Artículo III

Referencia: Lugo-Marín, J., Díez-Villoria, E. Magán-Maganto, M., Pérez-Méndez, L.; Alviani, M.; de la Fuente-Portero, J.A.; & Canal-Bedia, R. (2019). Spanish validation of the Autism Quotient Short Form Questionnaire for adults with Autism Spectrum Disorder. *Journal of autism and developmental disorders,* accepted for publication.

Título: Validación española de la forma abreviada del cuestionario 'Cociente Autista' en población adulta con Trastorno del Espectro Autista

Resumen

Introducción: Debido a la inclusión aun reciente de las formas más leves de Trastorno del Espectro Autista (TEA) en las clasificaciones diagnósticas, muchas personas adultas que presentan características relacionadas con el TEA pueden no haber sido identificadas en los primeros estadios del desarrollo. Si sabemos que esta población presenta una alta frecuencia de trastornos psiquiátricos co-ocurrentes, es probable que hayan establecido contacto con servicios de psiquiatría en la edad adulta, en los que la exploración diagnóstica se centra únicamente en la patología concomitante. Más aún, el solapamiento de las características relacionadas con el TEA con otras entidades diagnósticas de aparición en el principio de la edad adulta, especialmente los trastornos del espectro de la esquizofrenia (TEE), han podido sobreestimar la prevalencia de estos, desviando el foco de la intervención hacia los síntomas derivados en lugar de intervenir en las dificultades relacionadas con el TEA.

Objetivo: el objetivo del presente estudio fue adaptar y validar la versión abreviada del cuestionario Cociente Autista en una muestra de adultos de habla española.

Método: Un total de 46 adultos con TEA, 41 familiares de personas con TEA, 17 pacientes con TEE, y 190 adultos sin historia reportada de patología psiquiátrica completaron el cuestionario

Cociente Autista abreviado. Se estudió la fiabilidad a través del análisis de la consistencia interna y la fiabilidad test-retest. Un análisis factorial confirmatorio se llevó a cabo para poner a prueba el modelo factorial previamente propuesto por los autores del cuestionario. Las diferencias intergrupo en las puntuaciones del cuestionario, así como la correlación con una medida gold-standard de evaluación de TEA (ADOS-2) también fueron objeto de análisis. Por último, el análisis de las habilidades discriminatorias del test fue realizado mediante el estudio de curvas ROC.

Resultados: los resultados del análisis factorial confirmatorio fueron aceptables, pero no excelentes, con valores RMSEA cercanos a .07, y valores CFI y TLI entre.90 y .95, y valores WRMR > .90. El análisis de la consistencia interna mostró que el Cociente Autista abreviado presenta una muy buena estructura interna en los cuatro grupos, con valores entre $C\alpha$ =.79 y $C\alpha$ =.88, y una fiabilidad test-retest buena con valores entre *r* = .812 y *r* = .942. El cuestionario mostró una elevada validez convergente con la medida gold-standard de evaluación en TEA (ADOS-2) (*r* = .734, *p* = <.01) y un buen poder discriminante entre los distintos grupos (71.77%). Las diferencias entre grupos en las puntuaciones del cuestionario resultaron significativas, con el grupo TEA mostrando una puntuación significativamente mayor al resto de grupos. El punto de corte de >63 mostró buenas propiedades psicométricas en la detección de adultos con TEA versus no clínicos (sensibilidad de .98 y especificidad de .84) y TEA versus TEE (punto de corte >65, sensibilidad de .94 y especificidad de .77)

Conclusiones: el cuestionario Cociente Autista abreviado presenta propiedades psicométricas aceptables para la detección de adultos con TEA en población de habla española. Las puntuaciones en el cuestionario apoyan el argumento de un continuo dimensional en la manifestación de las características relacionadas con el TEA. El cuestionario puede ser útil en el proceso de evaluación diagnóstica de adultos con sospecha de TEA, así como en el diagnóstico diferencial con los TEE.

De: "Fred Robert Volkmar" <em@editorialmanager.com>
Asunto: Your Submission JADD-D-19-00182R2
Fecha: 20 de junio de 2019, 13:10:25 CEST
Para: "Ricardo Canal-Bedia" <rcanal@usal.es>
Responder a: "Fred Robert Volkmar" <jaddassist@yale.edu>

Dear Dr. Canal-Bedia,

We have completed our review of your revised manuscript: "Spanish validation of the Autism Quotient Short Form Questionnaire for adults with autism spectrum disorder". We appreciate your careful attention to the reviewers' concerns and feel that the manuscript is now ready for publication.

You will be contacted about proofs and offprints by Springer. Please remember to quote the manuscript number, JADD-D-19-00182R2, whenever inquiring about your manuscript.

Thank you for this interesting contribution. We are pleased that you chose to submit your work to the Journal of Autism and Developmental Disorders. We wish you the very best in your research and look forward to hearing from you again soon.
Sincerely,
Fred Robert Volkmar, MD
Editor in Chief
Journal of Autism and Developmental Disorders

Spanish validation of the Autism Quotient Short Form for adults with autism spectrum disorder

Abstract

The objective of this study was to adapt and validate the abbreviated version of the "Autism-Spectrum Quotient" (AQ-Short) in a sample of Spanish native adults. A total of 46 individuals with ASD, 41 ASD-relatives, 17 patients with schizophrenia spectrum disorders (SSD) and 190 non-clinical adults were administered the Spanish version of the AQ-Short. The results of the confirmatory factorial analysis found two high-order factors (Social Behaviour and Numbers/Patterns) and four subscales (Social Skills, Routines, Switching and Imagination). The reliability analysis showed very good internal structure and test-retest reliability. The AQ-Short also showed moderate convergent validity with ADOS-2. Differences by group were found in the ASD group when compared to other groups. Gender differences were only found in the non-clinical group.

Keywords: Autism Spectrum Quotient; ASD; Validity; Reliability; Factor Analysis; Schizophrenia Spectrum Disorders; Diagnosis

Spanish validation of the Autism Quotient Short Form for adults with autism spectrum disorder

The dimensional approach to autism spectrum disorder (ASD) assumes that the defining characteristics of ASD, that is, the difficulties in social communication and the repetitive behaviours and restricted interests, are the extreme expression of common features present in the general population (American Psychiatry Association, 2013). A disorder of the autistic spectrum would thus be considered as a diagnosis when the intensity of characteristics is severe enough that it would not allow the individual to adapt to their environment.

This dimensional approach has acquired great relevance in the last two decades due to its usefulness for both clinical practice and in the identification of endophenotypes that can be analysed in genetic research (Chakrabarti et al., 2009; Grzadzinski, Huerta, & Lord, 2013). That is why various studies have been conducted to determine the presence of autism-like traits in specific populations, such as in first-degree relatives of people with ASD, where a milder phenotype is often observed but with a profile similar to the defining traits of ASD, which is called the broader autism phenotype (BAP) (Bishop et al., 2004; Eyuboglu, Baykara, & Eyuboglu, 2018; Klusek, Losh, & Martin, 2014; Losh, Childress, Lam, & Piven, 2008; Micali, Chakrabarti, & Fombonne, 2004; Piven et al., 1990; Piven & Palmer, 1999; Ruta, Mazzone, Mazzone, Wheelwright, & Baron-Cohen, 2012; Ruzich et al., 2015).

But this broad phenotype is not restricted to family members of individuals with ASD, and numerous studies have provided evidence suggesting that autistic traits are continuously distributed in the general population (Constantino & Todd, 2003; Ruzich et al., 2015). This fact has led to the idea that individuals in the general population who score high on BAP measures should also express some degree of deficit in the same characteristic areas of autism. Following this approach, a wide variety of studies have attempted to identify subclinical characteristics of autism in the general population. For example, some studies have analyzed the difficulties in social functioning related to BAP (Jobe & White, 2007) or whether autistic traits affect

relationship satisfaction (Pollmann, Finkenauer, & Begeer, 2010), as well as social characteristics such as lower tendency to correspond to direct gaze (Chen & Yoon, 2011), and the reciprocal relation between face recognition and autistic traits (Halliday, MacDonald, Sherf, & Tanaka, 2014). Other studies have examined the relationship between measures of the BAP and cognitive characteristics associated with autism, such difficulties in perceptual speech processing (Stewart & Austin, 2009), or whether the capacity to engage in detailed visuospatial analysis, a frequent feature in autism, extend into the general population (Grinter, Van Beek, Maybery, & Badcock, 2009).

BAP is also a topic of special interest because it can help clinicians identify autism traits in patients with disorders other than ASD. For example Sizoo et al. (2009), investigated the presence of autism traits in patients with ADHD and substance use. BAP has also been studied in patients with social anxiety disorder (Hoekstra, Bartels, Cath, & Boomsma, 2008), in patients with obsessive-compulsive disorder (Hoekstra et al., 2008; Mito et al., 2014; Wikramanayake et al., 2018), patients with anorexia nervosa (Rhind et al., 2014), depressed patients (Takara & Kondo, 2014) and patients with schizophrenia spectrum disorders (Gillespie, Mitchell, & Abu-Akel, 2017; Naito, Matsui, Maeda, & Tanaka, 2010; Solomon et al., 2011).

The case of schizophrenia spectrum disorders (SSD) needs to be considered when approaching concurrent psychiatric disorders in ASD. This group has been repeatedly identified as having a higher presence of autistic-like traits than the general population (Hallerbäck, Lugnegård, & Gillberg, 2012; Lugnegård, Hallerbäck, & Gillberg, 2015; Naito et al., 2010; Wouters & Spek, 2011). This is not surprising, because autism was first conceptualised as a core feature of the schizophrenic disorder (Bleuler, 1911). Both disorders share a set of common characteristics that make differential diagnosis challenging. Difficulties in social communication (Couture et al., 2010), emotion recognition (Bölte & Poustka, 2003), mentalizing skills (Martinez et al., 2017; Pilowsky, Yirmiya, Arbelle, & Mozes, 2000), a high prevalence of formal thought disorders (Gaag, Caplan, Engeland, Loman, & Buitelaar, 2005; Solomon, Ozonoff, Carter, & Caplan, 2008), and findings at a neuropsychological level (Eack et al., 2013; Marinopoulou,

Lugnegård, Hallerbäck, Gillberg, & Billstedt, 2016) suggest the need for diagnostic instruments that allow clinicians to better differentiate between both disorders. Regarding the study of clinical overlap, numerous studies have shown interest in empirically explore this issue. When measuring the presence of autistic-like traits in adults with SSD, studies have reported a high prevalence of these, suggesting a direct relationship between both disorders (Barlati, Deste, Gregorelli, & Vita, 2018; Lugnegård et al., 2015; Spek & Wouters, 2010). In the same way, many people with ASD often develop psychotic-like symptoms (Blackshaw, Kinderman, Hare, & Hatton, 2001; Craig, Hatton, Craig, & Bentall, 2004; Jänsch & Hare, 2014), considering autistic characteristics a risk factor for receiving a diagnosis that falls into the schizophrenia spectrum (Lugo et al., 2018). Due to the late inclusion of mildest-forms of ASD (with an average intellectual quotient and preserved verbal language) in the diagnostic classifications, adults with a possible ASD, who currently have obvious adaptive difficulties, are entering the clinical setting in recent years. It is possible that in these people, now adults, no ASD symptoms were identified during their childhood (Lai & Baron-Cohen, 2015). Thus, when considering the overlap between the diagnostic criteria for SSD and those for ASD, it is not difficult to realize that many of them may have been misdiagnosed, overestimating the prevalence of SSD in people with ASD.

