The Autism Spectrum Quotient-Revised: A measure to better identify the Autism Spectrum Disorder presentation in females?

Alice Macoun

Supervisor: Professor Richard O'Kearney

January 2018

A thesis submitted in partial fulfilment of the requirements for the degree of Master of Clinical Psychology at The Australian National University

Except where due reference is made in the text, this thesis is my own original work. It has not been submitted for any other degree, or diploma at any university.

Acknowledgements

Following an intensive two-year period, today is the day: writing this page of acknowledgements is the finishing touch to my research paper. The topic of the paper is one that I have been passionate about for some time, so I am pleased to have had the opportunity to conduct this research and now finally share my work. I wish to reflect on those who have encouraged and supported me throughout the period in which this project was conceived, developed and finally came to fruition.

First and foremost, I would like to express my gratitude to those who participated in the study, including the autism practitioners in phase one of the study. I was overwhelmed by the final number of responses received and encouraged by the comments made by the practioners regarding the need for a study of this nature. A special thanks also goes to those who distributed the questionnaire link amongst work colleagues – Ben, Ian, Meg and Matt.

Second, I must express my appreciation to my research supervisor, Professor Richard O'Kearney, for believing in this project, particularly when others had told me it was of a PhD scope, and for being available to guide and reassure me whenever I encountered challenges. Beyond this project, thank you also for your advice in my navigation of post-university career decisions. Thanks also to Boris Bozumic and Conal Monaghan for advice on factor analysis.

Warmest thanks to my friends and fellow students for your support. In particular, Anita, thank you for your sympathetic ear and ongoing friendship despite distance.

Finally, thank you to my family and my partner. In particular, thank you to my parents for being the role models (in different ways) that you are and equipping me with a perfect writing environment, as well as early morning walks and coffee – a key recipe to success; and to my special partner, Yianni, you have been a tremendous support over the last two years. For your rationality, commitment, patience, sensitivity and unwavering belief in me, I cannot thank you enough.

The Autism Spectrum Quotient-Revised: A measure to better identify the Autism Spectrum Disorder presentation in females?

Alice Macoun[1] and Richard O'Kearney[1]

[1] Research School of Psychology, The Australian National University, Acton ACT 2601

Author for Correspondence: Alice Macoun, Research School of Psychology, The Australian National University, Acton ACT 2601, Australia; Phone: + 61 409 172 116; Email: alice.macoun@anu.edu.au

Funding: No research grants were received for the research.

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

Informed Consent: Informed consent was obtained from all individual participants included in the study.

Abstract

Literature reflects a growing concern that females may be "slipping through the net" because current understanding of the autistic phenotype and diagnostic approaches are derived primarily from research on males. The present study sought to develop a psychometrically sound revision of the Autism Spectrum Quotient (AQ), to improve its sensitivity in identifying adult females with high levels of Autism Spectrum Disorder (ASD) traits. In phase one of the study, 30 experienced ASD practitioners rated how relevant they thought 20 newly developed screening items were to ASD in females. Fifteen of these items were then added into the existing AQ, comprising the AQ-Revised. Alongside measures assessing convergent and divergent validity and levels of distress and impairment, 1007 non-ASD (55% female) and 45 ASD (51% female) participants completed the AQ-Revised. The AQ-Revised and the AQ were found to be psychometrically sound and both fit a 2-factor structure, based on factor analyses. No significant gender differences were apparent in non-ASD participants on the total AQ-Revised, however in the ASD group females scored significantly higher on the total AQ-Revised and AQ scores compared to males. The AQ-Revised was found to discriminate well between ASD and non-ASD cases and, with a cut off score of 29, captured 130 more females high on ASD traits than the existing AQ. These women scoring above 29 on the measure were found to have significantly higher levels of distress and functional impairment than those scoring below 29, but reported levels of distress and impairment were equivalent to women scoring above the recommended AQ cut off (32). The utility of the revised instrument for females and the concept of gender-specific ASD screening are discussed.

Keywords: AQ; Autism Spectrum Disorder (ASD); Autism; Gender; Females; Screening

Author for email correspondence: Alice Macoun at alice.macoun@anu.edu.au

Autism Spectrum Disorder (ASD) is a neurodevelopmental disorder characterised by persistent deficits in social communication and interaction, as well as restricted, repetitive patterns of behaviour, interests, or activities, which cause interference to one's ability to function (DSM-5, American Psychiatric Association, 2013). While significant impairment in these two domains must be present in order for diagnostic criteria to be met, there is growing evidence that the severity of autistic dimensions lies on a continuum ranging from severe whereby individuals require "very substantial support" (APA, 2013) to low, where the degree of impairment is insufficient for diagnosis (Baron-Cohen, 1995; Hoekstra, Bartels, Verweij, & Boomsma, 2007). This suggests heterogeneity of autistic symptomology across individuals with ASD and in the general population.

Because this heterogeneity presents challenges to the conceptualisation and assessment of an autistic syndrome or disorder, it is important to identify factors that may be associated with variability in ASD features. One heterogeneous feature of the autism spectrum is the marked sex discrepancy. Current diagnostic practices support a strong preponderance of males with epidemiological studies reporting an average male to female ratio for an ASD diagnosis of 4.3:1, which increases to 10:1 in high functioning autism (HFA; previously referred to as "Asperger's Disorder") and reduces to 1.95:1 among ASD individuals with a comorbid intellectual disability (ID; Fombonne, 2003; 2005; 2007; 2009). Statistics suggest that the frequency of ASD is 1% of the population (APA, 2013) with males being 3 times more likely than females to have the condition (Loomes, Hull, & Mandy, 2017). General population questionnaires measuring autistic traits have also consistently described the greater presence of autistic characteristics in males compared to females (Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001; Posserud, Lundervold, & Gillberg, 2006; Allison, et al., 2008; Williams et al., 2008; Constantino & Todd, 2003; Skuse et al., 2009). While the mechanisms underlying this sex distinction are not well understood,

there is growing evidence for a heterogeneity in autistic features that is associated with sex. There is recent recognition that this association may contribute to a bias towards the male phenotype in the identification of ASD (Goldman, 2013) with the DSM-5 acknowledging that females with ASD "without accompanying intellectual impairments or language delays may go unrecognised, perhaps because of subtler manifestations of social and communication difficulties" (APA, 2013, p. 57).

An emerging evidence base supports the existence of a female ASD phenotype; a composition of autistic traits, which has an imperfect fit with the conventional presentation and pathogenesis of ASD, derived primarily from research on males (Bargiela, Steward, & Mandy, 2016; Mandy et al., 2011; Lai et al., 2015). This evidence indicates sex differences in ASD across the areas of social communication difficulties; repetitive and restricted interests; and mood and behavioural characteristics. It should be noted that the longstanding underrepresentation of females in research, as well as clinical practice, has likely generated a male-biased understanding of ASD, thereby biasing the literature reviewed below. Further, it is likely that the females who participate in the research are those with more obvious cases, whose traits are more disruptive to the home, educational or occupational settings, thereby potentially further biasing the literature. Given the inverse relationship between autistic symptomology and ID (Skuse et al., 2009; Bartak & Rutter, 1976), and thus the possible confounding effect of intelligence, most of the literature examining sex differences in ASD focuses on HFA participants.

Social and Communication Differences

McLennan, Lord and Schloper (1993) examined sex differences in a sample of 21 females and 21 males with HFA. They found that while males experienced greater separation anxiety and impairments in reciprocal social interaction and communication prior to the age of 5, females experienced greater severity in social deficits during the adolescent period.

Further, in a similarly designed study, Holterman, Bolte, and Poustka (2007) reported that females experience more social problems, as well as more attention and thought problems compared to males. However, such differences were found in parent reports alone and, thus, susceptible to interpreting bias on the parents' behalf, who may have held certain expectations regarding socially desirable behaviour from daughters in contrast to sons. Knickmeyer, Wheelwright, and Baron-Cohen (2008) also found that females with ASD between the ages of 4 and 14 engaged in more pretend play than their male counterparts, suggesting superior imaginative play skills.

Further, in a qualitative analysis, which utilised Framework Analysis to examine the experience of females on the autism spectrum, Barigela, Steward, and Mandy (2016) found that, when in social situations, in an effort to hide their autistic traits and appear "normal", the women wore a "mask" or took on a certain "persona". However, the constant maintenance of the mask frequently resulted in exhaustion or confusion over their identity. The women also described actively learning appropriate social etiquette through different media sources. Referred to as the "Camouflage Hypothesis", this concept of females "camouflaging" their limitations in social understanding by observing and mimicking others is one of the most frequently discussed hypotheses to account for females being overlooked in the current diagnostic system (e.g., Dworzynski, Ronald, Bolton, & Happe, 2012; Kopp & Gilberg, 1992; Lord et al., 2000; Mandy et al., Bargiela et al., 2016; Ehlers & Gillberg, 1993; Gillberg, 1993). Despite the challenges experienced in social interaction, many authors (e.g., Barigela et al., 2016; Bauminger & Kasari, 2000; Attwood, 2007; Kopp & Gillberg, 1992) highlight that females with ASD do in fact desire friendship and, at times, feel lonely.

Repetitive and Restricted Interests

A number of studies have investigated sex differences in restricted interests and play. For example, in a study of 384 ASD males and 91 Females with ASD between the ages of 3

and 8 years, Lord et al. (2000) found that males engaged in higher rates of repetitive or restricted behaviours and inappropriate play compared to females, when IQ was controlled for. Nicholas et al. (2008) also found that ASD males had greater preoccupation with parts of objects, routines and rituals, as well as stereotyped mannerisms. Such results are supported by experienced clinicians and qualitative studies of females on the autism spectrum, including those conducted by Barigela, Steward, and Mandy (2016) and Baldwin and Costley (2016), who suggest that the special interests typically found in Females with ASD are more aligned with social and gender norms than corresponding interests of ASD males. Commonly reported interests in females on the spectrum include animals, nature, celebrities, fiction franchises, the arts and classical literature (Gould & Ashton-Smith, 2011; Attwood et al., 2006). Further, in a study conducted by Kopp and Gillberg (1992), which described six cases of HFA females aged between 6 and 10 years, all individuals displayed "atypically excessive" patterns of language use, including repetitive questioning of others and echolalia.

Mood and Behavioural Characteristics

Few studies have examined sex differences in the secondary manifestations of ASD, including in mood and behavioural characteristics. However, in one study by Hartley and Sikora (2009), female toddlers exhibited greater difficulties with sleep and anxious and depressed affect than their male counterparts (Hartley & Sikora, 2009). Consistent with this finding, also in a sample of ASD toddlers, Carter et al. (2007) reported a trend towards atypical depression and withdrawal in females compared to males. Further, in their study of children and adolescents with HFA, Mandy et al. (2011) found that females displayed greater internalising difficulties compared to males, who are reported to be typically outwardly aggressive when angry or stressed. However, this effect was found in the parent report only and was not apparent in the teacher report. Such findings are supported by Bargiela et al.'s. (2016) qualitative study with HFA female adults, which found 13 out of the 14 participants

reported clinically severe anxiety and elevated levels of distress. Regarding female presentation in the school environment Attwood et al. (2006) described Females with ASD as passive, shy and compliant, a description supported by females interviewed in Bargiela et al.'s (2016) study who outlined the negative outcomes of their perceived passivity in adulthood. Further, a number of studies have indicated a higher incidence of abnormal eating patterns, including comorbid Anorexia Nervosa, in Females with ASD in comparison to their male counterparts and neurotypical peers (Råstam, 2008; Gillberg, Råstam, Wentz & Gillberg, 2007; Zucker et al., 2007). While the literature in this specific area remains scarce, researchers have suggested that the abnormal eating patterns found in Females with ASD may be ingrained as a form of repetitive stereotyped behaviour (Hackler, 1986; Råstam, 2008).

