals are narrower in contrast to the BAPQ-StrModel.

Discussion

As far as we know, this is the first study that provides a Spanish version of a questionnaire to explore the BAP. Furthermore, this study extends previous findings on the factor structure of the BAPQ (e.g., Broderick et al. 2015; Ingersoll et al. 2011; Sasson et al. 2013a; Wainer et al. 2011) in a large sample of university students. All the obtained results were gender invariant; therefore, gender cannot be considered a variable that may influence them. Our results are in agreement with Hurley et al.'s (2007) internal structure validity evidences for the control group since they found three dimensions and moderate correlations for all the variables (namely, Aloof and Rigid r = .54, Aloof and Pragmatic Language r = .53, and Rigid and Pragmatic Language r = .51), although the correlation between Aloof and Pragmatic Language is rather higher in our study than in the BAPQ original study (Hurley et al. 2007). We consider that a possible severity disparity in the BAP traits, between control parents and university students, could also underlie this BAPQ internal structure difference (e.g., different levels of data dispersion or ceiling/floor effects).

Table 6 Regression coefficients [and 95% CI intervals] for BAPQ-StrModel above the diagonal

	Aloof	Rigid	Extraversion	Agreeableness	Openness	Enunciative Pragmatic	Textual Pragmatic	Interactive Pragmatic
Aloof			0.82 [0.70, 0.93]**	0.39 [0.21, 0.57]**	−0.31 [−0.52, −0.11]*	−0.34 [−0.55, −0.13]*	−0.50 [−0.72, −0.28]**	−0.25 [−0.45, −0.05]*
Rigid	−.58**		−0.12 [−0.19, −0.04]*	−0.01 [−0.12, 0.11]	−0.29 [−0.40, −0.17]**	−0.08 [−0.19, 0.04]	−0.10 [−0.22, 0.02]	−0.07 [−0.17, 0.04]
Pragmatic Language	−.80**	0.33**	0.05 [−0.09, 0.18]	0.06 [−0.15, 0.27]	−0.23 [−0.46, −0.01]	−0.75 [−0.97, −0.53]**	−0.88 [−1.11, −0.65]**	−0.86 [−1.06, −0.65]**
Extraversion								
Agreeableness			.00					
Openness			.01	.06				
Enunciative Pragmatic			.24**	.02	.15*			
Textual Pragmatic			.08	.06	.26**	.41**		
Interactive Pragmatic			.10	.13	.04	.45**	.51**	
R^2			.70	.12	.10	.32	.38	.50

Correlations and R^2 below the diagonal

*p < .05; **p < .001

Nevertheless, the current study evidences a bigger problem in the internal structure of the BAPQ. As we hypothesized, the severe misfit of the BAPQ-Model is caused mainly by problems in the Pragmatic Language dimension. The Pragmatic Language items are the main contributors to the global misfit of the BAPQ-Model (BAPQ items 2, 7, 14, 17, 20). In contrast, Aloof and Rigid dimensions show adequate fit with moderate factor loadings and adequate reliability estimations. Our results are similar to those found by Sasson et al.'s (2013a) who indicated that five out of seven salient cross-loadings were caused by the items of the Pragmatic Language factor (BAPQ items 2, 7, 11, 29 and 34).

The alternative measurement model, excluding the Pragmatic Language dimension, showed CFI estimations slightly below than the recommended cut-off to consider it adequate. However, this CFI estimation is clearly better than that obtained when Pragmatic Language was included in the model as well as than that reported for most of the studies where the BAPQ internal structure was analysed (e.g., Broderick et al. 2015). These results are in agreement with the idea that the main source of misfit in BAPQ dimensions is due to the Pragmatic Language scores. Moreover, we consider that the CFI misfit of the model with only Aloof and Rigid dimensions could be explained by the high ratio item/factor in each BAPQ dimension (see Kenny and McCoach 2003). These results, along with Aloof and Rigid adequate reliability estimations, support the use of their scores for the proposed goals, although fit indicators suggest that these measures could be improved.