Several instruments have been developed to assess the BAP (Landry & Chouinard, 2016). One of the most widely used is the Autism Quotient (AQ) (Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001). The AQ is a self-reported screening instrument with 50 items, in which the person must choose whether to agree or disagree with questions related to ASD characteristics (e.g., I prefer to do things the same way over and over again; I tend to have very strong interests). Factor analysis showed a theoretical model composed of five domains: social skills, attention to detail, attention switching, communication and imagination. According to the original validation, a score above 32 would suggest that the individual may have clinically significant levels of autistic traits, which does not mean that the individual actually has ASD, since in order to reach that conclusion, a more comprehensive evaluation would be needed to

identify the presence and clinical significance of core autism traits in that individual (Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001). In 2011, Hoekstra et al. developed an abbreviated 28 item version of the instrument (AQ-Short), in order to increase the efficiency of the test. Five subscales were identified in the AQ-Short: social skills, routines, switching, imagination, and numbers/patterns. These subscales were combined into two higher-order factors: social behavioural difficulties and a fascination with numbers/patterns. AQ-Short can also be valuable as a rapid assessment of autistic traits for screening purposes (Hoekstra et al., 2011). The study of these authors indicates that a score above 65 would merit a referral to a specialized service to confirm (or rule out) a diagnosis of autism.

Although the AQ has been adapted and validated in several languages, showing appropriate psychometric properties (do Egito, Ferreira, Gonçalves, & Osório, 2017; Hoekstra et al., 2008; Lau et al., 2013; Lepage, Lortie, Taschereau-Dumouchel, & Théoret, 2009; Wakabayashi, Baron-Cohen, Wheelwright, & Tojo, 2006), a validation for the Spanish-speaking population has not yet been conducted. The language validation is an essential process in ensuring the conservation of the psychometric properties of the tests. The process of translating and adapting a scale requires more than translation into the target language. It is necessary to ensure that the scores obtained with the translated version are equivalent to those obtained with the original test. It is therefore necessary for the Spanish-speaking population to provide answers on the psychometric characteristics and the clinical value of the AQ-Short questionnaire. It was, therefore, the aim of the present study to validate the AQ- Short for use with the Spanish adult general population. The characteristics of the Spanish AQ- Short, including test–retest reliability, internal consistency, convergent validity, and confirmatory factor analysis were studied. Also, group differences between individuals with ASD and other three groups of participants (ASD first-degree relatives, SSD patients and Non-Clinical comparison subjects) were analysed in order to explore the discriminative power of the AQ with the SSD and Non-Clinical groups and to explore the presence of the BAP, meaning subclinical autistic traits, in the ASD relatives and the SSD patients, thus supporting the dimensional approach of ASD.

METHODS

Participants

All participants were born in Spanish territory, had Spanish as their mother tongue, and were 18 years old or above. Four groups of participants were evaluated in order to establish between-group comparisons with the AQ-Short and other studied variables. The groups were as follows.

ASD group. Participants from this group (n = 46) were recruited from ASD community assistance centres and via online announcements in ASD associations webpages. All participants reported that they had received prior to evaluation a diagnosis of ASD regarding DSM-5 criteria (American Psychiatric Association, 2013), and provided the diagnosis signed by a qualified clinician. Below average IQ was set as an exclusion criterion, as it could affect the appropriate understanding of the test.

ASD-relatives group. The participants in this group (n = 41) were recruited via public and online announcements on ASD associations webpages.

SSD group. Participants from this group (17 patients with SSD) were recruited from the care units of the psychiatric service of the Hospital Universitario Nuestra Señora de Candelaria (Santa Cruz de Tenerife, Spain) and all of them fulfilled DSM-5 criteria for any SSD at the time of the study (American Psychiatric Association, 2013). They were asked to complete the AQ-Short. Excluding criteria were acute psychotic symptoms at the time of the evaluation, intellectual difficulties and an above-threshold score in an ASD gold-standard measure (Autism Diagnostic Observation Schedule-2) (Lord et al., 2012).

Non-clinical group. Data was collected from 190 non-clinical participants recruited via internet advertisements and by word of mouth through acquaintances of the authors. This group were explicitly defined as not having received an ASD diagnosis, nor any other psychiatric disorder diagnosis prior to the evaluation. Participants were directly asked whether they believed they

had an undiagnosed ASD, ruling out their scores for the data analysis, thus decreasing the risk of overestimating their scores in the AQ-Short.

For all participants, when not able to directly measure IQ, this was considered into the average range when reporting a higher qualification and/or the absence of a prior diagnosis of intellectual disability.

Instruments

Demographic Questionnaire: the following data was collected regarding the socio-demographic and clinical characteristics of all participants: age, gender, higher education level, work/academic status, psychopharmacological treatment, and psychiatric history.

The Spanish Autism-Spectrum Quotient (AQ-Short): The AQ-Short was translated into Spanish after obtaining permission from the original author, Professor Simon Baron-Cohen. . The translation was done by a native Spanish speaker, a professor of English philology who remained blind to the construct that was intended to be measured by the questionnaire. The instruction given to this translator was that, instead of being literal, the translation should seek semantic, linguistic and cultural equivalence. This first translation was revised by the first author of this article to verify that the items of this first Spanish version corresponded to the items of the original version. Subsequently, a bilingual psychologist specialized in ASD translated the Spanish version back into English. Afterwards, a panel of 4 bilingual professionals, experts in ASD, compared the original version of the AQ-Short and the translated version, discussed the points of discrepancy and reached agreements to introduce the necessary corrections in the Spanish version to reflect as accurately as possible the content of each item. The final version of the Spanish AQ is similar to the original English, maintaining the same format of 28 forced choice items. The Spanish version of the AQ-Short is composed of 28 items. The total AQ-Short score and the subscales were based on the original Likert responses (1 = "definitely agree", 2 = "slightly agree", 3 = "slightly disagree", and 4 = "definitely

disagree"). Items 2, 4, 5, 7, 10, 13, 14, 15, 16, 22, 23, 25 and 26 were reverse scored. In the original study by Hoekstra et al. (2011), the minimum score was 28 points and the maximum 112, with 65 points an acceptable cut-off point for distinguishing autistic-like traits, with sensitivity and specificity values of .97 and .82, respectively.

Reynolds Intelligence Screening Test (RIST) (Reynolds & Kamphaus, 2003b)**:** This is a screening test that estimates a general measure of the IQ in an age-range from 3 to 94 years. The RIST has its origin in the Reynolds Intelligence Assessment Scales (RIAS) (Reynolds & Kamphaus, 2003a) and is composed of two of its subtests: 'Guess what' (verbal subtest) and 'Odd-item out' (non-verbal subtest). The verbal subtest is a classic measure of crystallised intelligence, and the non-verbal subtest is closely related to the assessment of fluid intelligence. Both subtests have shown good psychometric properties. The Spanish RIST was used for this study, which was validated with more than 2,000 Spanish individuals and showed good psychometric properties (Santamarıa & Fernández Pinto, 2009).

Autism Diagnostic Observation Schedule (ADOS-2) (Lord et al., 2012)**:** The Autism Diagnostic Observation Schedule (ADOS-2) is a standardised and semi-structured assessment of communication, social interaction and play/imagination skills for people with suspected ASD. The scale is structured in five modules (T, 1, 2, 3 and 4), each created to be used as a function of the chronological age and language level of each individual. Each module is composed of a set of activities that provide standardised contexts, where an evaluator can observe (or not) the presence of certain social and communicative behaviours relevant to the diagnosis of ASD. After conducting the Module 4 (adults) protocol and obtaining the algorithm scores, there are three possible outcomes: autism, autism spectrum and non-ASD.

Positive and Negative Symptoms Scale (PANSS) (Kay, Fiszbein, & Opler, 1987)**:** This is a semi-structured clinical interview consisting of 30 items and four scales: Positive Symptoms (PANSS-P); Negative Symptoms (PANSS-N); Composite Scale (PANSS-C); and General Psychopathology (PANSS-GP). Each item is scored according to a Likert scale of seven degrees

of severity. The Spanish PANSS was used for this study, which has showed good psychometric properties in the diagnosis of SSD (Peralta & Cuesta, 1994).

Procedure

Participants were asked to complete the socio-demographic and AQ-Short questionnaires individually through a web survey application or in paper-and-pencil format. In order to evaluate the ASD diagnosis and the IQ, a single subgroup was extracted from ASD (n = 24) and non-clinical (n = 18) groups. These two subgroups, plus the whole SSD group (n = 17), were assessed individually with the ADOS-2 and the RIST by a clinical psychologist trained and accredited in the use of these instruments. The SSD group was also assessed with the PANSS in order to confirm the absence of acute psychotic symptomatology at the time of evaluation. Only the socio-demographic and AQ-Short questionnaires were administered to the ASD-relatives group. In order to explore test-retest reliability, randomly selected participants from the ASD (n = 26), SSD (n = 9) and non-clinical (n = 61) groups were asked to complete the AQ-Short a second time. The windows between the first and second evaluations were from 2 to 16 weeks.

Statistical analysis

Reliability was studied through the analysis of internal consistency (Cronbach's alpha), and the test-retest analysis with Spearman rho correlations and intraclass correlation coefficients (ICC3) (Shrout & Fleiss, 1979). Cronbach's α internal consistency measures were considered minimally acceptable when α= .65, acceptable when α= .70, and optimal when α= .80 (Nunnally & Bernstein, 1994).

Confirmatory factor analysis was conducted to test the factor model structure most commonly proposed for the AQ-Short. This factor structure had been tested previously in a control sample (Hoekstra et al., 2011) and in a sample of adults with clinically diagnosed ASD (Kuenssberg, Murray, Booth, & McKenzie, 2014). Due to the limited sample size of the ASD and SSD

groups the factor structure was tested with all the sample responses (and also replicated with the non-clinical sample excluding ASD and SSD). The model was estimated using Lavaan version 0.5-23.1097 (Rosseel, 2012), via robust diagonally least squares estimation (WLSMV) (Beauducel & Herzberg, 2006). WLSMV is recommended when there are four (or less) ordered response categories, as is the case for the AQ-Short response Likert-scale. Model fit was evaluated using the usual χ^2, as well as the comparative fit index (CFI) and Tucker Lewis index (TLI) as incremental fit indices, and the weighted root mean square residual (WRMR) and root mean squared error of approximation (RMSEA) as baseline fit indices.

Group differences in AQ scales and subscales by group were analysed using Kruskall Wallis, and the Dwass-Steel-Critchlow-Fligner (Douglas & Michael, 1991; Dwass, 1960; Steel, 1960) procedure was used for multiple pairwise comparisons. Group differences in AQ scales and subscales by gender were analysed using Welch's *t*-tests. Gender DIF (Differential Item Functioning) was analysed with the iterative hybrid ordinal logistic regression/item response theory (Choi, Gibbons, & Crane, 2011) because it can effectively handle the polytomous property of AQ items (4-point Likert scale).

The convergent validity of the AQ-Short was explored with a correlational analysis with the ADOS-2. We considered correlation coefficients small when $r = .10$, moderate when $r = .30$ and large when $r = .50$ (Cohen, 1988). Linear discriminant analysis was used to evaluate discriminant validity (i.e. the ability of AQ global and subscales to discriminate between groups with known differences).