While many studies report sex differences, albeit commonly subtle differences, there are other studies which have not found any evidence in support of a contrasting male and female ASD phenotype (Wing & Gould, 1979; Volkmar, Szatmari, Sparrow & 1993; Baron-Cohen et al., 2006, 2007; Wakabayashi et al., 2004, 2007; Tsai & Beisler, 1983; Pilowsky, Yirmiya, Schulman, & Dover 1998). That said, many of these earlier studies failed to control for sex discrepancy in intellectual functioning. Thus, sex difference in intellectual functioning may have obscured the true difference in ASD symptomology between males and females. It should be noted that while the difference between the terms is acknowledged, the terms "sex" and "gender" are used synonymously throughout the literature on this topic. "Gender" is used for the purpose of this study.

Implications for Identification of ASD

One important implication of these findings is that a differing female phenotype and developmental profile may be overlooked during the screening and diagnostic process, especially if clinicians approach evaluation with their own ASD expectations regarding the

association with gender (Goldman, 2013). Literature outlines that females on the high functioning end of the spectrum tend to be referred, and clinically identified, at a later age and the time frame from the initial presentation to a health care professional to formal diagnosis is significantly longer for females compared to their male counterparts (Giarelli et al., 2010; Siklos & Kerns, 2007). Females are also reported to be at a considerably elevated risk of their ASD going undiagnosed, with their difficulties missed entirely or mislabelled (Lai & Baron-Cohen, 2015). For example, females are frequently provided with other diagnoses prior to ASD, including Attention Deficit Hyperactivity Disorder, developmental delay, anxiety, Anorexia Nervosa, minimal brain dysfunction, and speech and motor concerns (Kopp & Gillberg, 1992, 1997; Nilsson, Gillberg, Gillberg, & Rastam, 1999; Miniscalco & Sanburg, 2010). Further, in a study by Dworzynski et al. (2012), females scoring highly on screeners for autistic traits were less likely than equivalent scoring males to meet clinical diagnostic criteria for autism. Failure to identify females on the spectrum can have serious implications for adult females including social exclusion and isolation, unmet support needs in education and occupational settings, high levels of mental health disorders and missed opportunity for identification with the autism community (Baldwin & Costley, 2016; Bargiela et al., 2016; Haney, 2016). In some cases, such challenges are reported to be compounded by the internal turmoil and distress of struggling to understand why for their entire childhood and adult life they have felt “different” (Baldwin & Costley, 2016; Haney, 2016; Dworzynski et al., 2012). With potential to mitigate the negative implications for females by earlier detection, appropriate, timely and gender sensitive ASD screening is, therefore, of crucial importance.

While a number of screening instruments designed to identify ASD are available, the value of these tools in the identification of Females with ASD is questionable because, like diagnostic criteria, most appear to have been developed based upon the traditional male

phenotype (Andersson, Gillberg, & Miniscalco, 2013; Haney, 2016). Only one study, to date, has attempted to address this issue. In a sample of 190 children aged between 3 and 18, Kopp and Gillberg (2011) identified and evaluated 18 new screening items, which they deemed would better capture the female phenotype of ASD. They added these items to the existing parent-rated screener for HFA children, the Autism Spectrum Screening Questionnaire (ASSQ; Ehlers, Gillberg, & Wing, 1999). Their preliminary findings suggested that certain single items on their revised measure were more frequently endorsed in girls with ASD than in boys. Specifically, compared to males, the parents of Females with ASD tended to respond affirmatively to the following four items: "difficulties completing daily activities"; "has a different voice or speech"; "interacts with mostly younger children"; and "avoids demands". Other highly predictive items included: "very determined" and "careless with regards to physical appearance and dress", a characteristic also frequently observed by experienced clinicians in the field (Attwood et al., 2006). While the study could not confirm that the score of the revised ASSQ better identified females, it made a significant contribution to the understanding of how the clinical phenotype of a girl with ASD may present. Kopp and Gillberg (2011) highlighted parent ratings of autistic symptomology, and potential associated interpreting bias, as a limitation of their study and recommended the development of specific gender-based assessment instruments as an important goal for future research, a recommendation shared by many researchers in the field (Baron Cohen et al., 2009; Posserud, Lundervold, Steijnen, Verhoeven, Stormark, & Gillberg, 2008; Williams et al., 2008).

With the insight provided by Kopp and Gillberg's (2011) study and the findings of the reviewed literature above, the present study sought to revise the Autism Spectrum Quotient (AQ; Baron-Cohen, Wheel-wright, Skinner, Martin, & Clubley, 2001), a commonly used measure of autistic traits in adults, by adding in new items deemed to be more sensitive to

females. As many Females with ASD report that others do not understand them, it is proposed that, being a self-report measure, the revised AQ will be less susceptible to interpreting bias and, therefore, may better capture the true experiences of females on the spectrum.

The AQ (Baron-Cohen et al., 2001) was originally developed as a quick and quantitative self-report measure for assessing autistic spectrum traits in adults of normal intelligence in both the autism community and general population. Since its development, it has been well validated in research, both in clinical samples (Baron-Cohen, et al., 2001) and the general population (Broadbent, Galic, & Stokes, 2013; Kloosterman, Keefer, Kelley, Summerfeldt, & Parker, 2011), and shown to be a reliable and valid ASD screening questionnaire in clinical practice (Woodbury-Smith, Robinson, Wheelwright, & Baron-Cohen, 2005). During the AQ's development, however, the factor structure was not empirically tested, and validity studies since have not come to a consensus on the best-fitting model; a 2-factor hierarchical structure (Hoekstra et al., 2008), 3-factor (Austin, 2005), 4-factor (Stewart & Austin, 2009), and 5-factor structure (Baron-Cohen et al., 2001) have each been proposed. Baron-Cohen et al. (2001) proposed a differentiation cut-off score of 32 and above as indication that an individual may be on the spectrum, therefore warranting a full diagnostic assessment. However, other studies including that with an Australian population (Broadbent et al., 2013) and another with a clinical population (Woodbury-Smith, et al., 2005) have proposed more conservative cut-off scores of 29 and 26 respectively. While there has been no clear indication that the AQ is sensitive to the male ASD presentation alone, males frequently score significantly higher on the measure compared to females (e.g., Baron-Cohen et al., 2001; Baron-Cohen et al., 2006; Auyeung, Baron-Cohen, Wheelwright, & Allison, 2008). Further, the need for more "gender specific" items for females in ASD screening instruments, or alternatively different gender based cut-off scores or norms, to

better capture the subtle difficulties that females face, has been made clear (Baldwin & Costley, 2016; Bell, Foster, & Mash, 2005; Rutter, Caspi, & Moffitt, 2003). To the author's knowledge, there has been no reported attempt to undertake this task with an adult population.

Aims

Against this background, the present study aims to: (1) refine and present a revised version of the AQ (the Autism Spectrum Quotient-Revised) which encompasses the new items that are proposed to be more sensitive to the female ASD phenotype than those included in the original AQ; (2) to test the psychometric properties, including internal consistency, convergent and divergent validity, and factor structure of the AQ-Revised for males and females using exploratory and confirmatory factor analyses to determine whether the new items fit within the existing AQ factor structure; (3) analyse, via gender separate item comparisons, whether the newly added items will be endorsed more often in females than in males, suggesting that they are effective in tapping into the female ASD phenotype; (4) evaluate the ability of the AQ-Revised to distinguish between ASD and non-ASD cases and generate an appropriate cut off score for the AQ-Revised using a Receiver Operating Curve (ROC); (5) determine whether the AQ-Revised outperforms the AQ in identifying females high on autistic traits; and (6) determine whether there are females without an ASD diagnosis who score highly on the AQ-Revised, Impact Scale and K-6, but are not picked up by the AQ, thereby indicating the existence of individuals in the population who are evading a diagnosis yet experiencing difficulty in functioning and heightened levels of distress associated with ASD symptoms (Barigiela et al., 2016; Baldwin & Costley, 2016).

Methods

Participants

A total of 1151 individuals consented to participate in the study. This larger group was comprised of two subgroups: (1) individuals recruited through the researcher's own network, university and the community via social media advertisements, and (2) individuals recruited through a data gathering website, Crowdflower. Participants recruited through the researcher's university – specifically the ANU Psychology Research Participation Scheme – received course credit and those recruited through Crowdflower were financially remunerated for correct participation. All other participants received no incentive for completion. Recruitment of the ASD individuals in the community was assisted by Autism Spectrum Australia (Aspect), which advertised the study on their website and social media channels. Participation in the study was on an entirely voluntary basis. Those who did not complete an entire scale in the questionnaire or who responded "other" to a question concerning gender were omitted from the sample, resulting in a final sample size of 1052, including a Community/Student sample size of 523 (203 males, 320 females; *M* age = 23.76, *SD* = 9.19), and a Crowdflower sample size of 529 (275 males, 254 females; *M* age = 35.56, *SD* = 12.10). Mean ages for self-nominated ASD participants and non-ASD participants across both the Crowdflower and Community/Student samples are displayed in Table 1. Those who indicated that they had been told by a professional that they had an intellectual disability and did not complete school beyond primary school were to be removed, as it could not be guaranteed that they could comprehend the instrument's contents. However, no participants met this criterion.

Table 1

Mean age for ASD and non-ASD participants

	Mean age in years	SD
Non-ASD total (*n* = 1007)	29.66	12.33
Non-ASD female (*n* = 551)	28.99	12.88
Non-ASD male (*n* = 465)	30.47	11.62
ASD total (*n* = 45)	29.91	10.35
ASD female (*n* = 23)	31.57	10.65
ASD male (*n* = 22)	28.10	9.96

Measures

Participants completed a set of demographic questions, which included questions relating to gender, highest level of education completed, delivery of education (mainstream or special needs) and the presence of an intellectual disability. Participants were also asked to indicate whether they had ever been diagnosed as having Autism Spectrum Disorder, Asperger's Disorder or Pervasive Developmental Disorder. Those who provided a positive response to this question were then asked questions relating to the professional who provided the diagnosis (e.g., psychologist, psychiatrist, paediatrician or other), their age at diagnosis, how many practitioners they saw prior to receiving the diagnosis and any comorbid diagnoses. After completing the demographic questions, all participants completed the following self-report questionnaires. Measures can be found in Appendix A.

Autism-Spectrum Quotient-Revised (AQ-Revised)

The AQ (Baron-Cohen et al., 2001) is an instrument designed to measure autistic traits in intellectually competent adults (α = .79). The AQ comprises 50 items and, utilising a forced choice format, is rated on a 4-point Likert scale ranging from *definitely agree* to *definitely disagree.* With a maximum total score possible of 50, a dichotomous scoring system (0/1) is used, where a score of 1 is granted if the respondent records the 'autistic trait'

either mildly or strongly and 0 is given for other responses. Twenty-four items are worded to produce an 'agree' response, and 26 a 'disagree' response, in a high scoring individual with HFA.

For the purpose of the present study, 15 items believed to reflect the female ASD presentation were added into the AQ. Permission to adapt the scale was granted from the original authors of the AQ via email (see Appendix B for a copy of the correspondence). The 15 questions were selected from an original 20 new items, created by the author of the study, based on the research reviewed in the paper's introduction regarding possible sex differences in the ASD phenotype. Thirty experienced ASD psychologists and clinical psychologists, who were recruited through the author's own network and through the Australian Psychological Society (APS) Autism Spectrum Disorder Practitioner List rated the 20 newly developed items in regard to their relevance to the female ASD presentation on a 5-point Likert scale (1=*extremely irrelevant*; 5=*extremely relevant*). The results were then analysed and five items were selected for elimination based on their comparatively low mean, mode and median scores and high variance scores across both the 29 respondents and a comparator who had reportedly assessed and treated in excess of 10 000 individuals on the spectrum, as well as the item’s request for a judgement about qualities rather than behaviour (see Appendix C for the questionnaire items, further information regarding the items selection process, details of the professionals and relevant statistics). Following the elimination process, 15 items from the original 20 remained. The new 65-item combined instrument (merging the original AQ and additional 15 questions) will be referred to as the Autism Spectrum Quotient-Revised (AQ-Revised). Based on qualitative differences outlined in the literature, the new questions encompass both behavioural ASD traits relating to the two ASD domains of impairment, social communication and restrictive, repetitive behaviours, as well as adaptive functions in daily living skills. These domains of interest were first identified

from examination of the literature and then the items were generated to reflect the constructs. The AQ-Revised follows the same format as the AQ and, to avoid a response bias, approximately half of the new items are worded to produce an 'agree' and half a 'disagree' response. The items which would be reverse scored were determined using an online random number generator. Items were also randomised with respect to the expected response from a high-scoring completer and the new questions were randomised within the original AQ. This was done by selecting a randomising setting on the online survey software through which the survey was created. Scoring is consistent with the original AQ scoring.