A more detailed analysis of the correlations between item uniqueness in the BAPQ-model (see Table 5) contributes to identify the main potential source of misfit in Aloof and Rigid scales. These results indicate a possible effect of content overlapping, that is, some of the items could be also referred to related contents apart from Aloofness and Rigidity. For example, items 19, 16, and 6 are all referred to cope with new experiences, so these items could be exploring also the preference for this kind of experiences no matter which scale they belong to.

Our second goal was to explore the validity evidences of the BAPQ scores based on its relationship with measures of other variables. In general terms, results from this study indicate that each BAPQ scale is associated to a number of personality traits and language skills as we expected. As it was hypothesized, validity evidences showed a strong negative association between Pragmatic Language and Enunciative Pragmatic, Textual Pragmatic and Interactive Pragmatic from PAQ (Rodríguez-Muñoz 2012). However, as it was highlighted, we strongly advise against the use of the BAPQ Pragmatic Language factor. It has negative internal structure validity evidences and the inclusion of Pragmatic Language factor in any prediction, together with Aloof, could lead to inaccurate and unexplainable results.

On the other hand, Aloof and Rigid scores have shown empirical evidences supporting their use. In addition to their favourable internal structure validity evidences, Aloof (it should be interpreted as lack of Aloof in the models of the current study) had a strong positive correlation with Extraversion and a moderate-low positive correlation with

Table 7 Regression coefficients [and 95% CI intervals] for AlRig-StrModel above the diagonal

	Aloof	Extraversion	Agreeableness	Openness	Enunciative Pragmatic	Textual Pragmatic	Interactive Pragmatic
Aloof		0.79 [0.74, 0.83]**	0.35 [0.25, 0.44]**	−0.15 [−0.26, −0.04]*	0.19 [0.08, 0.29]**	0.12 [0.02, 0.22]	0.35 [0.26, 0.44]**
Rigid	−.58**	−0.11 [−0.18, −0.04]*	0.00 [−0.11, 0.11]	−0.33 [−0.43, −0.22]**	−0.21 [−0.31, −0.1]**	−0.25 [−0.36, −0.14]**	−0.21 [−0.31, −0.11]**
Extraversion							
Agreeableness		.01					
Openness		.00	.06				
Enunciative Pragmatic		.18**	−.00	.20**			
Textual Pragmatic		.04	.03	.30**	.58**		
Interactive Pragmatic		.06	.08	.12*	.60**	.67**	
R^2		.71	.12	.09	.11	.10	.23

Correlations and R^2 below the diagonal

*p < .05; **p < .001

Agreeableness. Taking into account this information, our second hypothesis should be inverted (because the reverse Aloof interpretation) but accepted. Indeed, the strong correlation between Extraversion and Aloofness could be explained regarding to the negative pole of the Extraversion construct, which also encompasses negative qualifiers like loneliness, self-absorption, social withdrawal, etc., theoretically strongly related to Aloofness. We also found a moderate-low negative association between Rigidity and Openness, which confirm our original hypothesis. According to the Openness definition (see Costa and McCrae 1999), and Rigidity conceptualization (e.g., Hurley et al. 2007), both constructs should be slight and negatively correlated since they have few qualifiers in common (e.g., Costa and McCrae 1999).

However, on the other hand, Aloof dimension was moderately and negatively related to the PAQ subscales. These unexpected relationships have to be analysed together with the high Aloof-P. Language correlation and the wide confidence intervals (which sizes are a consequence of standard errors magnitudes) of some of the Aloof and Pragmatic Language regression coefficients. These results could be a direct consequence of collinearity between Aloof and Pragmatic Language factors (O'brien 2007). This fact is confirmed in the AlRig-StrModel model where an adequate (i.e., positive) relationship between Aloof and PAQ factors with narrower confidence intervals is showed. The results confirmed this hypothesis indicating that the Pragmatic Language malfunctioning not also affect to the internal structure validity evidences, but also to validity evidences of Aloof and Rigid scores based on the relationship with other variables.