A receiver operating characteristic (ROC) curve analysis was carried out to evaluate the AQ-Short cut points suggestive of an ASD diagnosis. As the primary interest was to quantify how accurately AQ-Short can discriminate between ASD and non-clinical subjects only those groups were included in the analysis. The area under the curve (AUC), accuracy, sensitivity and specificity for the Youden index-based cut-off scores were calculated, and ROC curve plots were used to draw the specificity versus the sensibility for the candidate threshold values between 0.0 and 1.0.

Unless otherwise specified, an alpha level of .05 was used to test for significance. All statistical analyses were performed using the computing environment R (R Core Team, 2013)

Ethics approval

Approval for this study was obtained from the Hospital Universitario Nuestra Señora de Candelaria Research Ethics Committee (PI-32/17). All participants gave written or web-based consent in accordance with the Declaration of Helsinki.

RESULTS

The socio-demographic and clinical characteristics of the four groups are shown in Table 1. A non-parametric one-way ANOVA Kruskal-Wallis found a significant difference regarding age between groups, with ASD-relatives being older than other participants (x^2 (3) = 45.1, $p < .001$). Differences regarding IQ reached significance for the IQ-Total ($w = 3.27$, $p < .002$) and non-verbal-IQ scores ($w = 4.03$, $p < .004$), with the ASD group having lower scores than the non-clinical participants. Global gender proportion differences were found as a function of Group ($\chi^2(3) = 47.5$; $p < .001$). A set of two tailed binominal tests showed proportion differences ($p < .05$) in the SSD group (more males than females) , and the ASD-relatives ($p < .01$) and non-clinical groups ($p < .001$) which were mostly female. No proportion differences were found in the ASD group ($p = .01$).

[Insert Table 1 about here]

Internal Consistency and Test-Retest Reliability

Internal consistency was assessed for the four groups of participants (Table 2). All AQ-Short scales showed satisfactory internal consistency for the four groups, ranging from $C\alpha$ =.79 to $C\alpha$ =.88. In the ASD group, items 14 (*When reading a story, I find it difficult to work out the character's intentions*) and 16 (*I notice patterns in things all the time*) correlated negatively with the total scale. In the SSD group, items 3 (*Trying to imagine something, I find it easy to create a picture in my mind*), 11 (*I find making up stories easy*) and 20 (*I find it easy to work out what someone is thinking or feeling*) also correlated negatively with the total scale.

[Insert Table 2 about here]

Table 3 shows the very good test-retest reliability of the AQ-Short version and all the subscales (ICC ranging from .90 to .97). No significant differences were found between the two measurement times across subscales, implying sufficient constancy.

[Insert Table 3 about here]

Factor analysis

As shown in Table 4, the model fit can be considered acceptable but not excellent, with good fit to the data according to RMSEA values close to .07 and CFI or TLI values between .90 and .95 (Hu & Bentler, 1998), but worse fit according to WRMR (values > 1.0); (Yu, 2002) and the null of perfect fit rejected (but see DiStefano, Liu, Jiang, and Shi (2018) for evidence of the unexpected behaviour of WRMR in some situations; and Brown (2014), for the relative value of chi-square to test the quality of the model).

Figure 1 shows the parameter estimates for the structural analysis of the AQ-Short.

[Insert Table 4 about here]

[Insert Figure 1 about here]

Intergroup differences in AQ-Short

A set of non-parametric Kruskall-Wallis H tests showed significant differences by group for the global score ($\chi^2(3) = 101$, $p < .001$) and also for the two high-order factors, Social Behaviour ($\chi^2(3) = 95$, $p < .001$) and Numbers/Patterns ($\chi^2(3) = 55.2$, $p < .001$), and the four subscales included in the Social Behaviour higher order factor: Social Skills ($\chi^2(3) = 60.4$, $p < .001$), Routines ($\chi^2(3) = 64.7$, $p < .001$), Switching ($\chi^2(3) = 85.6$, $p < .001$) and Imagination ($\chi^2(3) = 80.5$, $p < .001$) (see Table 5). Multiple comparisons between groups controlling for the error rate simultaneously for all contrasts (Dwass-Steel-Critchlow-Fligner procedure) showed significant differences in all scores for the ASD group when compared to others. The SSD and ASD-relatives groups did not differ on any AQ scale. Compared to the non-clinical sample, the SSD group scored higher in the AQ total score ($w = -4.92$, $p < .001$), Social Behaviour high-order factor ($w = -4.72$, $p < .001$) and the Switching ($w = -4.91$, $p < .001$) and Imagination ($w = -5.57$, $p < .001$) subscales. The ASD-relatives group scored significantly higher than the Non-Clinical group in the AQ total score ($w = -3.92$, $p < .006$) Social Behaviour high-order factor ($w = -4.08$, $p < .004$) and the Social Skills ($w = -3.35$, $p < .018$), Switching ($w = -3.78$, $p < .008$) and Imagination ($w = -4.40$, $p < .002$) subscales. Figure 2 shows the violin plots for the AQ total scale and intergroup differences (see Appendix A for other AQ scales).

[Insert Table 5 about here]

[Insert Figure 2 about here]

Convergent and Discriminant Validity

A correlation analysis was conducted to determine the relationship between AQ-Short and ADOS-2. The analysis showed a significant correlation between two instruments (r = .734, p = <.001), meaning moderate convergent validity. A subgroup analysis showed only a significant positive correlation in the Non-Clinical group (r = .7, p = <.001) (Figure 3)

[Insert Figure 3 about here]

Discriminant validity was explored by way of a linear discriminant analysis to find if a linear combination of the AQ-Short subscales can be used to correctly classify participants in the four groups. As can be seen in the prediction-accuracy figure (Figure 4), model's correct prediction was higher in the non-clinical group (178 out of 190 = 93.68%) than in the ASD group (32 out of 46 = 69.57%). This result indicates that AQ-Short has a better negative predictive than positive predictive value. In the case of ASD relatives and SSD groups, prediction-accuracy was quite low (2.44% and 0% respectively), as those cases were classified as Non-clinical according to AQ-Short subscales scores. Overall accuracy was 71.77%.

[Insert Figure 4 about here]

Cut-Off Scores of AQ Total

Table 6 shows predictive values for a range of potential cut-offs. In our study the total AQ-Short score of >63 gives a sensitivity of .98 and a specificity of .84. The area under the curve was .95, indicating excellent test accuracy. Consistently with previous discriminant analysis results, positive and negative predictive values were 0.6 and 0.994, respectively. Figure 5 shows the ROC curve for AQ total and subscales comparing ASD and Non-Clinical groups. Similar predictive properties were found when comparing participants reaching ASD threshold in the ADOS-2 (ASD n=24) to those who did not (SSD (n=17) and Non-Clinical (n=18), with a cut-off of >65 giving a sensitivity of 1 and specificity of .82. When comparing both clinical groups (ASD vs SSD) a cut-off point of >65were found, with a sensitivity of .94 and specificity of .77.

[Insert Table 6 about here]

[Insert Figure 5 about here]

DISCUSSION

To the extent of our knowledge, this study is the first Spanish validation of any version of the AQ questionnaire and the first replication in a non-English language of the abridged version of the AQ. Moreover, this study is the first validation of the AQ questionnaire which includes ASD diagnostic gold-standard measures not only in an ASD adult sample, but also other clinical/non-clinical groups.

The results of the reliability analysis suggest the good internal structure of the AQ-Short, supporting the results previously found in the original validation (Hoekstra et al., 2011). When comparing four groups in the AQ total scores, it seems that clinical groups (ASD, SSD) are less coherent in their responses, suggesting the need to identify the possible difficulties of these groups when understanding the items. An extension of the sample size could resolve this issue. Even so, both groups showed great reliability in the AQ-Short total score ($C\alpha$ = 0.8). When considering AQ-Short domains, the ASD and SSD groups showed lower than acceptable correlations (<.65) in the Routines, Switching and Imagination subscales. Three items weighing in the Imagination factor scale was found to negatively correlate with the total AQ-Short scale in the SSD group. It is plausible that some items from the Imagination factor correlate negatively with the total scale score in the SSD group, since people with SSD do not have a problem imagining *per se*, although the content and structure of what they imagine is different (Rasmussen & Parnas, 2015). The literature has provided abundant evidence that the ability to imagine clearly differentiates people with ASD from people with SSD (Crespi, Leach, Dinsdale, Mokkonen, & Hurd, 2016). The inverse relationship between imagination domain scores and total AQ-Short scores in the SSD group may be in accordance with the characteristic trend/imagery trait of people with SSD, which clinically differentiates them from people with ASD. This finding, however, should be confirmed with a larger sample of people with SSD, which, if so, would increase confidence in this instrument as a clinical resource with which to discriminate between the two disorders. It also raises the need to further study the abilities to imagine and to use imagination in a social context as relevant characteristics that should be different in both disorders.

Factor analysis showed good to very good model fitting to the two-factor model proposed by the original validation of the AQ-Short (Hoekstra et al., 2008; 2011). Two previous studies have suggested the existence of three underlying factors: social skills, details/patterns, and communication/mindreading (Austin, 2005; Hurst, Mitchell, Kimbrel, Kwapil, & Nelson-Gray, 2007). Another recent study (Murray, McKenzie, Kuenssberg, & Booth, 2017) tried to adjust a

bifactorial model confirming the AQ and the AQ-Short, assessing the extent to which the specific symptom areas measured by this questionnaire reflect the specific factors desired compared to a general factor of ASD. Their results indicate that for the AQ-Short, the covariance of the items mainly reflects the existence of a general factor rather than specific factors, with the exception of the items corresponding to the Numbers/Patterns scale. The finding of these authors is consistent with that obtained by Hoekstra et al. (2011) and by our study. Our study found a greater correlation between Social Behaviour and Numbers/Patterns (.62) which would question the hypothesis that Numbers/Patterns is a relatively different construct from that of Social Behaviour, as suggested in the study by Hoekstra et al. (2011).

The ASD group showed a significantly higher AQ-Short total score compared to other groups. This difference also reached statistical significance when comparing AQ subscales. The Routines and Numbers/Patterns subscales were clearly the best with which to discriminate ASD from other groups. Interestingly, the SSD and ASD-relatives groups did not differ in any AQ scale, thus suggesting a similar prevalence of BAP in both populations. The results found here support the dimensional approach to BAP in general population, with ASD participants reaching the highest scores, followed by SSD and ASD-relatives, and the non-clinical participants having the lowest BAP of all.

One of the main strengths of the present study was the use of gold-standard measures for ASD in a subsample of participants in order to study the association with AQ-Short scores. Correlation analysis indicated a moderate association between both measures when comparing all groups, but correlation analysis by group showed a positive significant correlation only in the non-clinical group, with both ASD and SSD showing non-significant correlations. This may be due to the small sample size and lower variability of ADOS-2 scores in both clinical samples, with the non-clinical group having a higher number of participants.

The original cut-off score reported by Hoekstra et al. (2011) was 65 points, with a sensitivity and specificity of .97 and .82, respectively. These results are similar to those found in the present study (cut-off point of >63, sensitivity of .98 and specificity of .84), supporting the

value of the AQ-Short as a screening test when the 50-item original AQ may be too demanding in terms of cognitive and time resources.