Empathy Quotient-Short (EQ-Short) and Systemizing Quotient-Short (SQ-Short)

The EQ-Short (Wakabayashi et al., 2006) is a 22-item measure developed to assess empathy ($\alpha = .88$), a component of social cognition considered to be impaired in ASD individuals. The SQ-Short (Wakabayashi et al., 2006) is a 25-item measure developed to assess systemising ($\alpha = .88$), which is the drive to analyse, understand, predict, control and construct rule-based systems, a drive which is considered to be strong in ASD individuals (Wheelwright et al., 2006). Both questionnaires have the same forced-choice, 4-point Likert scale format with some items worded to produce an 'agree' response, and others a 'disagree' response. Each of the empathy/systemising questions scores 2 points if the respondent selects the empathetic/systemising behaviour strongly and 1 point if mildly, making a maximum possible score of 44 and 50 on the EQ-Short and SQ-Short respectively, and a minimum of 0. The EQ-Short and SQ-Short have displayed greater reliability compared with their original 40-item scales (Wakabayashi et al., 2006). Using the longer scales, Wheelwright et al. (2006) report a strong negative correlation between the AQ and the Empathising Quotient and a moderate positive correlation between the AQ and the SQ-Revised. As such, these scales were used to inform on the convergent and divergent validity of AQ-Revised.

Impact Questionnaire

Created by the author of the study, the Impact Questionnaire was designed to assess the presence and degree of dysfunction in seven different domains of one's life. These included working/studying, sleeping, attention and concentration, interacting with others, feeling anxious and stressed, relationships and planning work or daily activities. Using a 1-month reference period, respondents are asked to rate how much trouble they have experienced in the different domains. The questionnaire is scored on a 4-point Likert scale ranging from 0 (*not at all*) to 3 (*very much*), with a maximum possible score of 21 and minimum of 0.

Kessler Psychological Distress Scale (K-6)

The K-6 (Kessler et al., 2002) is a widely used instrument designed to screen for psychological distress ($\alpha = .83$). Using a 30-day reference period, respondents are asked to rate how often they felt “nervous”, “hopeless”, “restless or fidgety”, “so depressed that nothing could cheer them up”, and “that everything was an effort” and “worthless”. The 6-item measure is scored on a 5-point Likert scale ranging from 0 (*none of the time*) to 4 (*all of the time*), with a maximum score possible of 24 and minimum of 0.

Procedure

Before commencing recruitment, ethics approval was obtained from the Human Australian National University Research Ethics Committee. Ethics approval was first granted for phase one of the study (Protocol number: 2016/435), the item selection process, whereby experienced ASD clinicians completed an online questionnaire. Clinicians were sent a link to the questionnaire via email, which was created using the online survey software, Qualtrics. Following analysis of the phase one responses and selection of the final 15 AQ-Revised items, ethics approval was granted for the primary phase of the study (Protocol number: 2016/771) and a second online questionnaire, again using Qualtrics, was created. This

questionnaire comprised of a participant information page, which provided information about the study and included contacts of support services for use if the questionnaire triggered any distress, consent to participate, demographics questions, the AQ-Revised, the EQ-Short, the SQ-Short, the Impact Questionnaire, the K-6 and a thank-you message, which again presented the contact details of support services. The questionnaire was completed by all participants online and was estimated by Qualtrics to have a 16-minute completion time.

Results

Overview of Exploratory Factor Analysis (EFA)

The Statistical Package for Social Sciences (SPSS; version 24.0) was used to analyse the data. As an ordinal scale, in comparison to a dichotomous scale, extracts responses from an individual that better approximates a continuous distribution and thereby provides more information for factor analysis (Kloosterman et al., 2010; Gorsuch, 1983), the full range of possible scores were used in the EFA and Confirmatory Factory Analysis (CFA). The purpose of the EFA was to determine whether the new items gelled with the existing AQ items and to identify any items failing to load onto sensible factors. The 65 items of the AQ-Revised were subjected to principal axis factoring extraction (PAF) with an oblimin rotation. This was considered the most appropriate approach, as it was reasonable to expect that the factors would be correlated (Field, 2013; Fabrigar, Wegener, MacCallun, & Strahan, 1999). The sample obtained from Crowdflower (n= 529, Females = 254, Males = 275) was used for this analysis. This group was considered sufficiently large for EFA, as the sample was over 300 participants (Tabachnick & Fidell, 2007) with a minimum of five participants per variable measured (Fabrigar et al., 1999).

Exploratory Factor Analysis of the AQ-Revised

Prior to performing PAF, the appropriate items in the scales were reverse scored and the suitability of the data for factor analysis was assessed. Inspection of the correlation matrix

revealed the presence of coefficients of .3 and above. The Kaiser-Meyer-Olkin value was .89, exceeding the recommended value of .6 (Kaiser, 1970, 1974; Tabachnick & Fidell, 2013) and Bartlett's Test of Sphericity (Bartlett, 1954) reached statistical significance, supporting the factorability of the correlation matrix. As PAF would also be conducted on each gender group separately, these measures were also assessed on the data after a gender split was performed. The Kaiser-Meyer-Olkin value for the females and males was .82 and .84 respectively and the Bartlett's Test of Sphericity reached statistical significance for both.

An initial analysis was conducted to obtain eigenvalues for each factor in the data. Fourteen factors had eigenvalues over Kaiser's criterion of one and in combination explained 44.55% of the variance. However, there were instances of cross loading and five factors consisted of only three or fewer items, making it a "weak and unstable" factor (Costello & Osborne, 2005). An inspection of the scree plot revealed a possible break at the fifth factor and inspection of the separate female and male scree plots indicated inflection at the fourth and fifth factors respectively (see Appendix D for the scree plots). Using Catell's (1966) scree test and considering previous research proposing 5-, 4-, 3- and 2-factor models for the AQ (Kloosterman et al., 2011), the decision to retain five factors for further investigation was made. PAF was conducted again with an oblimin rotation, this time extracting 2-, 3-, 4-, and 5-factor solutions, which explained 25.16%, 29.38%, 32.15% and 34.08% of the variance respectively.

Examination of the solutions and eventual elimination of items was undertaken in an iterative fashion based on a combination of empirical and conceptual considerations. Throughout the process, the author remained cognisant of the purpose of the analysis for the present study (namely to assess whether the new items fit sensibly with the existing AQ items) and aim of the study (namely to improve ASD screening for females). As such, both gender combined and gender separate PAF's were examined. The 5-factor combined gender

solution explained 34.90% of the variance, had only 3 cross loadings and the factors were broadly interpretable. However, nine items had factor loadings of less than .32, three of which were new items. Tabachnick and Fidell (2013) cite .32 as a decent rule of thumb for the minimum loading of an item, equating to approximately 10% overlapping variance with other items in the factor. Further, only four items loaded onto the final factor, all of which were negatively loaded, making it a weak and unstable factor. Taken together, this suggests that the solution was not optimal. Although the original AQ was designed to have five factors, the items proposed to fit into each of the factors did not load in this manner with 18 items loading onto the first factors and four on the last. Further, examination of the gender separate factor structures indicated the items were loading in different ways dependent upon gender, with regard to the factors they loaded onto and the direction in which they were loading, making them challenging to interpret despite the new items loading neatly into the different factors.

In the 4-factor model there were substantially more cross loadings, with seven in the combined gender analysis, five in the females and nine in the males. In the combined gender rotation, ten items failed to load onto any factor at .32, two of which were new items. Further, the second factor in particular was difficult to interpret and the last factor had only five items loading onto it, which ranged from .44 to .65 in strength of loading. Costello and Osbourne (2005) state that 5 or more strongly loading items (.50 or above) are desirable and indicate a solid factor, yet three of the items on the final factor were less than .5.

In the 3-factor model, while there were no cross loadings in the combined gender data and only two of the new items out of ten failed to load onto any factor below .32, the items loading onto each factor appeared mixed in terms of theme. Further, following the gender split, 14 of the items failed to load at .32 in the female data, five of which were new items, there were three cross loadings and the primary loadings in the final factor were in the

opposite direction from other items defining the same factor. In the male data, there were 6 cross loadings and eight items failing to load at .32, none of which, however, were new items.

The 2-factor model, while explaining less variance, having three cross-loadings, having low communalities comparatively to the other structures, which can indicate that the item is unrelated to the other items (Costello & Osbourne, 2005), and ten items, three being new items (items 57, 53 and 58), not loading onto any factors, the structure made sense with regards to ASD having two primary domains of impairment, as outlined in the DSM-5. Further, on all three scree plots, the largest break was at 2 factors. The female data revealed no cross loadings and three new items out of 11, items 58, 55 and 53, failing to load onto any factors at .32. The male data revealed 3 cross loadings and 1 new item out of ten not loading onto any factors at .32. One new item, item 59, appeared to negatively load on the 2-factor structure in the mixed gender and separate gender rotations. Items 53 and 55 also loaded negatively on the combined and male gender rotations. An additional new item, item 62, was found to be negative on the combined rotation, as well as the male rotation.

Taken together, the 2-factor structure was considered to best depict how the items clustered together. On the basis of items 57, 58, 55, 53 not loading on any factor in either the gender combined or separate rotations and displaying very low communalities (all below .2), these items were removed. The two new negatively loading items were also removed, as they were not acting in the manner in which they were designed and would have implications for scoring. Specifically, with item 59 loading negatively and at a low strength (<.5) in both genders it did not display any great utility in identifying females over males. Regarding item 62, this item was the weakest loading of the new items across all rotations (e.g., -.388 in the combined gender rotation). This item did not load negatively in the female gender rotation, indicating differences in male and female responding. Further, there was an existing AQ item, item 43, that loaded negatively in the male rotation and another, item 41, that had a

negative cross-loading. However, these items were not removed, as the aim of the study was to improve the existing measure by adding new items rather than removing existing items and there is not necessarily a negative implication for females in items loading negatively in males. It does, however, support the concept of differences in the response profile of items across gender.

The six items discussed in the following paragraph were removed and the analysis was repeated on the remaining 59 items. In the combined rotation, the 2-factor solution explained 26.21% of the variance with factor 1 contributing 15.11% and factor 2 contributing 11.10%. All of the new items fitted within the two factors above .32 across the combined and separate gender rotations and all items were positive, with the exception of one item in the male rotation which had a negative cross loading. The items loading onto each factor varied slightly across the combined and gender separate rotations. Table 2 displays the factor loadings after rotation with males and females combined (with values below .32 left blank for clarity).

The items clustering on factor 1 appeared to primarily focus on socialisation, and so was called "Socialisation". The items clustering on factor 2 appeared to focus on interests and the subtleties of communication, and so was called "Interests and Communication Subtleties".