Collectively, these results indicate that more research is needed to improve the psychometric properties of this Spanish version of the BAPQ. Nevertheless, Aloof and Rigid sub-scales can be used with certain guarantees. Our study supports the Sasson et al.'s (2013) research, which already shows problems in the internal structure of the BAPQ in a sample of parents with ASD children (i.e., they do not achieve a simple factor structure). Consequently, we consider that the psychometric problems we have detected in the questionnaire are not directly derived from the language adaptation process. To sum up, in this study we have presented the first Spanish version of the BAPQ. This test has shown similar psychometric properties than other languages BAPQ versions although our results only support the use of Aloof and Rigid dimensions to assess BAP. Finally, Pragmatic Language scores have shown negative validity evidences similar to those found in other versions of the BAPQ (e.g., Sasson et al. 2013a).

We would like to conclude highlighting that although Aloof and Rigid sub-scales can be used to study BAP in the Spanish population, it would be reasonable to make efforts to improve BAP measurements regardless of the language

version used. As a first step, we propose to run the item response theory (IRT) approach on BAPQ dimensions; these analyses allow placing items and persons in the same continuous. Thus, IRT would help to identify whether BAPQ items are really suitable to measure BAP in a non-clinical sample (if so, items should be distributed along the same construct levels than this sample). Additionally, we think that the best way to improve BAP measurements could be to construct a new assessment tool based on items which psychometric properties have been widely proven. Based on Wainer et al. (2011) study exploring the common structure underlying AQ, SRS, and BAPQ scores, it would be reasonable to select, among these scales, the items that prove their better psychometric properties for each dimension by using IRT or exploratory factor analysis in the CFA framework (see Brown 2006). Once a stable sub-domain internal structure is achieved, validity evidences could be tested as in the present research but extending the sample to the general population (including not only college students but also people at different ages and with different educational levels) as well as to the relatives of ASD patients.

Acknowledgments This study was supported by Grant PSI2015-65248-P from the Spanish Ministry of Economy and Competitiveness, co-funded by ERDF (FEDER) funds. Authors would like to thank to Rocio Rodríguez Robles for her expert English language editing help. MGG is supported by a pre-doctoral grant by the Ministry of Education, Culture and Sport (FPU2015-01562).

Author Contributions PSJ and MGG analysed the data and wrote the manuscript. MGG and AGR were responsible for data collection. PSJ, AFE and FC were responsible for the design of the study. All authors participated in the BAPQ adaption into Spanish language, contributed to data interpretation and assisted with writing the manuscript.

Compliance with Ethical Standards

Conflict of interest All authors declare no conflict of interest.

References

Amaral, D. G., Dawson, G., & Geschwing, D. H. (2011). *Austism spectrum disorders*. New York: Oxford University Press.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*. Washington: American Psychiatric Pub.

Austin, E. J. (2005). Personality correlates of the broader autism phenotype as assessed by the Autism Spectrum Quotient (AQ). *Personality and Individual Differences, 38*, 451–460. https://doi.org/10.1016/j.paid.2004.04.022.

Bailey, A., Palferman, S., Heavey, L., & Le Couteur, A. (1998). Autism: The phenotype in relatives. *Journal of Autism and Developmental Disorders, 28*, 369–392. https://doi.org/10.1023/A:1026048320785.

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): Evidence from asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders, 31*, 5–17. https://doi.org/10.1023/A:1005653411471.

Betancur, C. (2011). Etiological heterogeneity in autism spectrum disorders: More than 100 genetic and genomic disorders and still counting. *Brain Research, 1380*, 42–77. https://doi.org/10.1016/j.brainres.2010.11.078.

Bolton, P., Macdonald, H., Pickles, A., Rios, P. A., Goode, S., Crowson, M., … Rutter, M. (1994). A case-control family history study of autism. *Journal of child Psychology and Psychiatry, 35*, 877–900. https://doi.org/10.1111/j.1469-7610.1994.tb02300.x.

Broderick, N., Wade, J. L., Meyer, J. P., Hull, M., & Reeve, R. E. (2015). Model invariance across genders of the broad autism phenotype questionnaire. *Journal of Autism and Developmental Disorders, 45*, 3133–3147. https://doi.org/10.1007/s10803-015-2472-z.

Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York: Guilford Press.

Bruining, H., De Sonneville, L., Swaab, H., De Jonge, M., Kas, M., van Engeland, H., & Vorstman, J. (2010). Dissecting the clinical heterogeneity of autism spectrum disorders through defined genotypes. *PLoS ONE, 5*, e10887. https://doi.org/10.1371/journal.pone.0010887.