Limitations

Our study has some limitations which need to be considered. First, the small sample sizes may have biased the results found here, especially those regarding cut-off points for screening ASD. Also, the AQ-Short is a self-report measure. It has been suggested that people with ASD might have poor insight when asked about their own behaviours (Johnson, Filliter, & Murphy, 2009). As the AQ-Short items point to preferences rather than behaviours, dependence on a reliable self-knowledge on individual difficulties is overcome (Baron-Cohen et al., 2001).

ASD diagnosis must be based in both standardised measures and clinical judgment, and this is imperative when it comes to research. Only a few participants of the ASD group in this study were assessed by a clinician who specialised in ASD using the current gold-standard measure to support diagnosis. The self-description of the rest of the participants on having received an ASD diagnosis in the past was enough to fulfill this inclusion criteria. As can be easily inferred, some participants may not really have fulfilled a clinical ASD diagnosis, thus confounding the validity of the group scores. The opposite effect can also be found, where non-diagnosed participants who think they may have an unidentified ASD diagnosis try to fulfill their prophecy by overestimating their chance of falling into the autism spectrum when answering the AQ-Short. This issue was addressed in our study by asking participants directly whether they believed they had an undiagnosed ASD, ruling out their scores for the data analysis, thus decreasing the risk of overestimating their scores in the AQ-Short.

Finally, it is important to consider the differences found in gender-ratios between clinical vs. non-clinical groups in our study. The SSD group was mostly male, while the ASD-relatives and non-clinical groups were predominantly female. This may have affected the results of the AQ-

Short scores found in our study, which pointed to a higher prevalence of BAP in male participants, but only in the Non-Clinical group.

In order to explore these differences, a robust ANOVA (10% trimmed means) showed a significant effect of Group factor (Q = 198; p < .001) but not for Gender nor the interaction Group x Gender. In any case, a series of Welchs t tests were conducted to test for gender differences in each group, finding a significant effect of gender only in the non-clinical group (t (84) = 3.29, p < .01, d = 0.51, 95% CI [0.22, 0.92]*)*, indicating a higher proportion of ASD characteristics for males (M = 51.4; SE=1.03) than for females (M = 57.5; SE=1.52). An additional analysis of item differential functioning by gender was conducted in all participants, showing that the AQ-Short did not exhibit overall item response differences between males and females (see Appendix B). Further research is needed into gender differences in the scores obtained in the AQ, being this aim out of the scope of the present study.

Clinical implications

The results of the present study suggest the AQ-Short is a reliable instrument for the screening of ASD in Spanish native-speaker adults. It shows very good internal structure and a good convergence with diagnostic ASD gold-standard measures. Further, the AQ-Short seems to clearly differentiate ASD from other specific populations (SSD, ASD-relatives), and is thus a useful instrument in the screening stage. As this is the only reported validation of the AQ-Short in a non-English language, more evidence on the accuracy of this instrument in other languages is needed.

Acknowledgments: The authors wish to thank all the participants for their selfless collaboration in this study. Also, board members of Asociación Síndrome de Asperger Islas Canarias (ASPERCAN) and Confederación Asperger España (CONFAE) for their help in the diffusion of this study. M.D. Inocencio Díaz Martínez and Cl. Psc. Beatriz Ferrera, facultative clinicians of

the Hospital Universitario Nuestra Señora de Candelaria and AFES, respectively, for their collaboration in the recruitment stage. Finally, the authors would like to thank Prof. Dr Blanca Martín Torres and PhD Student Clara Janicel Fernandez for their participation in the Spanish translation of the AQ-Short, and PhD Patricia García Primo for reviewing the manuscript

Author Contributions: JLM and MMM participated in the study design, data collection, statistical analyses and manuscript preparation; LPM participated in the study design and statistical analyses. MA and JFP participated in manuscript preparation and review. EDV and RCB participated in the study design, statistical analyses and manuscript preparation. All authors read and approved the final manuscript.

Funding: This research was conducted as a part of the first author's PhD project. The authors were supported by a research grant awarded by the Spanish Ministry of Economy and Competitiveness (grant PSI2016-80575-R), and European Union. DGSANCO. Ref.: SANCO/2014/C2/035.

Conflict of interest: The authors declare that they have no conflict of interest.

Ethical Approval: All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

REFERENCES

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*. 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*(2), 451–460.

Barlati, S., Deste, G., Gregorelli, M., & Vita, A. (2018). Autistic traits in a sample of adult patients with schizophrenia: prevalence and correlates. *Psychological Medicine*, 1–9.

Baron-Cohen, S. (2002). The extreme male brain theory of autism. *Trends in Cognitive Sciences*, *6*(6), 248–254.

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): Evidence from asperger syndrome/high-functioning autism, malesand females, scientists and mathematicians. *Journal of Autism and Developmental Disorders*, *31*(1), 5–17.

Beauducel, A., & Herzberg, P. Y. (2006). On the performance of maximum likelihood versus means and variance adjusted weighted least squares estimation in CFA. *Structural Equation Modeling*, *13*(2), 186–203.

Bishop, D. V. M., Maybery, M., Maley, A., Wong, D., Hill, W., & Hallmayer, J. (2004). Using self- report to identify the broad phenotype in parents of children with autistic spectrum disorders: a study using the Autism- Spectrum Quotient. *Journal of Child Psychology and Psychiatry*, *45*(8), 1431–1436.

Blackshaw, A. J., Kinderman, P., Hare, D. J., & Hatton, C. (2001). Theory of mind, causal attribution and paranoia in Asperger syndrome. *Autism*, *5*(2), 147–163. https://doi.org/10.1177/1362361301005002005

Bleuler, E. (1950). Dementia praecox or the group of schizophrenias.

Bölte, S., & Poustka, F. (2003). The recognition of facial affect in autistic and schizophrenic subjects and their first-degree relatives. *Psychological Medicine*, *33*(5), 907–915.

Brown, T. A. (2014). *Confirmatory factor analysis for applied research*. Guilford Publications.

Chakrabarti, B., Dudbridge, F., Kent, L., Wheelwright, S., Hill- Cawthorne, G., Allison, C., …

Baron- Cohen, S. (2009). Genes related to sex steroids, neural growth, and social–emotional behavior are associated with autistic traits, empathy, and Asperger syndrome. *Autism Research*, *2*(3), 157–177.

Chen, F. S., & Yoon, J. M. D. (2011). Brief report: broader autism phenotype predicts spontaneous reciprocity of direct gaze. *Journal of Autism and Developmental Disorders*, *41*(8), 1131–1134.

Choi, S. W., Gibbons, L. E., & Crane, P. K. (2011). Lordif: An R package for detecting differential item functioning using iterative hybrid ordinal logistic regression/item response theory and Monte Carlo simulations. *Journal of Statistical Software*, *39*(8), 1.

Cohen, J. (1988). Statistical power analysis for the behavioural sciences. Hillsdale, NJ: erlbaum.

Constantino, J. N., & Todd, R. D. (2003). Autistic traits in the general population: a twin study. *Archives of General Psychiatry*, *60*(5), 524–530.

Couture, S. M., Penn, D. L., Losh, M., Adolphs, R., Hurley, R., & Piven, J. (2010). Comparison of social cognitive functioning in schizophrenia and high functioning autism: More convergence than divergence. *Psychological Medicine*, *40*(4), 569–579. https://doi.org/10.1017/S003329170999078X

Craig, J. S., Hatton, C., Craig, F. B., & Bentall, R. P. (2004). Persecutory beliefs, attributions and theory of mind: Comparison of patients with paranoid delusions, Asperger's syndrome and healthy controls. *Schizophrenia Research*, *69*(1), 29–33. https://doi.org/10.1016/S0920-9964(03)00154-3

Crespi, B., Leach, E., Dinsdale, N., Mokkonen, M., & Hurd, P. (2016). Imagination in human social cognition, autism, and psychotic-affective conditions. *Cognition*, *150*, 181–199.

DiStefano, C., Liu, J., Jiang, N., & Shi, D. (2018). Examination of the weighted root mean square residual: Evidence for trustworthiness? *Structural Equation Modeling: A Multidisciplinary Journal*, *25*(3), 453–466.

do Egito, J. H. T., Ferreira, G. M. R., Gonçalves, M. I., & Osório, A. A. C. (2017). Brief Report:

Factor Analysis of the Brazilian Version of the Adult Autism Spectrum Quotient. *Journal of Autism and Developmental Disorders*, 1–7.

Douglas, C. E., & Michael, F. A. (1991). On distribution-free multiple comparisons in the one-way analysis of variance. *Communications in Statistics-Theory and Methods*, *20*(1), 127–139.

Dwass, M. (1960). Some k-sample rank-order tests. *Contributions to Probability and Statistics*.

Eack, S. M., Bahorik, A. L., McKnight, S. A. F., Hogarty, S. S., Greenwald, D. P., Newhill, C. E., … Minshew, N. J. (2013). Commonalities in social and non-social cognitive impairments in adults with autism spectrum disorder and schizophrenia. *Schizophrenia Research*, *148*(1–3), 24–28. https://doi.org/10.1016/j.schres.2013.05.013

Eyuboglu, M., Baykara, B., & Eyuboglu, D. (2018). Broad autism phenotype: theory of mind and empathy skills in unaffected siblings of children with autism spectrum disorder. *Psychiatry and Clinical Psychopharmacology*, *28*(1), 36–42.

Gaag, R. J. van der, Caplan, R., Engeland, H. van, Loman, F., & Buitelaar, J. K. (2005). AControlled Study Of Formal Thought Disorder in Children with Autism and Multiple Complex Developmental Disorders. *Journal of Child & Adolescent Psychopharmacology*, *15*(3), 465–476.

Gillespie, S. M., Mitchell, I. J., & Abu-Akel, A. M. (2017). Autistic traits and positive psychotic experiences modulate the association of psychopathic tendencies with theory of mind in opposite directions. *Scientific Reports*, *7*(1), 6485.

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*(4), 670–677.

Grzadzinski, R., Huerta, M., & Lord, C. (2013). DSM-5 and autism spectrum disorders (ASDs): an opportunity for identifying ASD subtypes. *Molecular Autism*, *4*(1), 12.

Hallerbäck, M. U., Lugnegård, T., & Gillberg, C. (2012). Is autism spectrum disorder common

in schizophrenia? *Psychiatry Research*, *198*(1), 12–17. https://doi.org/10.1016/j.psychres.2012.01.016

Halliday, D. W. R., MacDonald, S. W. S., Sherf, S. K., & Tanaka, J. W. (2014). A reciprocal model of face recognition and autistic traits: evidence from an individual differences perspective. *PloS One*, *9*(5), e94013.

Hoekstra, R. A., Bartels, M., Cath, D. C., & Boomsma, D. I. (2008). Factor structure, reliability and criterion validity of the Autism-Spectrum Quotient (AQ): a study in Dutch population and patient groups. *Journal of Autism and Developmental Disorders*, *38*(8), 1555–1566.

Hoekstra, R. A., Vinkhuyzen, A. A. E., Wheelwright, S., Bartels, M., Boomsma, D. I., Baron-Cohen, S., … van der Sluis, S. (2011). The construction and validation of an abridged version of the autism-spectrum quotient (AQ-Short). *Journal of Autism and Developmental Disorders*, *41*(5), 589–596.