Table 2

Factor loadings and parameter estimates from the PAF (n=529) and CFA (n=523) of the AQ-Revised

Factor	Items	PAF	CFA
Socialisation	Q47. I enjoy meeting new people. *	.77	1.00
	Q38. I am good at social chit-chat. *	.75	1.13
	Q44. I enjoy social occasions. *	.74	1.10
	Q11. I find social situations easy. *	.74	1.25
	Q17. I enjoy social chit-chat. *	.72	1.12
	Q65. Socialising is easy; I never feel lonely and desire friendship. *	**.70**	**.99**
	Q63. Friendships were easy for me during adolescence. *	**.61**	**.93**
	Q10. In a social group, I can easily keep track of several different people's conversations. *	.60	.99
	Q15. I find myself drawn more strongly to people than to things. *	.58	.75

	Q50. I find it very easy to play games with children that involve pretending. *	.52	.50
	Q34. I enjoy doing things spontaneously. *	.51	.66
	Q26. I frequently find that I don't know how to keep a conversation going.	.49	1.07
	Q22. I find it hard to make new friends.	.47	1.12
	Q3. If I try to imagine something, I find it very easy to create a picture in my mind.	.45	.28
	Q8. When I'm reading a story, I can easily imagine what the characters might look like. *	.45	.51
	Q27. I find it easy to "read between the lines" when someone is talking to me. *	.45	.87
	Q32. I find it easy to do more than one thing at once. *	.45	.78
	Q61. I typically cope with demands by tackling them head on. *	**.44**	**.54**
	Q37. If there is an interruption, I can switch back to what I was doing very quickly. *	.43	.67
	Q1. I prefer to do things with others rather than on my own. *	.43	.66
	Q48. I am a good diplomat. *	.43	.73
	Q36. I find it easy to work out what someone is thinking or feeling just by looking at their face. *	.43	.68
	Q14. I find making up stories easy. *	.42	.50
	Q40. When I was young, I used to enjoy playing games involving pretending with other children. *	.36	.54
	Q46. New situations make me anxious.	.35	.66
	Q25. It does not upset me if my daily routine is disturbed. *	.34	.63
	Q31. I know how to tell if someone listening to me is getting bored. *	.33	.76
	Q.24. I would rather go to the theatre than a museum. *	-	-
	Q.28. I usually concentrate more on the whole picture, rather than the small details. *	-	-
	Q.43. I like to plan any activities I participate in carefully.	-	-
Interests and Communication Subtleties	Q39. People often tell me that I keep going on and on about the same thing.	.63	1.00
	Q33. When I talk on the phone, I'm not sure when it's my turn to speak.	.62	1.43
	Q7. Other people frequently tell me that what I've said is impolite, even though I think it is polite.	.61	1.05
	Q16. I tend to have very strong interests, which I get upset about if I can't pursue.	.61	.90
	Q20. When I'm reading a story, I find it difficult to work out the characters' intentions.	.61	.93
	Q52. I have been told that I have an unusual tone of voice (childish, high pitched or hoarse).	**.59**	**.76**
	Q51. When I was young, I spent more time in my own fantasy world that in reality.	**.56**	**0.99**
	Q45. I find it difficult to work out people's intentions.	.54	1.26
	Q35. I am often the last to understand the point of a joke.	.53	1.00
	Q64. During my schooling years, I recall desiring the exclusive attention of one particular friend.	**.53**	**.86**
	Q4. I frequently get so strongly absorbed in one thing that I lose sight of other things.	.53	1.02
	Q18. When I talk, it isn't always easy for others to get a word in edgeways.	.48	.61
	Q56. Sometimes I feel confused about my identity.	**.47**	**.87**
	Q41. I like to collect information about categories of things (e.g. types of car, types of bird, types of train, types of plant, etc.).	.46	1.00
	Q54. In social situations, I wear a mask or take on a persona other than my own.	**.46**	**1.31**

Q9. I am fascinated by dates.	.45	.66
Q23. I notice patterns in things all the time.	.44	.55
Q5. I often notice small sounds when others do not.	.44	.75
Q42. I find it difficult to imagine what it would be like to be someone else.	.42	1.02
Q2. I prefer to do things the same way over and over again.	.39	.97
Q6. I usually notice car number plates or similar strings of information.	.37	.72
Q19. I am fascinated by numbers.	.36	.76
Q60. During school, I was described as passive, shy and/or compliant.	**.36**	**1.01**
Q21. I don't particularly enjoy reading fiction.	.35	.37
Q13. I would rather go to a library than a party.	-	-
Q30. I tend to notice details that others do not.	-	-
Q12. I don't usually notice small changes in a situation, or a person's appearance. *	-	-
Q49. I am not very good at remembering people's date of birth. *	-	-
Q29. I am not very good at remembering phone numbers. *	-	-

Note. * Denotes a reverse scored item; All displayed CFA parameter estimates are significant ($p < .01$); Newly developed screening items are in boldface.

Exploratory Factor Analysis of the Original AQ

A PAF of the original AQ was also run with genders combined and separate. Inspection of the correlation matrix revealed the presence of some coefficients of .3 and above, the Kaiser-Meyer-Olkin value was .89 and Bartlett's Test of Sphericity reached statistical significance, supporting the factorability of the correlation matrix. The Kaiser-Meyer-Olkin value for the females and males separately was .82 and .85 respectively and the Bartlett's Test of Sphericity reached statistical significance for both. An initial analysis was conducted to obtain eigenvalues for each factor in the data. Eleven factors in the combined gender analysis had eigenvalues over Kaiser's criterion of one and in combination explained 43.64% of the variance. Inspection of the scree plot revealed a break at four factors. Inspection of the separate female and male scree plots indicated inflection at the fourth factor, however, a large break at two factors. This break was most pronounced in the male scree plot (see Appendix D for the scree plots). None of the PAFs supported a five-factor solution, as proposed by the authors of the AQ. Further, examination of the pattern matrices for 2-, 3- and 4 factor solutions did not reveal that the items from the specified AQ subscales fit together in the defined factors. In addition, there were some items that failed to load on any factor at .32

or higher, suggesting that the removal of some items may improve the performance of the questionnaire; however, this was beyond the scope of the present study. Without the removal of items, the 4-factor model displayed high cross-loadings in the combined gender analysis and the 3-factor model lacked conceptual congruence with the other items on the factors. Thus, despite displaying some low communalities and explaining only 25.58% of the variance, the AQ, according to PAF with the current sample, appears to fit a 2-factor model, based on low cross loadings and the scree plot.

Overview of the Confirmatory Factor Analysis

The CFA with weighted least squares means and variance adjusted estimation (WLSMV) was conducted in "R" 2.4.2 with the statistical package "lavaan". WLSMV was selected as the estimation method due to the questionnaire having only 4 Likert items (Muthen & Kaplan, 1985). The purpose of the CFA was to cross-validate the 2-factor model of the AQ-Revised, as outlined above. The CFA was performed with data from the student/community subgroup of participants (n= 523, Females =320, Males =203). Consistent with the above, with the sample being over 300 participants (Field, 2013; Tabachnick & Fidell, 2007) and a minimum of five participants per variable measured (Fabrigar et al., 1999), the sample size was deemed more than sufficient for CFA. The power of the sample to reject bad models based on Root Mean Square Error of Approximation (RMSEA = .08) and retain good models (RMSEA=.05) was strong, with the power estimate being 1.00, using the test of close fit (MacCallum, Browne, & Sugawara, 1996).

Confirmatory Factor Analysis

The two factors were specified as follows: Socialisation (items 47, 38, 44, 11, 17, 65, 63, 10, 15, 50, 34, 26, 22, 8, 3, 32, 27, 61, 48, 37, 36, 1, 14, 40, 46, 25, 31, 24, 28, and 43) and Interests (items 39, 33, 7, 16, 20, 52, 45, 51, 64, 35, 4, 18, 56, 41, 54, 9, 23, 5, 42, 2, 6, 60, 19, 21, 13, 30, 12, 49, and 29). Model fit was determined by a number of goodness-of-fit

indices (see Table 3), including the Comparative Fit Index (CFI), the Non-Normed Fit Index (NNFI), the Root Mean Square Error of Approximation (RMSEA), the Standardised Root Mean Square Residual (SRMR) and the ratio of χ^2 (chi-square) goodness-of-fit to the degrees of freedom.

The CFI and NNFI both range from 0 to 1 with higher values representing a superior model fit (Socha, Cooper, & Mccord, 2010) and values equal to or greater than .90 considered acceptable for a well-fitting model (Bentler, 1992; Bentler & Bonnet, 1980). The RMSEA and SRMR also range from 0 to 1; however, lower values represent a superior model. Values equal to or less than .07 (Steiger, 2007) and .08 respectively for the RMSEA and SRMR statistics are deemed acceptable (Hu & Bentler, 1999).

The χ^2 (chi-square) goodness-of-fit statistic assesses the proposed model against the alternative that the variables are correlated merely by chance (Bentler & Bonett, 1980). A failure to reject this test would indicate that the residual covariance estimate is equal to a matrix that contains zeros alone – a signal of perfect model specification (Socha et al., 2010). Thus, rejection of this test (i.e., a significant *p*-value) would suggest a poor model fit. This test, however, is sensitive to sample size and frequently rejects the model when large samples are used (Bentlet & Bonnett, 1980; Jöreskog and Sörbom, 1993). Due to the sensitivity of the χ^2, Wheaton, Muthen, Alwin, and Summers (1977) proposed a relative/normed chi-square (χ^2/df) that, in an effort to minimise the impact of sample size, takes the ratio of the χ^2 to the degrees of freedom. While currently there is no consensus regarding an acceptable ratio for this statistic, Wheaton and colleagues (1977) recommend 5 or less as sufficient. Row 1 of Table 3 presents the goodness-of-fit for the model and Table 2 presents the parameter estimates. As presented in Table 3, the model fell short of the full recommended criteria for good fit. To explore how the model could better fit the data, parameter estimates were examined and items with low magnitude of loadings were identified. With the removal of 7

items that displayed low magnitude of loadings (30, 49, 29, 12, 3, 21 and 24), all from the existing AQ, the model was found to adequately fit the data (see row 2 of Table 3). The χ^2 was significant for both models (with p<.0005), however, as noted, this test is sensitive to sample size. The correlation between the factors was moderately strong, r=.67 and Cronbach's alphas for the Socialisation and Interests and Communications Subtleties subscales were .90 and .83 respectively.

Table 3

Fit Indices

Model	χ^2	*df*	χ^2/df	CFI	NNFI	RMSEA	SRMR
2 Factor Model	4741.49*	1648.00	2.88	.89	.89	.06	.07
2 Factor Model with 7 items removed	3956.81*	1323.00	2.99	.91	.90	.06	.08

Note. *p<.0005

Characteristics of the AQ-Revised

The remainder of the statistical tests were conducted on the entire sample, with the AQ and AQ-Revised scored dichotomously (0/1). The two samples were examined for significant differences on the AQ-Revised prior to combining the Crowdflower and Community/Student samples. On average, the Crowdflower sample scored higher on total AQ-Revised scores (M = 25.02, SD = 8.73) than those participants in the Community/Student sample (M = 22.53, SD = 9.85). The difference was significant, t(1032.05) = -4.337, p < .00 (two tailed), however, it was a weak effect, d = 0.27 according to Cohen (1969). Thus, it was considered acceptable to combine the samples.

A histogram of the 59-item AQ-Revised and the two factors suggested an absence of a normal distribution (see Appendix E for histograms). The Kolmogorov-Smirnov and Shapiro-Wilk statistics were significant for all, suggesting a violation of the assumption of normality. However, these tests are considered unreliable in sample sizes above 300 and

using histograms and absolute skewness and kurtosis values is considered best practice for larger samples (Kim, 2013). Kim (2013) proposes an absolute value in excess of 2 or an absolute kurtosis above 7 as reference values for determining non-normality. On this basis, the data is considered to have a non-normal positive skew, which is expected given the lower frequency of individuals high on ASD traits in the normal population compared to those lower in traits, and a slightly platykurtic distribution within the normal range (see Table 4). With large samples, as in the present study, Tabanachnick and Fidell (2007) state that skewness will not 'make a substantive difference in the analysis'. Tabachnick and Fidell (2007) also suggest that if factor analysis is used in a descriptive manner to summarise relationships in a large set of variables, as in the present study, the assumptions surrounding variable distributions, such as normality and linearity, do not apply, and the *t* test is said to be robust to moderate departure from normality (Kim, 2013). Due to the number of statistical analyses carried out, a significance level of alpha = 0.01 was selected.