Chen, F. S., & Yoon, J. M. (2011). Brief report: Broader autism phenotype predicts spontaneous reciprocity of direct gaze. *Journal of Autism and Developmental Disorders, 41*, 1131–1134. https://doi.org/10.1007/s10803-010-1136-2.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*, 233–255. https://doi.org/10.1207/s15328007sem0902_5.

Constantino, J. N., & Gruber, C. P. (2007). *Social responsiveness scale (SRS)*. Los Angeles, CA: Western Psychological Services.

Constantino, J. N., & Todd, R. D. (2003). Autistic traits in the general population: A twin study. *Archives of General Psychiatry, 60*, 524–530. https://doi.org/10.1001/archpsyc.60.5.524.

Constantino, J. N., & Todd, R. D. (2005). Intergenerational transmission of subthreshold autistic traits in the general population. *Biological Psychiatry, 57*, 655–660. https://doi.org/10.1016/j.biopsych.2004.12.014.

Costa, P. T., & McCrae, R. R. (1989). *The NEO-PI/NEO-FFI manual supplement*. Odessa, FL: Psychological Assessment Resources.

Costa, P. T., & McCrae, R. R. (1992). Normal personality assessment in clinical practice: The NEO Personality Inventory. *Psychological Assessment, 4*, 5. https://doi.org/10.1037/1040-3590.4.1.5.

Costa, P. T., & McCrae, R. R. (1995). Persons, place and personality: Career assessment using the revised NEO Personality Inventory. *Journal of Career Assessment, 3*(2), 123–139.

Costa, P. T., & McCrae, R. R. (1999). *Inventario de Personalidad NEO revisado (NEO PI-R) e Inventario NEO reducido de cinco factores (NEO-FFI) [NEO PI-R, Revised NEO Personality Inventory and NEO Five-Factor Inventory (NEO-FFI)]*. Madrid: TEA Editions.

D'Cruz, A. M., Ragozzino, M. E., Mosconi, M. W., Shrestha, S., Cook, E. H., & Sweeney, J. A. (2013). Reduced behavioral flexibility in autism spectrum disorders. *Neuropsychology, 27*, 152. https://doi.org/10.1037/a0031721.

Dawson, G., Estes, A., Munson, J., Schellenberg, G., Bernier, R., & Abbott, R. (2007). Quantitative assessment of autism symptom-related traits in probands and parents: Broader phenotype autism symptom scale. *Journal of Autism and Developmental Disorders, 37*(3), 523–536. https://doi.org/10.1007/s10803-006-0182-2.

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods, 41*, 1149–1160. https://doi.org/10.3758/brm.41.4.1149.

Fombonne, E. (2009). Epidemiology of pervasive developmental disorders. *Pediatric Research, 65*, 591–598. https://doi.org/10.1203/PDR.0b013e31819e7203.

Ganz, J. B. (2007). Classroom structuring methods and strategies for children and youth with autism spectrum disorders. *Exceptionality, 15*, 249–260. https://doi.org/10.1080/09362830701655816.

Geschwind, D. H. (2009). Advances in autism. *Annual Review of Medicine, 60*, 367–380. https://doi.org/10.1146/annurev.med.60.053107.121225.

Grinter, E. J., Van Beek, P. L., Maybery, M. T., & Badcock, D. R. (2009). Brief report: Visuospatial analysis and self-rated autistic-like traits. *Journal of Autism and Developmental Disorders, 39*, 670–677. https://doi.org/10.1007/s10803-008-0658-3.

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2009). *Multivariate data analysis: A global perspective* (7th ed.). Upper Saddle River: Prentice Hall.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*, 1–55. https://doi.org/10.1080/10705519909540118.

Hurley, R. S., Losh, M., Parlier, M., Reznick, J. S., & Piven, J. (2007). The broad autism phenotype questionnaire. *Journal of Autism and Developmental Disorders, 37*, 1679–1690. https://doi.org/10.1007/s10803-006-0299-3.

Ingersoll, B. (2010). Broader autism phenotype and nonverbal sensitivity: Evidence for an association in the general population. *Journal of Autism and Developmental Disorders, 40*, 590–598. https://doi.org/10.1007/s10803-009-0907-0.