Hu, L., & Bentler, P. M. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. *Psychological Methods*, *3*(4), 424.

Hurst, R. M., Mitchell, J. T., Kimbrel, N. A., Kwapil, T. K., & Nelson-Gray, R. O. (2007). Examination of the reliability and factor structure of the Autism Spectrum Quotient (AQ) in a non-clinical sample. *Personality and Individual Differences*, *43*(7), 1938–1949.

Jänsch, C., & Hare, D. J. (2014). An investigation of the "jumping to conclusions" data-gathering bias and paranoid thoughts in Asperger syndrome. *Journal of Autism and Developmental Disorders*, *44*(1), 111–119. https://doi.org/10.1007/s10803-013-1855-2

Jobe, L. E., & White, S. W. (2007). Loneliness, social relationships, and a broader autism phenotype in college students. *Personality and Individual Differences*, *42*(8), 1479–1489.

Johnson, S. A., Filliter, J. H., & Murphy, R. R. (2009). Discrepancies between self-and parent-perceptions of autistic traits and empathy in high functioning children and adolescents on the autism spectrum. *Journal of Autism and Developmental Disorders*, *39*(12), 1706–1714.

Kay, S. R., Fiszbein, A., & Opler, L. A. (1987). The Positive and Negative Syndrome Scale

(PANSS) for Schizophrenia. *Schizophrenia Bulletin*, *13*(2), 261–276. https://doi.org/10.1093/schbul/13.2.261

Klusek, J., Losh, M., & Martin, G. E. (2014). Sex differences and within-family associations in the broad autism phenotype. *Autism*, *18*(2), 106–116.

Kuenssberg, R., Murray, A. L., Booth, T., & McKenzie, K. (2014). Structural validation of the abridged Autism Spectrum Quotient–Short Form in a clinical sample of people with autism spectrum disorders. *Autism*, *18*(2), 69–75.

Lai, M.-C., & Baron-Cohen, S. (2015). Identifying the lost generation of adults with autism spectrum conditions. *The Lancet. Psychiatry*, *2*(11), 1013–27. https://doi.org/10.1016/S2215-0366(15)00277-1

Landry, O., & Chouinard, P. A. (2016). Why we should study the broader autism phenotype in typically developing populations. *Journal of Cognition and Development*, *17*(4), 584–595.

Lau, W. Y.-P., Gau, S. S.-F., Chiu, Y.-N., Wu, Y.-Y., Chou, W.-J., Liu, S.-K., & Chou, M.-C. (2013). Psychometric properties of the Chinese version of the Autism Spectrum Quotient (AQ). *Research in Developmental Disabilities*, *34*(1), 294–305.

Lepage, J.-F., Lortie, M., Taschereau-Dumouchel, V., & Théoret, H. (2009). Validation of French-Canadian versions of the Empathy Quotient and Autism Spectrum Quotient. *Canadian Journal of Behavioural Science/Revue Canadienne Des Sciences Du Comportement*, *41*(4), 272.

Lord, C., Rutter, M., DiLavore, P. C., Risi, S., Gotham, K., & Bishop, S. (2012). Autism diagnostic observation schedule–Second edition (ADOS-2). *Los Angeles: Western Psychological Services*.

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*(4), 424–433.

Lugnegård, T., Hallerbäck, M. U., & Gillberg, C. (2015). Asperger syndrome and schizophrenia: Overlap of self-reported autistic traits using the Autism-spectrum Quotient (AQ). *Nordic Journal of Psychiatry*, *69*(4), 268–274. https://doi.org/10.3109/08039488.2014.972452

Marín, J. L., Rodríguez-Franco, M. A., Chugani, V. M., Maganto, M. M., Villoria, E. D., & Bedia, R. C. (2018). Prevalence of Schizophrenia Spectrum Disorders in Average-IQ Adults with Autism Spectrum Disorders: A Meta-analysis. *Journal of Autism and Developmental Disorders*, *48*(1), 239–250.

Marinopoulou, M., Lugnegård, T., Hallerbäck, M. U., Gillberg, C., & Billstedt, E. (2016). Asperger Syndrome and Schizophrenia: A Comparative Neuropsychological Study. *Journal of Autism and Developmental Disorders*, *46*(7), 2292–2304.

Martinez, G., Alexandre, C., Mam-Lam-Fook, C., Bendjemaa, N., Gaillard, R., Garel, P., … Krebs, M.-O. (2017). Phenotypic continuum between autism and schizophrenia: evidence from the movie for the assessment of social cognition (MASC). *Schizophrenia Research*, *185*, 161–166.

Micali, N., Chakrabarti, S., & Fombonne, E. (2004). The broad autism phenotype: findings from an epidemiological survey. *Autism : The International Journal of Research and Practice*, *8*(1), 21–37. https://doi.org/10.1177/1362361304040636

Mito, H., Matsuura, N., Mukai, K., Yanagisawa, Y., Nakajima, A., Motoyama, M., … Matsunaga, H. (2014). The impacts of elevated autism spectrum disorder traits on clinical and psychosocial features and long-term treatment outcome in adult patients with obsessive-compulsive disorder. *Comprehensive Psychiatry*, *55*(7), 1526–1533.

Murray, A. L., McKenzie, K., Kuenssberg, R., & Booth, T. (2017). Do the Autism Spectrum Quotient (AQ) and Autism Spectrum Quotient Short Form (AQ-S) primarily reflect general ASD traits or specific ASD traits? A bi-factor analysis. *Assessment*, *24*(4), 444–457.

Naito, K., Matsui, Y., Maeda, K., & Tanaka, K. (2010). Evaluation of the validity of the Autism Spectrum Quotient (AQ) in differentiating high-functioning autistic spectrum disorder from schizophrenia. *The Kobe Journal of Medical Sciences*, *56*(3), E116-24.

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric Theory (McGraw-Hill Series in Psychology)* (Vol. 3). McGraw-Hill New York.

Peralta, V. M., & Cuesta, M. J. Z. (1994). Validation of positive and negative symptom scale (PANSS) in a sample of Spanish schizophrenic patients. *Actas Luso-Españolas de Neurología, Psiquiatría y Ciencias Afines*, *22*(4), 171–177.

Pilowsky, T., Yirmiya, N., Arbelle, S., & Mozes, T. (2000). Theory of mind abilities of children with schizophrenia, children with autism, and normally developing children. *Schizophrenia Research*, *42*(2), 145–155. https://doi.org/10.1016/S0920-9964(99)00101-2

Piven, J., Gayle, J., Chase, G. A., Fink, B., Landa, R., Wzorek, M. M., & Folstein, S. E. (1990). A family history study of neuropsychiatric disorders in the adult siblings of autistic individuals. *Journal of the American Academy of Child & Adolescent Psychiatry*, *29*(2), 177–183.

Piven, J., & Palmer, P. (1999). Psychiatric disorder and the broad autism phenotype: evidence from a family study of multiple-incidence autism families. *American Journal of Psychiatry*.

Pollmann, M. M. H., Finkenauer, C., & Begeer, S. (2010). Mediators of the link between autistic traits and relationship satisfaction in a non-clinical sample. *Journal of Autism and Developmental Disorders*, *40*(4), 470–478.

Rasmussen, A. R., & Parnas, J. (2015). Pathologies of imagination in schizophrenia spectrum disorders. *Acta Psychiatrica Scandinavica*, *131*(3), 157–161.

Reynolds, C. R., & Kamphaus, R. W. (2003a). Reynolds intellectual assessment scales (RIAS). *Lutz, FL: Psychological Assessment Resources*.

Reynolds, C. R., & Kamphaus, R. W. (2003b). Reynolds intellectual screening test. *Lutz, FL:*

Psychological Assessment Resources.

Rhind, C., Bonfioli, E., Hibbs, R., Goddard, E., Macdonald, P., Gowers, S., … Treasure, J. (2014). An examination of autism spectrum traits in adolescents with anorexia nervosa and their parents. *Molecular Autism*, *5*(1), 56.

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling and more. Version 0.5–12 (BETA). *Journal of Statistical Software*, *48*(2), 1–36.

Ruta, L., Mazzone, D., Mazzone, L., Wheelwright, S., & Baron-Cohen, S. (2012). The Autism-Spectrum Quotient—Italian version: A cross-cultural confirmation of the broader autism phenotype. *Journal of Autism and Developmental Disorders*, *42*(4), 625–633.

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*(1), 2.

Santamarıa, P., & Fernández Pinto, I. (2009). RIAS Escalas de Inteligencia de Reynolds y RIST Test de Inteligencia Breve de Reynolds Manual. *Adaptación Española. Madrid, Spain: TEA Ediciones*.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: uses in assessing rater reliability. *Psychological Bulletin*, *86*(2), 420.

Sizoo, B. B., van den Brink, W., Eenige, M. G., Koeter, M. W., van Wijngaarden-Cremers, P. J. M., & van der Gaag, R. J. (2009). Using the autism-spectrum quotient to discriminate autism spectrum disorder from ADHD in adult patients with and without comorbid substance use disorder. *Journal of Autism and Developmental Disorders*, *39*(9), 1291–1297. https://doi.org/10.1007/s10803-009-0743-2

Solomon, M., Olsen, E., Niendam, T., Ragland, J. D., Yoon, J., Minzenberg, M., & Carter, C. S. (2011). From lumping to splitting and back again: Atypical social and language development in individuals with clinical-high-risk for psychosis, first episode

schizophrenia, and autism spectrum disorders. *Schizophrenia Research*, *131*(1–3), 146–151. https://doi.org/10.1016/j.schres.2011.03.005

Solomon, M., Ozonoff, S., Carter, C., & Caplan, R. (2008). Formal thought disorder and the autism spectrum: relationship with symptoms, executive control, and anxiety. *Journal of Autism and Developmental Disorders*, *38*(8), 1474–1484.

Spek, A. A., & Wouters, S. G. M. (2010). Autism and schizophrenia in high functioning adults: Behavioral differences and overlap. *Research in Autism Spectrum Disorders*, *4*(4), 709–717. https://doi.org/10.1016/j.rasd.2010.01.009

Steel, R. G. D. (1960). A rank sum test for comparing all pairs of treatments. *Technometrics*, *2*(2), 197–207.

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*(3), 224–228.

Takara, K., & Kondo, T. (2014). Autism spectrum disorder among first-visit depressed adult patients: diagnostic clues from backgrounds and past history. *General Hospital Psychiatry*, *36*(6), 737–742. https://doi.org/10.1016/j.genhosppsych.2014.08.004

Team, R. C. (2013). R: a language and environment for statistical computing. Version 3.0. 1. R Foundation for Statistical Computing. Vienna, Austria.

Wakabayashi, A., Baron-Cohen, S., Wheelwright, S., & Tojo, Y. (2006). The Autism-Spectrum Quotient (AQ) in Japan: a cross-cultural comparison. *Journal of Autism and Developmental Disorders*, *36*(2), 263–270.

Wheelwright, S., Auyeung, B., Allison, C., & Baron-Cohen, S. (2010). Defining the broader, medium and narrow autism phenotype among parents using the Autism Spectrum Quotient (AQ). *Molecular Autism*, *1*(1), 10. https://doi.org/10.1186/2040-2392-1-10

Wikramanayake, W. N. M., Mandy, W., Shahper, S., Kaur, S., Kolli, S., Osman, S., … Fineberg, N. A. (2018). Autism spectrum disorders in adult outpatients with obsessive

compulsive disorder in the UK. *International Journal of Psychiatry in Clinical Practice*, *22*(1), 54–62.