Table 4

Descriptive Statistics for the Variables

Scale	Mean (*N* = 1052)	Standard Deviation	Z Skewness	Z Kurtosis
AQ-Revised (Total)	23.78	9.38	.31	-.124
Socialisation	11.26	6.30	.41	-.65
Interests	12.06	5.35	.50	-.12

Mean Scores

Independent samples t-tests were conducted to compare the AQ-Revised total score for males and females in both non-ASD and ASD individuals. There was no significant difference in scores for non-ASD males ($M = 23.85$, $SD = 8.60$) and females ($M = 22.65$, $SD = 9.17$; $t\ (1005) = -2.11$, $p = .04$, two tailed), at the $p = .01$ level. However, a significant difference was found in scores for ASD males ($M = 33.45$, $SD = 6.88$) and females ($M =$

40.30, *SD* = 9.72; *t* (39.68) = 2.74, *p* = .009), two tailed) with females scoring higher. An independent samples t-test also revealed a significant difference in total AQ-Revised scores for ASD (*M* = 36.96, *SD* = 9.04) and non-ASD (*M* = 23.19, *SD* = 8.96; *t* (1050) = 10.08, *p* < .000, two tailed), with the ASD group scoring higher, thereby indicating predictive validity.

Independent samples t-tests were also conducted to compare the AQ total score for males and females in both non-ASD and ASD individuals. There was a significant difference in scores for non-ASD males (*M* = 20.39, *SD* = 7.35) and females (*M* = 19.01, *SD* = 7.84; *t* (1005) = -2.87, *p* = .004, two tailed), with males scoring higher. A significant difference was also found in scores for ASD males (*M* = 28.18, *SD* = 6.18) and females (*M* = 34.13, *SD* = 8.36; *t* (40.46) = 2.72, *p* = .01, two tailed) on the AQ total score, with females scoring higher. An independent samples t-test also revealed a significant difference in total AQ scores for ASD (*M* = 31.22, *SD* = 7.89) and non-ASD (*M* = 19.64, *SD* = 7.65; *t* (1050) = 9.93, *p* < .000, two tailed), with the ASD group scoring higher.

Further, independent samples t-tests were conducted on each of the nine new items for males and females in both non-ASD and ASD individuals. While most of the item mean scores (7/10 in the ASD group and 5/10 in the non-ASD group and one equal) were higher in females than males, non-significant differences between the genders were found on the majority of items (see Appendix F for means and standard deviations for all nine items across groups). In the non-ASD group, only one item, item 65 "*Socialising is easy; I never feel lonely and desire friendship*" (reverse scored) was significant between the genders (*t* (969.94) = 2.54, *p* = .01, two tailed) with higher scores in females (*M* = .54, *SD* = .50) compared to males (*M* = .46, *SD* = .50). A second item in this group, item 56 "*Sometimes I feel confused about my identity*" displayed a trend towards significance (*t* (979.78) = 1.83, *p* = .07, two tailed) with higher scores in females (*M* = .40, *SD* = .49) compared to males (*M* = .35, *SD* = .48). In the ASD group, two items significantly differed between males and females. Item 65

"*Socialising is easy; I never feel lonely and desire friendship*" (reverse scored; t (32.79) = 3.32, p = .002, two tailed) displayed higher scores in females (M = .91, SD = .29) compared to males (M = .50, SD = .51). Similarly, item 61 "*I typically cope with demands by tackling them head on*" (reverse scored; t (36.67) = 3.19, p = .003, two tailed) displayed significantly higher scores in females (M = .87, SD = .34) compared to males (M = .45, SD = .51). These results indicate that these items were more frequently endorsed in females than males.

Scale Reliability

Cronbach's alpha coefficients for the AQ and AQ-Revised in the present study are presented in Table 5 and indicate that the contents of the instruments have a high internal consistency.

Table 5

Internal Consistencies (Cronbach's alpha) for the AQ and AQ-Revised

Scale	Non-ASD (n = 1007)	ASD (n = 45)	Entire Sample (N = 1052)
AQ	.83	.84	.84
AQ-Revised	.85	.85	.87

Validity

The relationship between the EQ-Short, SQ-Short and the AQ-Revised were examined by testing the correlations between each pair of questionnaires. This was also done for the AQ. To account for the non-normal distribution and the measures being Likert scales in design, the non-parametric Spearman's rho statistic was used to assess validity. With the exception of the AQ-Revised and SQ-Short (p = .015), all of the correlations were significant at the p = <.01 level in the expected directions and strengths: r_s = -.62 for the AQ-Revised and EQ-Short, r_s = .08 for the AQ-Revised and SQ-Short, r_s = -.62 for the AQ and EQ-Short and r_s = .12 for the AQ and SQ-Short.

Cut-Off Scores

In the present study, distribution-free, non-parametric ROC curves, were used to determine appropriate cut-off points for males and females on the new AQ-Revised measure. Self-nomination of an ASD diagnosis based on a prior clinician's diagnosis was used to establish the presence/absence of ASD. The prevalence of ASD in the sample, using self-report of a diagnosis, was 4.28% (23 females, 22 males of a total 1052). Table 6 displays sensitivity (proportion of self-nominated ASD participants who are accurately identified as such) and specificity (the proportion of non-ASD participants who are accurately identified as such) values for a number of potential cut-off scores on the AQ-Revised. Figures 1 and 2 provide the ROC curves for each gender. The area under the curve (AUC) for the females was .90, indicating that the measure is an "excellent" test with favourable sensitivity and specificity statistics. The AUC for males was .81, indicating that it has "good" discriminatory ability. In designing a screening test, a sensitivity value of .8 or more is required, otherwise those who meet criteria for the disorder will be missed (Habibzadeh, Habibzadeh, & Yadollahie, 2016). Considering this, and the rule that a useful threshold score would discriminate ASD and non ASD-cases with as many true positive and as few false positives as possible, also used by Baron-Cohen et al. (2001), a cut off score of 28.5 on the AQ-Revised appears to be the best in the current sample for both males and females. Such a score provides, in females, a 17.4% false negative rate and a 28.3% false positive rate, and in males, an 18.2% false negative rate and a 31.8% false positive rate. Given that failure to pick up an ASD diagnosis has serious consequences, sensitivity is of greater importance than specificity in this screening instrument. As a score of 28.5 is impossible (due to the dichotomous scoring of 0 or 1), this cut off score is rounded up to 29. For females, at a cut off of 29, the AQ-Revised has a Positive Predictive Value (PPV) of 10.9% and a Negative Predictive Value (NPV) of 99.0%. Therefore, there is a 10.9% chance that a participant who

met the threshold of 29 actually reported having ASD, but only a 1% chance that a female responder who scores below the threshold of 29 actually has ASD. This is compared to the AQ with a threshold score of 32, which was found to have a PPV of 31.1% and a NPV of 98.3%, for females.

Figure 1

ROC curve for the AQ-Revised in Females

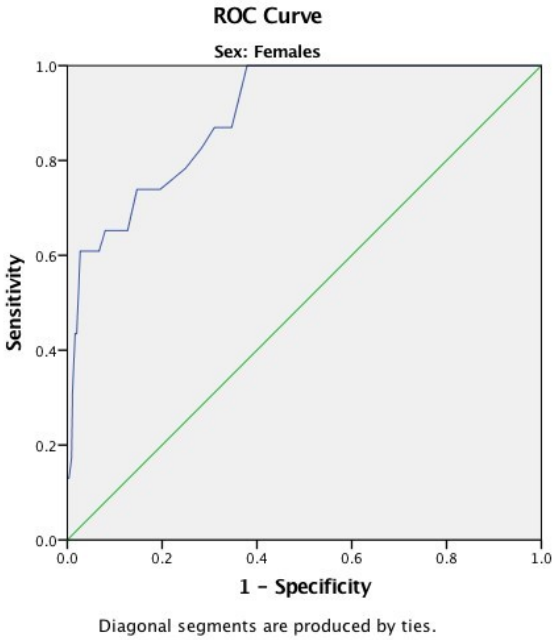


Figure 2

ROC curve for the AQ-Revised in Males

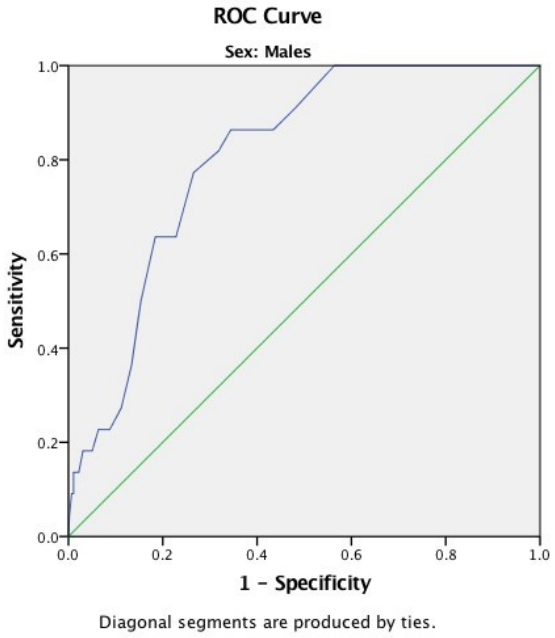


Table 6

Detailed report of diagnostic statistics for the AQ-Revised

Cut-off Point	Females		Males	
	Sensitivity (%)	Specificity (%)	Sensitivity (%)	Specificity (%)
≥.00	-	-	100	0.0
≥2.00	100	0	-	-
≥3.00	-	-	100	0.2
≥3.50	100	0.2	-	-
≥4.50	100	0.4	-	-
≥5.50	100	1.3	100	0.9
≥6.50	100	2.2	100	1.1
≥7.50	100	3.3	100	1.8
≥8.50	100	4.7	100	3.1
≥9.50	100	6.7	100	4.6
≥10.50	100	8.7	100	5.7
≥11.50	100	10.2	100	7.7
≥12.50	100	13.4	100	9.9
≥13.50	100	18.5	100	13.4
≥14.50	100	20.9	100	15.4
≥15.50	100	24.5	100	19.7
≥16.50	100	29.8	100	23.7
≥17.50	100	32.7	100	25.7
≥18.50	100	36.3	100	28.7
≥19.50	100	39.4	100	32.7
≥20.50	100	43.6	100	35.3
≥21.50	100	48.1	100	39.5
≥22.50	100	51.2	100	43.6
≥23.50	100	54.3	95.5	47.8
≥24.50	100	57.9	90.9	52.0
≥25.50	100	62.1	86.4	56.6
≥26.50	87.0	65.3	86.4	60.3
≥27.50	87.0	69.0	86.4	65.6
≥28.50	82.6	71.7	81.8	68.2
≥29.50	78.3	75.1	77.3	73.5
≥30.50	73.9	80.4	63.6	77.2
≥31.50	73.9	83.1	63.6	81.6
≥32.50	73.9	85.3	50.0	84.6
≥33.50	65.2	87.3	36.4	86.6
≥34.50	65.2	90.6	27.3	88.8
≥35.50	65.2	92.0	22.7	91.2
≥36.50	60.9	93.3	22.7	92.5
≥37.50	60.9	94.9	22.7	93.6
≥38.50	60.9	95.6	18.2	95.0
≥39.50	60.9	96.6	18.2	96.9
≥40.50	60.9	97.3	13.6	97.8
≥41.50	52.2	97.6	13.6	98.2
≥42.50	43.5	98.0	-	-
≥43.00	-	-	13.6	98.9
≥44.50	43.5	98.4	9.1	98.9
≥45.50	39.1	98.5	9.1	99.1
≥46.50	34.8	98.7	-	-
≥47.00	-	-	9.1	99.3
≥47.50	30.4	98.9	-	-
≥48.50	17.4	99.1	4.5	99.8
≥49.50	13.0	99.6	-	-
≥50.00	-	-	0	100
≥51.50	13.0	100	-	-
≥54.00	8.7	100	-	-
≥56.00	0	100	-	-

Females High on ASD Traits

Frequency statistics revealed that the AQ, at the cut off score of 32 proposed by Baron-Cohen et al. (2001), was found to capture 45 females in the current sample. The AQ-Revised, using the cut off score of 29 identified by the ROC curve, was found to capture 175 females, indicating that the new measure and cut-off score identifies individuals high on ASD traits that the AQ does not and is, therefore, a more sensitive screening instrument. The number of females high on ASD traits according to the new instrument and not identified by the AQ was a total of 130. Independent t-tests were conducted to determine whether these 130 females high on ASD traits and not picked up by the AQ displayed high levels of distress, as measured by the K-6, and impact on functioning, as measured by the Impact Questionnaire, compared with those females who did not display high ASD traits according to the AQ-Revised (i.e. below 29). There was a significant difference in K-6 scores for the females picked up by the AQ-Revised as being high on ASD traits (a score equal to or above 29; $n = 130$; $M = 10.11$, $SD = 5.50$) and those scoring below 29 ($n = 399$; $M = 6.47$, $SD = 4.76$; $t\,(527) = 7.28$, $p < .000$, two tailed) with the former scoring significantly higher. There was also a significant difference in Impact Questionnaire scores for the females picked up by the AQ-Revised as being high on ASD traits ($n = 130$; $M = 11.22$, $SD = 5.14$) and those scoring below 29 ($n = 399$; $M = 7.73$, $SD = 4.60$; $t\,(527) = 7.30$, $p < .000$, two tailed), with those scoring above 29 scoring significantly higher. These statistics suggest that there are females in the population high on ASD traits not being picked up by the existing AQ, yet experiencing levels of distress and impairment in functioning above their lower AQ-Revised scoring counterpart.