Ingersoll, B., Hopwood, C. J., Wainer, A., & Donnellan, M. B. (2011). A comparison of three self-report measures of the broader autism phenotype in a non-clinical sample. *Journal of Autism and Developmental Disorders, 41*, 1646–1657. https://doi.org/10.1007/s10803-011-1192-2.

Jeste, S. S., & Geschwind, D. H. (2014). Disentangling the heterogeneity of autism spectrum disorder through genetic findings. *Nature Reviews Neurology, 10*, 74–81. https://doi.org/10.1038/nrneurol.2013.278.

Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous Child, 2*, 217–250.

Kenny, D. A., & McCoach, D. B. (2003). Effect of the number of variables on measures of fit in structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal, 10*(3), 333–351. https://doi.org/10.1207/s15328007sem1003_1.

Klusek, J., Losh, M., & Martin, G. E. (2014). Sex differences and within-family associations in the broad autism phenotype. *Autism, 18*, 106–116. https://doi.org/10.1177/1362361312464529.

Landry, O., & Chouinard, P. A. (2016). Why we should study the broader autism phenotype in typically developing populations. *Journal of Cognition and Development, 17*, 584–595. https://doi.org/10.1080/15248372.2016.1200046.

Lenroot, R. K., & Yeung, P. K. (2013). Heterogeneity within autism spectrum disorders: What have we learned from neuroimaging studies? *Frontiers in Human Neuroscience*. https://doi.org/10.3389/fnhum.2013.00733.

Losh, M., Adolphs, R., Poe, M. D., Couture, S., Penn, D., Baranek, G. T., & Piven, J. (2009). Neuropsychological profile of autism and the broad autism phenotype. *Archives of General Psychiatry, 66*, 518–526. https://doi.org/10.1001/archgenpsychiatry.2009.34.

Losh, M., Childress, D., Lam, K., & Piven, J. (2008). Defining key features of the broad autism phenotype: A comparison across parents of multiple-and single-incidence autism families. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics, 147*, 424–433. https://doi.org/10.1002/ajmg.b.30612.

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods, 4*(1), 84–99. https://doi.org/10.1037/1082-989x.4.1.84.

McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah, NJ: Erlbaum.

Meera, S. S., Girimaji, S. C., Seshadri, S. P., Philip, M., Shivashankar, N., Morgan, P., & Piven, J. (2015). Translation of the Broad Autism Phenotype Questionnaire to an Indian language: A description of the process. *Asian Journal of Psychiatry, 15*, 62–67. https://doi.org/10.1016/j.ajp.2015.04.013.

Muñiz, J., Elosua, P., & Hambleton, R. K. (2013). Guidelines for test translation and adaptation: Second edition [Directrices para la traducción y adaptación de los tests: Segunda edición]. *Psicothema, 25*, 151–157.

Muñoz, F. J. R. (2012). La conciencia pragmática de adultos con síndrome de Asperger. Revista de Logopedia. *Foniatría y Audiología, 32*, 21–31. https://doi.org/10.1016/j.rlfa.2012.01.002.

O'brien, R. M. (2007). A caution regarding rules of thumb for variance inflation factors. *Quality & Quantity, 41*, 673–690. https://doi.org/10.1007/s11135-006-9018-6.

Pisula, E., & Ziegart-Sadowska, K. (2015). Social communication and language deficits in parents and siblings of children with ASD—A short review. Autism. Rijeka: InTech. https://doi.org/10.5772/59134.

Piven, J., Palmer, P., Jacobi, D., Childress, D., & Arndt, S. (1997a). Broader autism phenotype: Evidence from a family history study of multiple-incidence autism families. *American Journal of Psychiatry, 154*, 185–190. https://doi.org/10.1176/ajp.154.2.185.

Piven, J., Palmer, P., Landa, R., Santagelo, S., Jacobi, D., & Childress, D. (1997b). Personality and language characteristics in parents from multiple-incidence autism families. *American Journal of Medical Genetics, 74*, 398–411. https://doi.org/10.1002/(sici)1096-8628(19970725)74:4<398::aid-ajmg11>3.0.co;2-d

Piven, J., Wzorek, M., Landa, R., Lainhart, J., Bolton, P., Chase, G. A., & Folstein, S. (1994). Personality characteristics of the parents of autistic individuals. *Psychological Medicine, 24*, 783–795. https://doi.org/10.1017/S0033291700027938.