Wouters, S. G. M., & Spek, A. A. (2011). The use of the Autism-spectrum Quotient in differentiating high-functioning adults with autism, adults with schizophrenia and a neurotypical adult control group. *Research in Autism Spectrum Disorders*, *5*(3), 1169–1175. https://doi.org/10.1016/j.rasd.2011.01.002

Yu, C.-Y. (2002). *Evaluating cutoff criteria of model fit indices for latent variable models with binary and continuous outcomes* (Vol. 30). University of California, Los Angeles Los Angeles.

Figure 1. Factor structure of the AQ-Short, including factor correlation and factor loadings (N=294; all subjects)

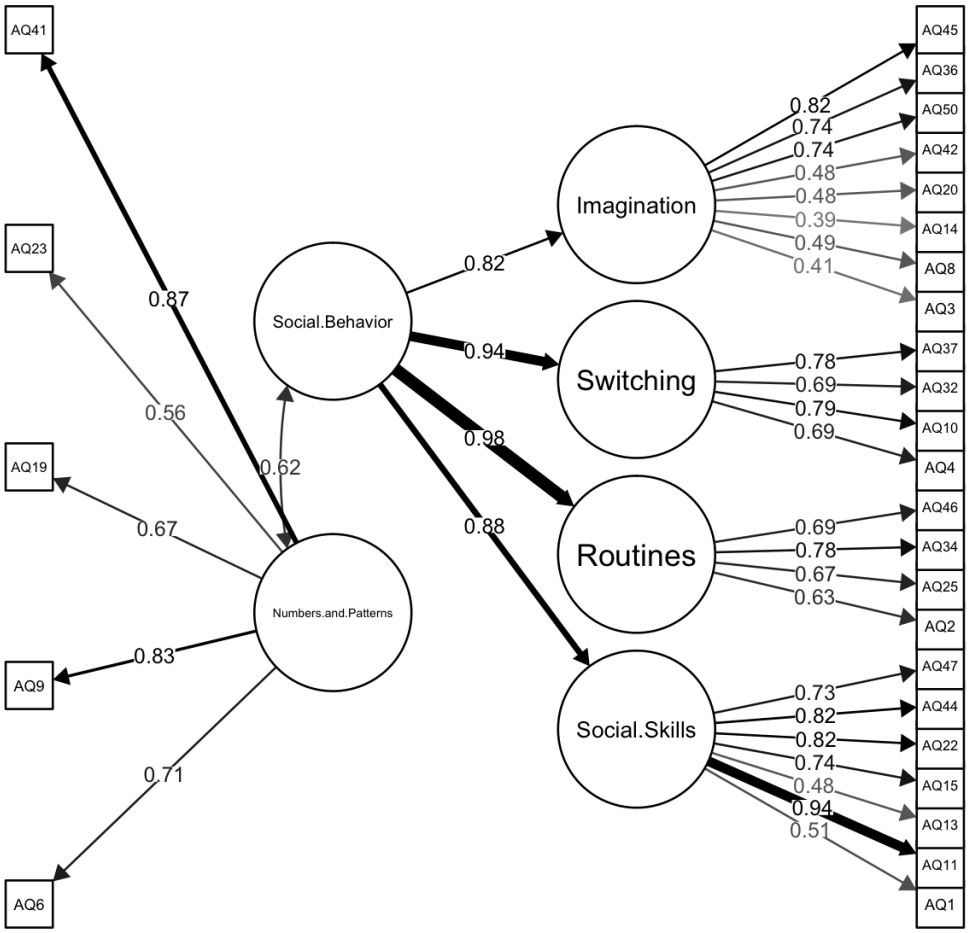
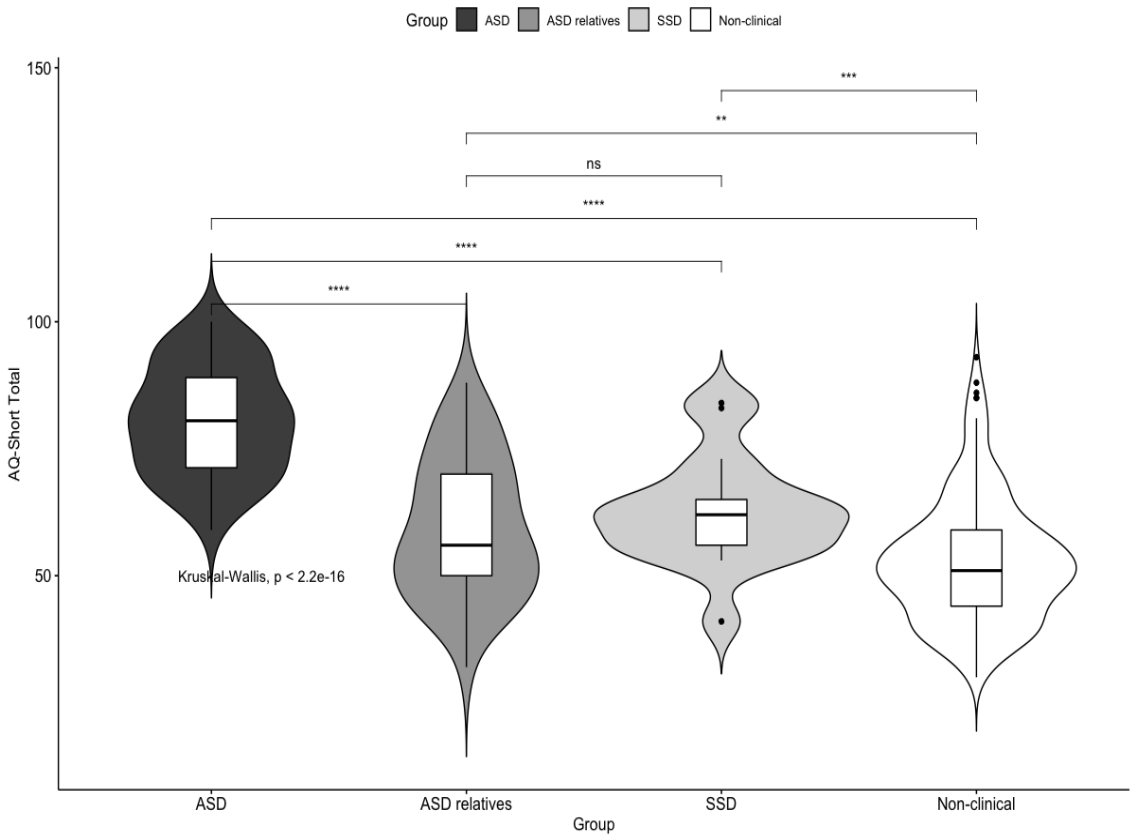


Figure 2. AQ total inter-group differences (non-parametric one-way ANOVA Kruskal-Wallis)



* $p < .05$ ** $p < .01$ *** $p < .001$

Figure 3. Scatter plot with line of best fit (95% confidence interval) correlation analysis of AQ-Short and ADOS-2 score by group.

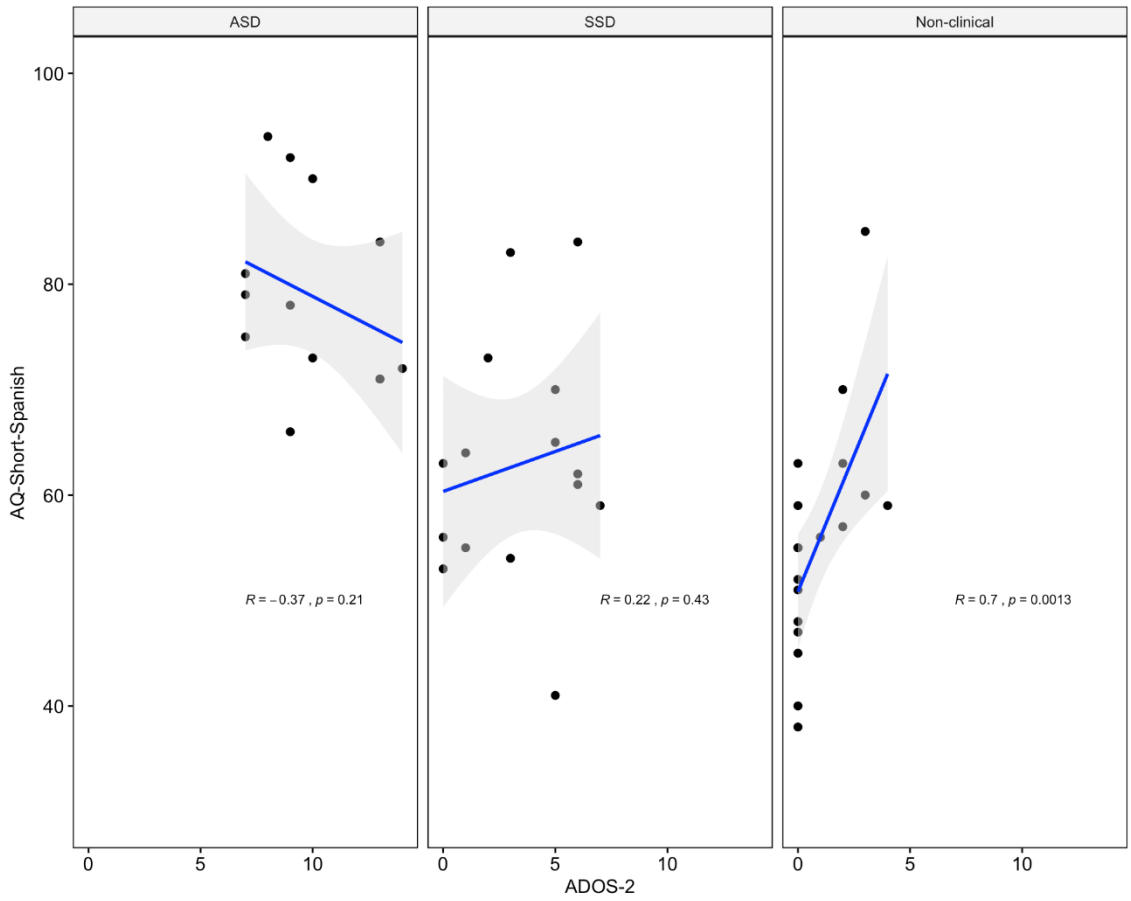


Figure 4. Group prediction-accuracy derived from the linear discriminant analysis (all AQ-Short subscales as predictors). Diagonal shows the number of observed cases correctly predicted by the model as belonging to the corresponding group.