The difference between the 130 females not captured by the AQ, but captured by the AQ-Revised, and the 45 females that the AQ did identify on the K-6 and Impact Questionnaire was also examined. There was no significant difference in the 130 females not

captured by the AQ (M = 10.11, SD = 5.50) and the 45 females who were (M = 9.78, SD = 5.38; t (173) = .52, p = .73, two tailed) on the K-6. Similarly, there was no significant difference in the 130 females not captured by the AQ (M = 11.22, SD = 5.14) and the 45 females who were (M = 11.67, SD = 4.48; t (173) = 0.52, p = .60, two tailed) on the Impact Questionnaire. This suggests that the level of distress and impairment experienced by the two groups is not different.

Discussion

Research has outlined a need for more "gender specific" items for females in ASD screening instruments and/or different gender based threshold scores, to better capture the subtle challenges that females face and prevent adverse consequences of a missed or delayed diagnosis or a misdiagnosis (Baldwin & Costley, 2016; Bell, Foster, & Mash, 2005; Rutter, Caspi, & Moffitt, 2003). This study sought to address this need and reports on the development and refinement of the AQ-Revised, a new screening instrument for autistic traits in adults, encompassing nine new items that were selected with the assistance of expert psychologists in the field with a view to better capture Females with ASD, and psychometrically evaluate both the existing AQ and the AQ-Revised.

The results of the study support the body of literature and anecdotal evidence that proposes gender differences in ASD by way of differences in item loadings found in the EFA, significant gender differences across total AQ-Revised and AQ mean scores in the ASD participants and non-ASD participants in the AQ, and some significant gender differences across the newly developed AQ-Revised items in both ASD and non-ASD cases. This illustrates the importance of researching symptom gender differentiation at the individual item level when examining screening methods and diagnostic processes for males and females with ASD.

Evaluation of Psychometric Properties

The AQ-Revised and AQ were found to display high internal consistency and convergent and divergent validity with measures assessing domains that are reported to be deficits and strengths in ASD individuals in similar strengths to previous studies evaluating the AQ (e.g., Wheelwright et al., 2006). The internal consistency was found to be slightly higher in the AQ-Revised, compared to the AQ. While this could indicate that the new items have improved the internal consistency of the instrument, this could be a reflection of the AQ-Revised being a longer measure, as it is known that Cronbach's alpha tends to decrease as the number of items reduces (Steiner & Norman, 2008). With both measures displaying strong divergent and convergent validity, support is rendered for the view that empathising is a deficit and systemising abilities are superior in individuals high on autistic traits (Wakabayashi et al., 2006). Given that systemising abilities are stronger in males than females (Wheelwright et al., 2006) and the new items were designed to better identify females, it is not surprising that the strength of the correlation between the AQ-Revised and SQ-Short was weaker in comparison to the AQ in the present study and that conducted by Wheelwright et al. (2006). Future studies evaluating the psychometric qualities of the AQ-Revised should be mindful of this and explore other instruments to assess the divergent validity of the measure.

Regarding the factor structure, both gender combined and gender separate EFAs of the AQ-Revised revealed that a 2-factor model was the best fit, with the removal of six of the new items, and the new items fit neatly with the existing AQ items in this factor structure. Theoretically, this model is consistent with the DSM-5 criteria for ASD, which outlines two primary domains of impairment in a positive diagnosis of ASD. While the 59 items of the ASD-Revised did not all fit neatly within these two domains, they did broadly fit. An attempt to confirm this 2-factor structure was made using CFA with genders combined, which

revealed that the model fell short of meeting recommended criteria for overall good fit. One way to improve the psychometric quality and factor structure of the AQ-Revised would be to eliminate poorly performing items from the original AQ, as the new AQ-Revised items that did not perform well were removed at the EFA stage. With the removal of seven AQ items with low magnitude of loadings an adequately fitting factor structure was derived (see Table 3). While the purpose of the study was the addition of new items to the existing measure to improve the sensitivity for females rather than the removal of items, further research may consider improving the psychometric qualities of the measure across sexes with the removal of existing AQ items that do not perform well.

The EFA of the AQ revealed that, in contrast to the proposal made by the creators of the measure (Baron-Cohen et al., 2001), the AQ did not fit a 5-factor structure. Rather, consistent with the findings of Hoestra et al. (2008), a 2-factor structure appeared to be a better fit. However, this conclusion is made on the basis of an EFA only and requires further testing through CFA. The EFA also revealed that some AQ items did not load appropriately, suggesting that the AQ could perform well with fewer items.

It should be noted that the gender separate EFAs revealed that some items loaded differently dependent upon gender, with regard to strength of loading and direction, suggesting that the factor structure varies depending upon gender, thereby lending support to the concept of gender differentiated response profiles on ASD screeners (Kopp & Gillberg, 2011). While for the purpose of the study the decision was made to conduct a gender combined CFA for the AQ-Revised, future research should investigate the factor structure of ASD screeners for each gender separately – a task which no study has attempted to date.

Gender Differentiation in Item Responding

Consistent with previous findings (e.g., Baron-Cohen et al., 2001; Baron-Cohen et al., 2006; Auyeung, Baron-Cohen, Wheelwright, & Allison, 2008), males scored significantly

higher on the total AQ score compared to females in the non-ASD group. However, in the ASD group, females scored significantly higher than males, which was surprising. This result may be understood with regards to qualitative reports from females explaining that their unusual behaviour and difference made sense following receiving a diagnosis (Baldwin & Costley, 2016; Bargiela et al., 2016). Thus, their greater insight into their behaviour due to their diagnosis, and possible treatment, may mean that they better understood what the questions were asking of them and explain the higher total mean score for females in the ASD group. In the AQ-Revised, there was *no* significant difference between genders in the total AQ-Revised score in the non-ASD group, but, similarly with the AQ, a significant difference in the ASD group with females scoring higher was found. It is possible that, because the AQ was developed prior to the emergence of the literature base supporting gender differences, it was based on the male ASD phenotype. Yet with the inclusion of items more relevant to females, the AQ-Revised closed the gap of significance between males and females of non-ASD individuals with regards to total mean score.

While there was no significant difference in males and females with regards to the total AQ-Revised score, higher mean scores for females were found across the majority of items, with one item in the non-ASD group and two items in the ASD group reaching significance. Such results indicate that these specific items were more often endorsed in females than males and the response profiles differed between males and females, thereby contributing important information regarding the autistic phenotype in Females with ASD. Specifically, item 65, "*Socialising is easy; I never feel lonely and desire friendship*" (reverse scored), was found to have a significantly higher mean in females in contrast to males in both the non-ASD and ASD groups. This finding lends support to the difficulty in social understanding and functioning (Baldwin & Costley, 2015), yet desire for friendship (Barigela et al., 2016), that female adults high on autistic traits are reported to possess, as well as a

greater importance on social connection by females. The mean for item 61, "*I typically cope with demands by tackling them head on*" (reverse scored), was also found to be significantly higher in females than males in the ASD group. This finding lends support to the literature outlining day-to-day functioning as a struggle for adult females on the spectrum (Baldwin & Costley, 2015). With the underlying tone of this question being avoidance, and avoidance being a hallmark characteristic of anxiety, this result could also indicate an elevated presence of comorbid anxiety in females high on ASD traits, a mental health challenge found to have a high prevalence in adult Females with ASD (Baldwin & Costley, 2015). These results suggest that in the screening and diagnostic process for ASD in females, attention should be paid to these two areas.

Interestingly, one of the items designed to tap into the use of "camouflaging" and coping in social situations, "*In a social situation, I wear a mask or take on a persona other than my own*", which remained in the final questionnaire following the EFA, was not found to significantly differ between genders in either group. This is in contrast to anecdotal suggestion and qualitative report (e.g., Dworzynski, Ronald, Bolton, & Happe, 2012; Kopp & Gilberg, 1992; Lord et al., 2000; Mandy et al., Bargiela et al., 2016; Ehlers & Gillberg, 1993; Gillberg, 1993) that a key feature of the ASD female presentation is an ability to "camouflage" and mask social difficulties in social situations with one way of doing so being to take on a persona or cultivate an image different to their own (Barigela et al., 2016). Different interpretations are plausible for the absence of a significant finding. First, it is possible that this item failed to operationalise the construct of masking. Second, with the mean score for this item in both the ASD and non-ASD groups being only a difference of .01, it is possible that males too employ a masking strategy: however, such a concept, being described as a female coping strategy, has never been explored in males. Thus, further quantitative research into the concept of masking in ASD across genders is required.

Prediction, Cut-Off Score and Performance of the AQ-Revised Over the AQ

The AQ-Revised differentiated well between ASD cases and non-ASD cases. This was demonstrated by significant group differences between ASD and non-ASD cases in total AQ-Revised scores, as measured by an independent t-test, and area under the ROC curve for both males and females in the high ranges, indicating that the measure is clinically useful with predictive validity in discriminating ASD cases from non-ASD individuals. Based on the ROC curve, it is recommended that a cut-off score of 29 would discriminate between ASD and non-ASD cases while ensuring the number of false positives are minimised, and simultaneously avoid cases "slipping through the net". At the score of 29 there is a 28.3% and 31.8% false positive rate for females and males respectively. While these false positive rates are lower than some psychometric evaluations of the existing AQ, such as that by Woodbury-Smith et al. (2005) which allowed 48.15% (cut off score = 26), it is higher than others such as that by Broadbent et al. (2013) which only allowed 1% (cut off score = 29). As less than 1% of the population are reported to have ASD, it could be argued that the false positive rates found in the present study are unacceptably high. However, given that the purpose of the study was to better identify females for full diagnostic assessment, making sensitivity more important than specificity, and the concern that current diagnostic practices developed based on the male ASD phenotype are not adequately capturing females, the current prevalence statistics may not be valid and thus a higher false positive rate may be acceptable.

Further regarding the performance of the revised measure, the AQ-Revised, with a cut off score of 29, was found to have a slightly higher NPV than the AQ, with a cut off score of 32 for females. As it is desirable to have a high NPV in a screening test, the AQ-Revised displays superior performance than the existing AQ in this population (Maxim, Niebo, & Utell, 2014).