Prutting, C. A., & Kittchner, D. M. (1987). A clinical appraisal of the pragmatic aspects of language. *Journal of Speech and Hearing Disorders, 52*, 105–119. https://doi.org/10.1044/jshd.5202.105.

Raykov, T., & Marcoulides, G. A. (2004). Using the delta method for approximate interval estimation of parameter functions in SEM. *Structural Equation Modeling: A Multidisciplinary Journal, 11*, 621–637. https://doi.org/10.1207/s15328007sem1104_7.

Raykov, T., & Marcoulides, G. A. (2011). *Introduction to psychometric theory*. New York: Routledge.

Ronald, A., Happe, F., Bolton, P., Butcher, L. M., Price, T. S., Wheelwright, S., … Plomin, R. (2006). Genetic heterogeneity between the three components of the autism spectrum: A twin study. *Journal of the American Academy of Child & Adolescent Psychiatry, 45*, 691–699. https://doi.org/10.1097/01.chi.0000215325.13058.9d.

Ruzich, E., Allison, C., Smith, P., Watson, P., Auyeung, B., Ring, H., & Baron-Cohen, S. (2015). Measuring autistic traits in the general population: A systematic review of the Autism-Spectrum Quotient (AQ) in a nonclinical population sample of 6,900 typical adult males and females. *Molecular Autism, 6*, 2. https://doi.org/10.1186/2040-2392-6-2.

Sasson, N. J., Lam, K. S. L., Childress, D., Parlier, M., Daniels, J. L., & Piven, J. (2013a). The broad autism phenotype questionnaire: Prevalence and diagnostic classification. *Autism Research, 6*, 134–143. https://doi.org/10.1002/aur.1272.

Sasson, N. J., Nowlin, R. B., & Pinkham, A. E. (2013b). Social cognition, social skill, and the broad autism phenotype. *Autism, 17*, 655–667. https://doi.org/10.1177/1362361312455704.

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online, 8*, 23–74.

Shi, L.-J., Ou, J.-J., Gong, J.-B., Wang, S.-H., Zhou, Y.-Y., Zhu, F.-R., … Luo, X.-R. (2015). Broad autism phenotype features of Chinese

parents with autistic children and their associations with severity of social impairment in probands. *BMC Psychiatry, 15*, 168. https://doi.org/10.1186/s12888-015-0568-9.

Stewart, M. E., & Austin, E. J. (2009). The structure of the Autism-Spectrum Quotient (AQ): Evidence from a student sample in Scotland. *Personality and Individual Differences, 47*, 224–228. https://doi.org/10.1016/j.paid.2009.03.004.

Stewart, M. E., & Ota, M. (2008). Lexical effects on speech perception in individuals with "autistic" traits. *Cognition, 109*, 157–162. https://doi.org/10.1016/j.cognition.2008.07.010.

Sucksmith, E., Roth, I., & Hoekstra, R. A. (2011). Autistic traits below the clinical threshold: Re-examining the broader autism phenotype in the 21st century. *Neuropsychology Review, 21*, 360–389. https://doi.org/10.1007/s11065-011-9183-9.

Thurstone, L. L. (1947). Factorial analysis of body measurements. *American Journal of Physical Anthropology, 5*, 15–28. https://doi.org/10.1002/ajpa.1330050103.

Wainer, A. L., Ingersoll, B. R., & Hopwood, C. J. (2011). The structure and nature of the broader autism phenotype in a non-clinical sample. *Journal of Psychopathology and Behavioral Assessment, 33*, 459–469. https://doi.org/10.1007/s10862-011-9259-0.

Wakabayashi, A., Baron-Cohen, S., & Wheelwright, S. (2006). Are autistic traits an independent personality dimension? A study of the Autism-Spectrum Quotient (AQ) and the NEO-PI-R. *Personality and Individual Differences, 41*, 873–883. https://doi.org/10.1016/j.paid.2006.04.003.