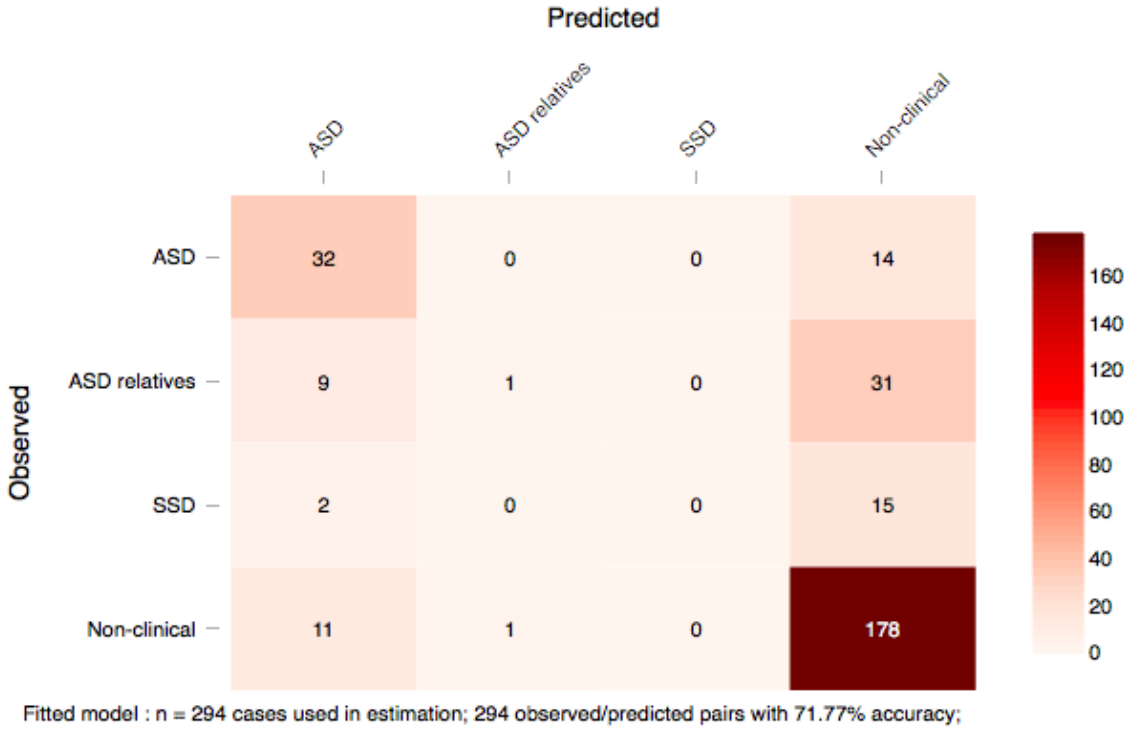
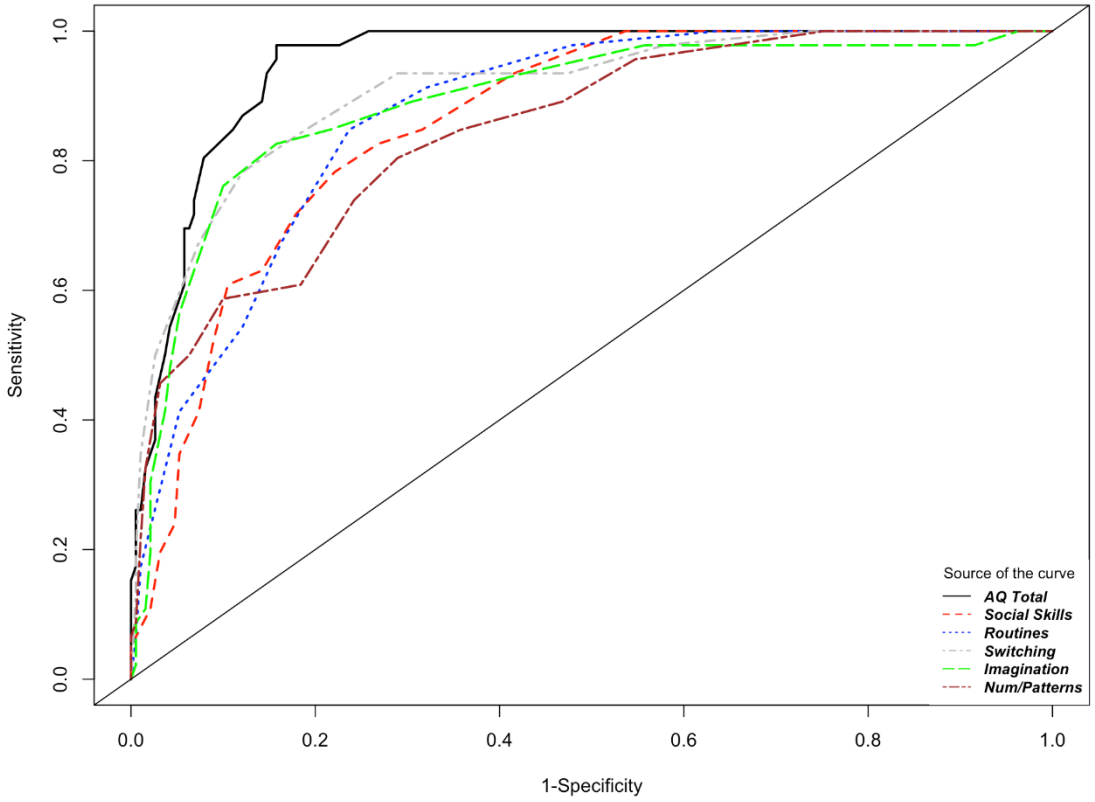


Figure 5. Receiver operating characteristic (ROC) curves illustrating the ability of the full AQ scale and AQ domains to identify any ASD cases at alternative cut-off points.



N = 236 (ASD versus Non-Clinical groups)

Table 1. Characteristics of the participants in the study and differences

	Group[1]				Differences[2]
	ASD (n = 46)	ASDR (n = 41)	SSD (n = 17)	NC (n = 190)	
Age [mean (sd)]	31.81 (9.11)	46.64 (9.73)	34.36 (9.28)	34.28 (10.11)	ASDR > (ASD, SSD, NC) **
IQ mean T scores (sd)					
N of IQ calculation	24	0	17	18	
Verbal	47.08 (13.8)	-	50.47 (10.5)	52.06 (10.8)	
Non-verbal	46.33 (12.1)	-	51.29 (9.2)	56.17 (7.5)	ASD< NC **
Full-IQ	93.75 (16)	-	101.53 (13.5)	105.56 (13.6)	ASD < NC *
Gender					
Male	63,0%	26,8%	82,4%	22,6%	
Female	37,0%	73,2%	17,6%	77,4%	
Education level					
Primary	2,2%	2,4%	11,8%	,5%	
Secondary	10,9%	2,4%	11,8%	3,2%	
Bachelor	23,9%	14,6%	35,3%	18,4%	
Professional Training	26,1%	29,3%	29,4%	9,5%	
University	28,3%	51,2%	11,8%	68,4%	
Work/Academic status					
Student	28,3%	0,0%	41,2%	22,1%	
Worker	19,6%	68,3%	11,8%	57,9%	
Student+Worker	8,7%	2,4%	0,0%	7,9%	
Unemployed	30,4%	12,2%	17,6%	6,8%	
Home worker	2,2%	7,3%	0,0%	4,2%	
Pensioner	2,2%	2,4%	29,4%	,5%	
Retired	0,0%	4,9%	0,0%	,5%	
Psychofarmacological treatment					
Any psychofarmacological treatment	28,3%	9,8%	88,2%	0,0%	
Antipsychotic	10,9%	0,0%	82,4%	0,0%	
Antidepressant	13,0%	4,9%	29,4%	0,0%	
Anxiolytic	13,0%	4,9%	35,3%	0,0%	
Hypnotics	2,2%	0,0%	29,4%	0,0%	
Mood stabilizer	0,0%	2,4%	17,6%	0,0%	
Metiylphenidate	2,2%	0,0%	0,0%	0,0%	
Psychiatric disorders					
Any psychiatric disorder (other than ASD)	58,7%	39,0%	100,0%	0,0%	
Substance Use Disorders	2,2%	0,0%	0,0%	0,0%	
Schizophrenia Spectrum Disorders	4,3%	0,0%	100,0%	0,0%	
Mood Disorders	17,4%	22,0%	17,6%	0,0%	
Anxiety Disorders	19,6%	34,1%	17,6%	0,0%	
Eating Disorders	6,5%	0,0%	0,0%	0,0%	
Personality Disorders	4,3%	0,0%	5,9%	0,0%	
Attention Deficit/Hyperactivity Disorder	17,4%	0,0%	0,0%	0,0%	

1: ASD: Autism Spectrum Disorder; ASDR: ASD relatives; SSD: schizophrenia spectrum disorders NC: Non-Clinical

2: Based on Dwass-Steel-Critchlow-Fligner pairwise comparisons: * $p < .05$; ** $p < .01$

Table 2. Internal consistencies (Cronbach's Alpha) for all five subscales of the AQ-short.

AQ scales	ASD (n = 46)	ASD relatives (n = 41)	SSD (n = 17)	Non-clinical (n = 190)
Social Behavior (23 items)	.78	.88	.78	.88
Social Skills (7 items)	.76	.79	.82	.83
Routines (4 items)	.42	.66	.45	.70
Switching (4 items)	.56	.61	.41	.68
Imagination (8 items)	.63	.71	.33	.64
Numbers/Patterns (5 items)	.72	.74	.64	.75
AQ-Short Total (28 items)	.79	.88	.80	.88

Table 3. Independent samples t-tests, means, standard deviations, Spearman rho correlation and intraclass correlation coefficients for the test-retest reliability of the AQ and its subscales for ASD (n = 26), SSD (n = 9) and Non-clinical (n = 61) participants.

AQ scales	Mean T1 (sd)	Mean T2 (sd)	*t*	*p*	r_s	ICC
Social Behavior	50.33 (13.8)	50.32 (15.1)	0.02	.99	.94***	.97***
Social Skills	15.34 (5.4)	15.28 (5.5)	0.28	.78	.91***	.96***
Routines	9.55 (3.1)	9.66 (3.2)	-0.53	.60	.82***	.90***
Switching	9.19 (3.2)	9.13 (3.2)	0.33	.74	.81***	.91***
Imagination	16.25 (4.5)	16.26 (5.2)	-0.04	.97	.81***	.91***
Numbers/Patterns	10.48 (4.1)	10.50 (4.1)	-0.09	.92	.83***	.93***
AQ-Short Total	60.81 (16.3)	60.82 (18.1)	-0.02	.99	.94***	.97***

T1 = time one, T2 = time two

*** $p < .001$

Table 4. Fit values based on WLSMV extractions for second-order Social Skills + Numbers/Patterns AQ-S model

Sample	**WLSMV χ^2**	**df**	**RMSEA (90% CI)**	**CFI**	**TLI**	**WRMR**
All subjects (N = 294)	828.94**	345	0.069 (.063 - .075)	.942	.937	1.293
Non-clinical (N = 231)	671.99**	345	0.064 (.057 - .071)	.913	.905	1.211

** p < .001; WLSMV = robust weighted least squares; RMSEA=root mean square error of approximation; CFI = comparative fit index; TLI = Tucker-Lewis Index; WRMR = weighted root mean square residual.

Table 5. Independent samples t-tests means and standard deviations of the AQ total score and its subscales for all groups.