Females Evading a Diagnosis

Using the cut off score of 29, the study revealed that the AQ-Revised captured 130 females high on ASD traits that the AQ failed to identify. These females were found to display significantly higher levels of distress and functional impairment, as measured by the K-6 and Impact Questionnaire, compared to the females in the study scoring below 29. This finding is striking, as it indicates that there are females in the population not being identified by the existing AQ, yet experiencing higher levels of distress and challenge across different life domains than those in the population lower on ASD traits. It should be noted, however, that while being significantly higher than those scoring below 29, the mean score for females scoring above 29 (M = 10.11) did not reach the score of 13 on the K-6, which indicates "probable *serious* mental illness" (Kessler et al., 2002). Further, results indicated that the 130 females identified by the AQ-Revised and not by the AQ did not experience difference in levels of distress or impairment compared to those picked up by the AQ. While no prior studies have made comparisons with neurotypical individuals, these results are consistent with previous findings that females high on ASD traits struggled in different domains of life, such as daily activities, working and interacting with others, and experienced emotional difficulties prior to any diagnosis (Baldwin & Costley, 2016; Bargiela et al., 2016). The results also highlight the importance of appropriate and timely screening for females high on ASD traits prior to their mental health challenges reaching clinically severe levels.

Strengths and Limitations

The present study has both strengths and limitations. Strengths of the study include the diversity of the sample (i.e., it encompassed students, members of the community and members of a data gathering website both inside and outside Australia), utilisation of both exploratory and confirmatory factor analysis, its psychometric evaluation of both an existing and revised measure, and the inclusion of instruments with established reliability. Moreover,

this study is the first to use the body of research examining sex differences in autistic traits practically, in an attempt to improve the ASD screening process for adult females. Further, in the context of most of the current research focusing on ASD in children, this study is a contribution to a gap in the literature surrounding ASD symptomology in adults, a gap which studies such as Pellicano, Dinsmore, and Charman (2014) have called for researchers to address.

Despite these strengths, however, several limitations do exist. First, the sample of ASD participants was small. Thus, findings should be interpreted with care. Second, age was not accounted for in the study. While previous studies evaluating the AQ have asserted that no significant age effects of the AQ exist (Baron-Cohen, Hoekstra, Knickmeyer, & Wheelwright, 2006; Auyeung, Baron-Cohen, Wheelwright, & Allison, 2008), Broadbent et al. (2013) reported a significant age effect on the total AQ score. Arguing against the claim that the instrument is not influenced by age, this suggests that sensitivity to age may exist within the AQ and, therefore potentially the AQ-Revised. Further, the different recruitment strategies utilised may limit the generalisability of the results, as different recruitment strategies may have resulted in the selection of individuals with particular characteristics or interests. Thus, future research should confirm ASD gender differences using consistent recruitment modalities with larger samples and assess for effect of age.

A limitation in the generation of the items should also be considered. In the field of questionnaire development, is common practice for items to be created following focus group discussions. However, practicality prevented this from occurring in the current study and the use of a rating questionnaire with experienced practitioners in the field who rated the items generated by the author was considered the best substitution for the gold standard approach. It is possible that this approach then undermined the validity of the items generated.

In addition, with the central tenet of the research being the additional of items that would make the AQ more sensitive to the female presentation, Item Response Theory and factorial invariance procedures could have been used to further develop the gender specific items for assessment and screening, specifically to assess the discriminability of the new items across genders and compared against the existing AQ items. As it stands, this study is a step towards the improvement of screening tools for both genders, but particularly for females.

A further limitation of the study, due to the questionnaire methodology employed, was reliance on respondents' self-selection and self-reporting on the basis that they did or did not have ASD and also did not have ID. There was no direct means for verifying the accuracy of respondents' self-report; however, measures were established to ensure the integrity of the sample to the greatest degree feasible. Specifically, respondents were asked questions surrounding the process of their ASD diagnosis. Regarding ID, while the initial screening mechanism of asking participants whether they had ever been told by a professional that they had an ID could only be successful in proportion to the cooperation of participants and self-awareness, there were also intrinsic features of the questionnaire's methodology that would almost certainly have presented difficulty for an individual possessing an ID. For example, the questionnaire was lengthy, thereby requiring concentration to complete, and the questions, while developed to be concise and clear, required the comprehension of a literate adult. Further, criteria were established for determining whether a respondent could comprehend the questionnaire and the data displayed that no individuals reported ID and also did not attain formal education beyond primary school. That said, in assuming skill in reading comprehension, the inappropriateness of the AQ-Revised for individuals with low IQ is a limitation of the instrument. This limitation also applies to the existing AQ.

A further limitation of the self-report nature of the study was the potential for possession of high ASD traits to impair the participant's ability to judge their own communicative or social behaviour, due to mind-reading issues known as "mindblindness", which could impact on responses (Baron-Cohen, 1995). If this were to occur, it would result in a participant scoring lower on the measure, rating their own behaviour as more appropriate than it may in fact be. Thus, any inaccuracies of this nature, if anything, would result in a more conservative estimate of the individual's true AQ-Revised score. However, in designing the AQ, Baron-Cohen et al. (2001) accounted for this through the inclusion of questions which asked about an individual's preference, rather than purely asking them to make a judgement about their own social and communicative behaviours. The newly created items in the present study further aimed to combat response bias by wording half the items with a positive response set and half negative. Future studies could further address these limitations by encompassing a direct observation component or cross-validation techniques.

Conclusions

While there is room for improvement and validation studies must take place with larger ASD populations prior to use, the AQ-Revised appears to be a psychometrically sound instrument and have utility as a brief screener of autistic traits for adults of normal intelligence – with a cut off point of 29 warranting referral for a diagnostic assessment. While only two new items were found to have significant gender differences across ASD and non-ASD individuals, all new items went through a rigorous evaluation process, being assessed by experts in the field as relevant to the established female ASD phenotype, were found to structurally fit with the existing AQ items in factor analyses and, combined with the AQ, were found to capture many more females than the AQ alone. As such, they are considered a very useful addition to the AQ. As a whole, the AQ-Revised, with a cut off point of 29, is effective in identifying females who may lie on the spectrum and, thus the study is a step in

the right direction towards reducing the risk, and associated negative consequences, of ASD going undiagnosed in females. More broadly, the study contributes to the body of literature on the autism spectrum in adults by lending support to the idea that ASD symptomology should be studied with genders separately.

***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. This article does not contain any studies with animals performed by any of the authors.*

References

Allison, C., Baron-Cohen, S., Wheelwright, S., Charman, T., Richer, J., Pasco, G., & Brayne, C. (2008). The Q-CHAT (Quantitative Checklist for Autism in Toddler): A normally distributed quantitative measure of autistic traits at 18-24-months of age: Preliminary report. *Journal of Autism and Developmental Disorders, 38,* 1414-1425. doi: 10.1007/s10803-007-0509-7

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.). Washington BC: American Psychiatric Association.

Andersson, G. W., Gillberg, C., & Miniscalco, C. (2013). Pre-school children with suspected autism spectrum disorders: Do girls and boys have the same profiles? *Research in Developmental Disabilities, 34,* 413-422. doi: 10.1016/j.ridd.2012.08.025

Attwood, T. (2007). Relationship problems of adults with Asperger's Syndrome. *Good Autism Practice,* 8, 13-19

Attwood, T., Grandin, T., Bolick, T., Faherty, C., Iland, L., Mcllwee Myers, J., Snyder, R., Wagner, S., & Wrobel, M. (2006). *Asperger's and Girls*. Arlington, Texas: Future Horizons, Inc

Auyeung, B., Baron-Cohen, S., Wheelwright, S., & Boomsma, D. I. (2008). The autism spectrum quotient: Children's version (AQ-Child). *Journal of Autism and Developmental Disorders, 38,*1230-1240. doi: 10.1002/aur.1651

Baldwin, S., & Costley, D. (2016). The experiences and needs of female adults with high-functioning autism spectrum disorder. *Autism, 20,* 483-495. doi:10.1177/1362361315590805

Bartak, L., & Rutter, M. (1976). Differences between mentally retarded and normally intelligent autistic children. *Journal of autism and childhood schizophrenia, 6,* 109-120. doi: 10.1007/BF01538054

Bargiela, S., Steward, R., William, M. (2016). The experiences of late-diagnosed women with autism spectrum conditions: An investigation of the female autism phenotype. *Journal of Autism and Developmental Disorders, 46,* 3281-3294. doi:10.1007/s10803-016-2872-8

Baron-Cohen, S. (1995). *Mindblindness: An essay on autism and theory of mind.* Boston: MIT Press/Bradford Books

Baron-Cohen, S., Hoekstra, R. A., Kickmeyer, R., & Wheelwright, S. (2006). The autism spectrum quotient (AQ) – adolescent version. *Journal of Autism and Developmental Disorder, 36,* 343-350. doi: 10.1007/s10803-006-0073-6

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. doi: 10.1037/t00350-000

Bartlett, M. S. (1954). A note on the multiplying factors for various chi square approximations. *Journal of the Royal Statistical Society, 16,* 296-298

Bauminger, N., & Kasari, C. (2000). Loneliness and friendship in high-functioning children with autism. *Child Development, 71,* 447-456. doi:10.1111/1467-8624.00156

Bell, D. J., Foster, S. L., & Mash, E. J. (2005). Understanding behavioural and emotional problems in girls. In D. J. Bell, S. L. Foster, & E. J. Mash (Eds.), *Handbook of behavioural and emotional problems in girls* (pp. 1-24). New York, NY: Kluwer Academic/Plenum Publishers

Bentler, P. M. (1992). On the fit of models to covariances and methodology to the Bulletin. *Psychological Bulletin, 112,* 400-404. doi:10.1037/0033-2909.112.3.400

Bentler, P. M., & Bonnet, D. C. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin, 88,* 588-606. doi: 10.1037/0033-2909.88.3.588

Broadbent, J., Galic, I., Stokes, M. A. (2013). Validation of autism spectrum quotient adult version in an australian sample. *Autism Research and Treatment, 2013,* 1-7. doi: 10.1155/2013/984205

Carter, A. S., Black, D. O., Tawani, S., Connolly, C. E., Kadlee, M. B., & Tager-Flusberg, H. (2007). Sex differences in toddlers with autism spectrum disorders. *Journal of Autism and Developmental Disorder, 37,* 86-97. doi: 10.1007/s10803-006-0331-7

Cattell, R. B. (1966). The scree tests for number of factors. *Multivariate Behavioural Research, 1,* 245-276. doi: 10.1207/s15327906mbr0102_10

Cohen, J. (1969). *Statistical Power Analysis for the Behavioural Sciences.* New York, NY: Academic Press

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

Costello, A. B., & Osbourne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment Research & Evaluation, 10,* 1-9

Dworzynsky, K., Ronald, A., Bolton, P., & Happe, F. (2012). How different are girls and boys above and below the diagnostic threshold for autism spectrum disorders? *Journal of the American Academy of Child and Adolescent Psychiatry, 51*, 788-797. doi: 10.1016/jaac.2012.05.018

Ehlers, S., & Gillberg, C. (1993). The epidemiology of Asperger syndrome. A total population study. *Archives of General Psychiatry, 60,* 524-530

Ehlers, S., Gillberg, C., & Wing, L. (1999). A screening questionnaire for Asperger syndrome and other high-functioning autism spectrum disorders in school age children. *Journal of Autism and Developmental Disorders, 29,* 129-141

Fabringar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods, 4,* 272-299. doi:10.1037/1082-989X.4.3.272

Field, A. (2013). *Discovering statistics using IBM SPSS statistics: (and sex and drugs and rock "n" roll)* (4th ed.). Los Angeles, CA: Sage

Fombonne, E. (2003). Epidemiological surveys of autism and other pervasive developmental disorders: An update. *Journal of Autism and Developmental Disorders, 33,* 365-382. doi:10.1023/A:1025054610557

Fombonne, E. (2005). The changing epidemiology of autism. *Journal of Applied Research in Intellectual Disabilities, 18,* 281-294. doi:10.1111/j.1468-3148.2005.00266.x

Fombonne, E (2007). Epidemiological surveys of pervasive developmental disorders. In F. R. Volkmar (Ed.), *Autism and pervasive developmental disorders* (2nd Ed., pp. 33-68). New York, NY: Cambridge University Press.