	Group[1]				
AQ scales	**ASD (n = 46)**	**ASDR (n = 41)**	**SSD (n = 17)**	**NC (n = 190)**	**Differences[2]**
Social Behavior	66.37 (9.2)	49.80 (11.9)	52.29 (9.4)	43.84 (10.8)	ASD > (ASDR, SSD,NC)** SSD > NC ** ASDR > NC **
Social Skills	20.28 (4.3)	15.22 (4.5)	15.12 (4.9)	13.59 (4.5)	ASD > (ASDR, SSD,NC)** ASDR > NC **
Routines	12.61 (2)	9.41 (2.8)	9.76 (2.4)	8.62 (2.7)	ASD > (ASDR, SSD,NC)**
Switching	12.33 (2.5)	8.54 (2.6)	9.65 (2.4)	7.38 (2.5)	ASD > (ASDR, SSD,NC)** SSD > NC ** ASDR > NC **
Imagination	21.15 (4.3)	16.63 (4.7)	17.76 (3.2)	14.25 (3.7)	ASD > (ASDR, SSD,NC)** SSD > NC ** ASDR > NC **
Numbers/Patterns	14.09 (3.8)	9.44 (3.5)	10.29 (3.2)	8.95 (3.2)	ASD > (ASDR, SSD,NC)**
AQ Total	80.46 (10.6)	59.24 (13.7)	62.59 (10.7)	52.79 (12.2)	ASD > (ASDR, SSD,NC)** SSD > NC ** ASDR > NC **

1: ASD: Autism Spectrum Disorder; ASDR: ASD relatives; SSD: schizophrenia spectrum disorders NC: Non-Clinical

2: Based on Dwass-Steel-Critchlow-Fligner pairwise comparisons: * $p < .05$; ** $p < .01$

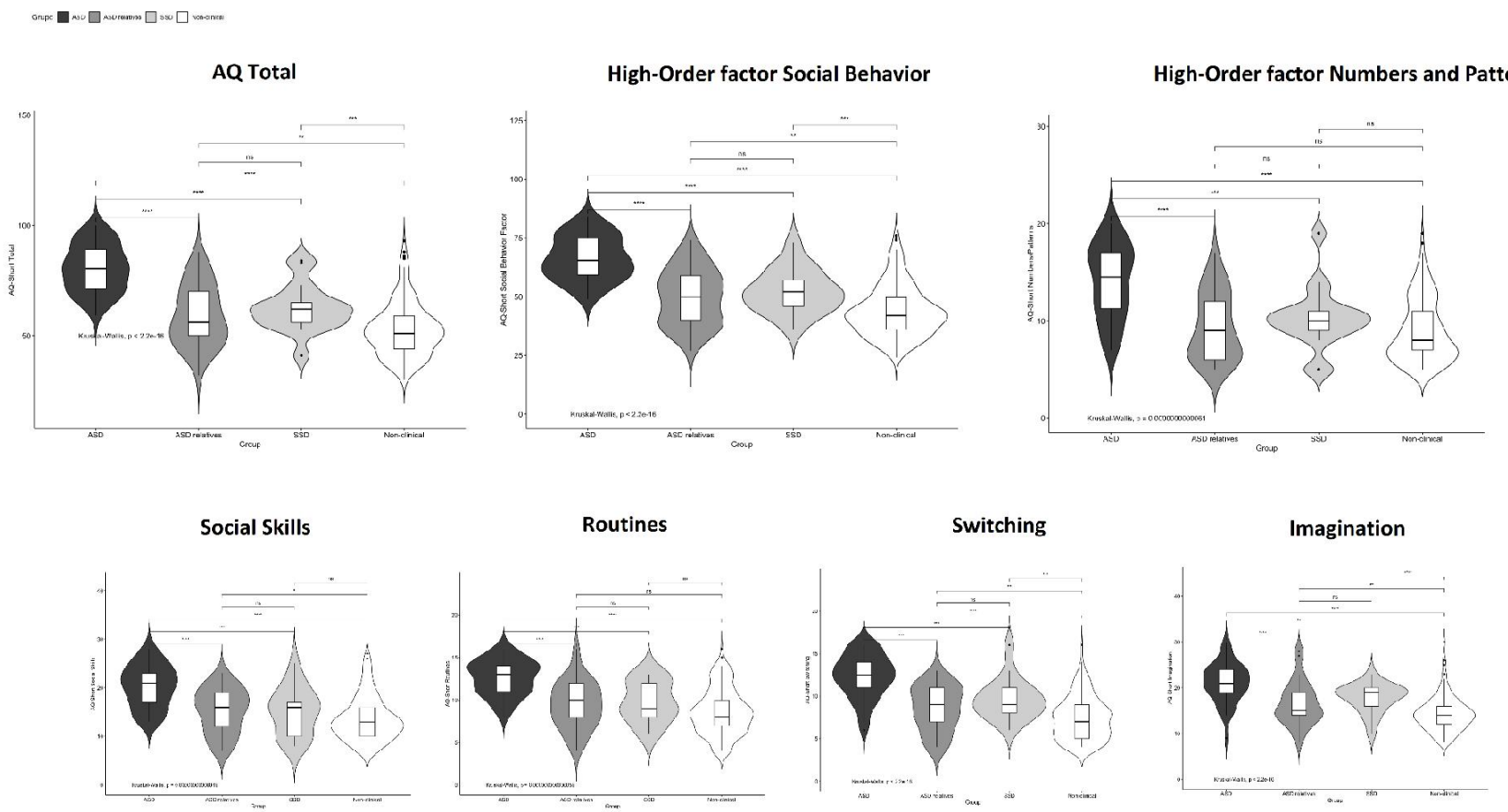
Table 6. Area under the curve (AUC), accuracy, sensitivity and specificity, positive (PPV) and negative (NPV) predictive values for the Youden index-based cut-off scores for full scale and five domains

AQ Scale	Cut-off	AUC (95% CI)	Accuracy (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)
AQ Total	63.5	0.948 (0.948-0.974)	0.869 (0.868-0.87)	0.978 (0.936-1.02)	0.842 (0.79-0.894)	0.6 (0.489-0.711)	0.994 (0.982-1.006)
Social Skills	16.5	0.859 (0.859-0.91)	0.78 (0.778-0.781)	0.783 (0.663-0.902)	0.779 (0.72-0.838)	0.462 (0.351-0.572)	0.937 (0.899-0.975)
Routines	10.5	0.87 (0.87-0.918)	0.78 (0.778-0.781)	0.848 (0.744-0.952)	0.763 (0.703-0.824)	0.464 (0.358-0.571)	0.954 (0.921-0.987)
Switching	10.5	0.908 (0.908-0.955)	0.86 (0.859-0.861)	0.783 (0.663-0.902)	0.879 (0.833-0.925)	0.61 (0.486-0.735)	0.944 (0.909-0.978)
Imagination	17.5	0.889 (0.889-0.945)	0.839 (0.838-0.84)	0.826 (0.717-0.936)	0.842 (0.79-0.894)	0.559 (0.441-0.677)	0.952 (0.92-0.985)
Numbers/Patterns	10.5	0.844 (0.844-0.904)	0.729 (0.727-0.73)	0.804 (0.69-0.919)	0.711 (0.646-0.775)	0.402 (0.302-0.502)	0.938 (0.898-0.977)

N = 236 (ASD versus Non-Clinical groups)

Appendix A

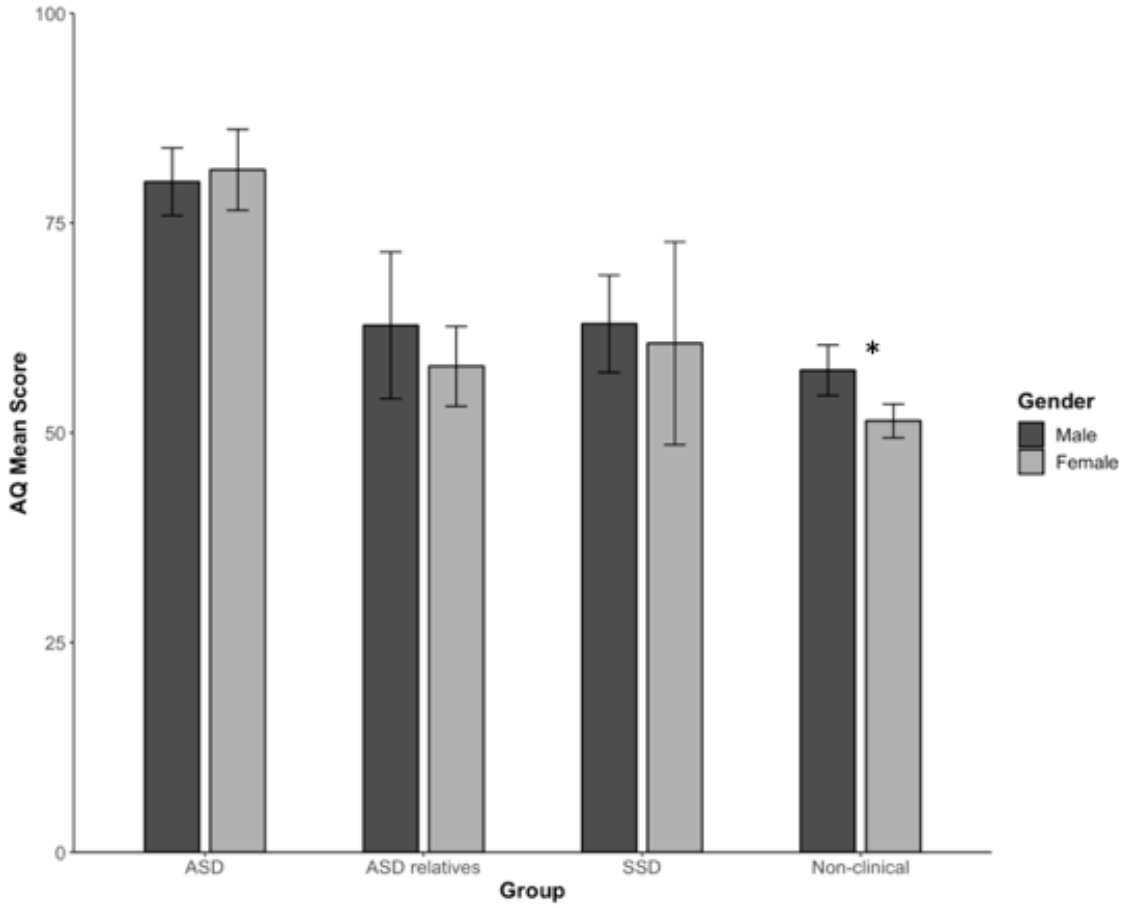
Figure 1. Violin plots for AQ inter-group differences (non-parametric one-way ANOVA Kruskal-Wallis)



* $p < .05$ ** $p < .01$ *** $p < .001$

Appendix B

Figure 3 shows the mean AQ-Short total scores for the four groups sorted by gender.



* $p < .05$

Differential item functioning (DIF) was used to examine whether an item performed differently for the male group than for the female group. Hybrid ordinal logistic regression has shown good power for detecting DIF, but inflated type I error rates (i.e., items with very small DIF valued as DIF items.) have also been reported with large samples. For this reason, DIF was considered present considering the likelihood ratio χ^2 tests (statistical significance) in conjunction with effect size measures (pseudo R^2 statistic >.20) and proportional β_1 change >10%.

Nine items were flagged for DIF (χ^2 significance as criteria), but none of those items also met the other two stablished criteria. As shown in Figure 1, a difference in test characteristic curves (TCC) between males and females can be seen for DIF items, however, as can be seen in the right figure of the panel, the absolute magnitude of the difference is very small and mainly due to a decrease in the probability of response for females at the lower level of the lower extreme ability level and an increase in the probability of females at the upper extreme ability level.

Figure 1. Impact of gender DIF on test characteristics curves (TCC)