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

Giarelli, E., Wiggins, L. D., Rice, C. E. Levy, S. E., Kirby, R. S., Pinti-Martin, J., & Mandell, D. (2010). *Disability and Health Journal, 3,* 107-116. doi:10.1016/j.dhjo.2009.07.001

Gillberg, I. C., Råstam, M., Wentz, E., & Gillberg, C. (2007). Cognitive and executive functions in anorexia nervosa ten years after onset of eating disorder. *Journal of Clinical & Experimental Neuropsychology: Official Journal of the International Neuropsychological Society, 29,* 170-178. doi: 10.1080/13803390600584632.

Goldman, S. (2013). Opinion: Sex, gender and the diagnosis of autism – A biosocial view of the male preponderance. *Research in Autism Spectrum Disorders, 7,* 675-679. doi: 10.1016/j.rasd.2013.02.006

Gorsuch, R. L. (1983). *Factor Analysis.* Hillsdale, NJ: Erlbaum

Habibzadeh, F., Habibzadeh, P., & Yadollahie, M. (2016). On determining the most appropriate test cut-off value: The case of tests with continuous results. *Biochemical Medicine, 26,* 297-307. doi: 10.11613/BM.2016.034

Hackler, J. (1986). Treatment of compulsive eating disorders in an autistic girl by combining behaviour therapy and pharmacotherapy. Case report. *Zeitschrift fur Kinder- und Jugendpsychiatrie, 14,* 220-227

Haney, J. L. (2016). Autism, females, and the DSM-5: Gender bias in autism diagnosis. *Social Work in Mental Health, 14,* 396-407. doi: 10.1080/15332985.2015.10318558

Hartley, S. L., & Sikora, D. M. (2009). Sex differences in autism spectrum disorder: An examination of developmental functioning, autistic symptoms, and coexisting behaviour problems in toddler. *Journal of Autism and Developmental Disorders, 39,* 1715-1722. doi: 10.1007/s10803-009-0810-8

Hoekstra, R. A., Bartels, M., Verweij, C. J. H., & Boomsma, D. I. (2007). Heritability of autistic traits in the general population. *Archives of Pediatrics & Adolescent Medicine, 161,* 372-377. doi:10.1001/archpedi.161.4.372

Hoekastra, 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,* 1555-1566. doi: 10.1007/s10803-008-0538-x

Holterman, M., Bolte, S., & Poustka, F. (2007). Autism spectrum disorders: Sex differences in autistic behaviour domains and coexisting psychopathology. *Developmental*

Medicine and Children Neuropsychology, 49, 361-366. doi: 10.1111/j.1469-8749.2007.00361.x

Hu, L., & Bentler, P. (1999). Cutoff criteria for fit indices in covariance structures analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6,* 1-55. doi: 10.1080/10705519909540118

Jöreskog, K., & Sörbom, D. (1993). *Introduction. In testing structural equation models.* Newbury Park, CA: Sage

Kaiser, H. (1970). A second generation little jiffy. *Psychometrika, 35,* 401-415. doi: 10.1007/BF02291817

Kaiser, H. (1974). An index of factorial simplicity. *Psychometrika, 39,* 31-36. doi: 10.1007/BF02291575

Kessler, R. C., Andrews, G., Hiripi, E., Mroczek, D. K., Normand, S . L., …Zaslavsky, A. M. (2002). Short screening scales to monitor population prevalences and trends in non-specific psychological distress. *Psychological Medicine, 32,* 959-976. doi: 10.1017/S0033291702006074

Kloosterman, P. H., Keefer, K. V., Kelley, E. A., Summerfeldt, L. J., & Parker, J. D. A. (2011). Evaluation of the factor structure of the autism-spectrum quotient. *Personality and Individual Differences, 50,* 310-214. doi:10.1016/j.paid.2010.10.015

Knicmeyer, R.C., Wheelwright, S., Baron-Cohen, S. B. (2007). Sex-typical play: Masculinization/defeminisation in Girls with an Autism Spectrum Condition. *Journal of Autism and Developmental Disorders, 38,* 1028-1035. doi: 10.1007/s10803-007-0475-0

Koenig, K., & Tsatsanis, K. D. (2005). Pervasive developmental disorders in girls. In D. J. Bell, S. L. Foster, & E. J. Mash (Eds.), *Handbook of behavioural and emotional problems in girls* (pp. 211-237). New York, NY: Kluwer Academic/Plenum Publishers.

Kopp, S., & Gillberg, G. (1992). Girls with social deficits and learning problems: Autism, atypical Asperger syndrome of a variant of these conditions. *European Child & Adolescent Psychiatry, 1,* 89-99. doi: 10.1007/BF02091791

Kopp, S., & Gillberg, G. (1997). Selective mutism: A population-based study: A research note. *Journal of Child Psychology and Psychiatry, 38,* 257-262. doi: 10.1111/j.1469-7610.1997.tb01859.x

Kopp, S., & Gillberg, C. (2011). The Autism Spectrum Screening Questionnaire (ASSQ)-Revised Extended Version (ASSQ-REV): An instrument for better capturing the autism phenotype in girls? A preliminary study involving 191 clinical cases and community controls. *Research in Developmental Disabilities. 32,* 2875-2888. doi:10.1016/j.ridd.2011.05.017

Kopp, S., & Gillberg, C. (2011). The Autism Spectrum Screening Questionnaire (ASSQ)-Revised Extended Version (ASSQ-REV): An instrument for better capturing the autism phenotype in girls? A preliminary study involving 191 clinical cases and community controls. *Research in Developmental Disabilities. 32,* 2875-2888. doi:10.1016/j.ridd.2011.05.017

Lai, M. C., & Baron-Cohen, S. (2015). Identifying the lost generation of adults with autism spectrum conditions. *Lancet Psychiatry, 2,* 1013-1027. doi: 10.1016/S2215-0366(15)00277-1

Lai, M. C., Lombardo, M., Suckling, J., Ruigrok, A., Chakrabarti, B., Ecker, C., … Baron-Cohen, S. (2013). Biological sex affects the neurology of autism. *Brain, 136,* 2799-2815. doi:10.1093/brain/awt216

Loomes, R., Hull, L., Mandy, W. P. L. What is the males-to-female ratio in Autism Spectrum Disorder? A systematic review and meta-analysis. *Journal of American Academy of Child and Adolescent Psychiatry, 56,* 466-474. doi: 10.1016/j-jaac.2017.03.013

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Leventhal, B. L., DiLavore, P. C. … Rutter, M. (2000). The autism diagnostic observation schedule – generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders, 30,* 205-223. doi: 10.1023/A:1005592401947

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods, 1,* 130-149. doi: 10.1037/1082-989X.1.2.130

Mandy, W., Chilvers, R., Chowdhury, U., Salter, G., Seigal, A., & Skuse, D. (2011). Sex differences in autism spectrum disorder: Evidence from a large sample of children and adolescents. *Journal of Autism and Developmental Disorders, 42,* 1304-1313. doi: 10.1007/s10803-011-1356-0

Maxim, L. D., Niebo, R., & Utell, M. J. (2014). Screening tests: A review with examples. *Inhalation Toxicology, 26,* 811-828. doi: 10.3109/08958378.2014.955932

McLennan, J. D., Lord, C., & Schopler, E. (1993). Sex differences in higher functioning people with autism. *Journal of Autism and Developmental Disorders, 23,* 217-227. doi:10.1007/BF01046216

Miniscalo, C., & Sanberg, A. D. (2010). Basic reading skills in Swedish children with late developing language and with or without autism spectrum disorder or ADHD. *Research in Developmental Disabilities, 31,* 1054-1061. doi: 10.1016/j.ridd.2010.04.004

Muthen, B., & Kaplan, D. (1985). Comparison of some methodologies for the factor analysis of non-normal Likert variables. *British Journal of Mathematical and Statistical Psychology, 38,* 171-189.

Nicholas, J. S., Charles, J. M., Carpenter, L. A., King, L. B., Jenner, W., & Spratt, E. G. (2008). Prevalence and characteristics of children with autism-spectrum disorders. *Annals of Epidemiology, 18,* 130-136. doi:10.1111/dmcn. 22_13241

Nilsson, E. W., Gillberg, C., Gillberg, I. C., & Råstam, M. (1999). Ten-year follow-up of adolescent-onset anorexia: Personality disorders. *Journal of the American Academy of Child & Adolescent Psychiatry, 38,* 1389-1395. doi: 10.1097/00004583-199911000-00013

Pilowsky, T., Yirmiya, N., Shulman, C., & Dover, R. (1998). The autism diagnostic interview-revised and the childhood autism rating scale: Differences between diagnostic systems and comparison between genders. *Journal of Autism and Developmental Disorders, 28,* 143-151. doi: 10.1023/A:10260922632466

Posserud, M. B., Lundervold, A. J., & Gillberg, C. (2006). Autistic features in a total population of 7-9-year-old children assessed by the ASSQ (Autism Spectrum Screening Questionnaire). *Journal of Child Psychology and Psychiatry, and Allied Disciplines, 47*, 167-175. doi: 10.1111/j.1469-7610.2005.01462.x

Posserud, M. B., Lundervold, A. J., Steijnen, M. C., Verhoeven, S., Stormark, K. M., & Gillberg, C. (2008). Factor analysis of the autism spectrum screening questionnaire. *Autism, 12*, 99-112. doi: 10.1177/1362361307085268

Råstam, M. (2008). Eating disturbances in autism spectrum disorders with a focus on adolescent and adult years. *Clinical Neuropsychiatry, 5,* 31-42.

Rutter, M., Caspi, A., & Moffitt, T. (2003). Using sex differences in psychopathology to study causal mechanisms: Unifying issues and research strategies. *Journal of Child Psychology & Psychiatry & Allied Disciplines, 44,* 1092. doi:10.1111/1469-7610.00194

Siklos, S., & Kerns, K. A. (2007). Assessing the diagnostic experiences of a small sample of parents and children with autism spectrum disorders. *Research in Developmental Disabilities, 28,* 9-22. doi: 10.1016/j.ridd.2005.09.003

Skuse, D. H., Mandy. W., Steer, C., Miller, L.L., Goodman, R., Lawrence, K., … Golding, J. (2009). Social communication competence and functional adaptation in a general population of children: Preliminary evidence for sex-by-verbal IQ differential risk. *Journal of the American Academy of Child and Adolescent Psychiatry, 48,* 128-137. doi:10.1097/CHI.0b013e31819176b8

Socha, A., Cooper, C. A., & McCord, D. M. (2010). Confirmatory factor analysis of the M5-50: An implementation of the international personality item pool item set. *Psychological Assessment, 22,* 43-49. doi: 10.1037/a0017371

Steiger, J. H. (2007). Understanding the limitations of global fit assessment in structural equation modelling. *Personality and Individual Differences, 42,* 893-898. doi: 10.1016/j.paid.2006.09.017

Steiner, D. L., & Norman, G. R. (2005). *Health Measurement Scales a Practical Guide to Their Development and Use.* New York: Oxford University Press

Tabachncik, B. G., & Fidell, L. S. (2013). *Using multivariate statistics, 6th ed.* Boston: Pearson

Tsai, L., & Beisler, J. (1983). The development of sex differences in infantile autism. *The British Journal of Psychiatry, 142,* 373. doi: 10.1192/bjp.142.4.373

Volmar, F. R., Szatmari, P., & Sparrow, S. S. (1993). Sex differences in pervasive developmental disorders. *Journal of Autism and Developmental Disorders, 23,* 579-591. doi: 10.1007/BF01046103

Wakabayashi, A., Baron-Cohen, S., Wheelwright, S., Goldenfeld, N., Delaney, J., Fine, D., … Weil, L. (2006). Development of short forms of the Empathy Quotient (EQ-Short) and the Systemizing Quotient (SQ-Short). *Personality and Individual Differences, 41,* 929-940. doi: 10.1016/j.paid.2006.03.017

Wheaton, B., Muthen, B., Alwin, D. F., & Summers, G. (1977). Assessing reliability and stability in panel models. *Sociological Methodology, 8,* 84-136. doi: 10.2307/270754

Wheelwright, S., Baron-Cohen, S., Goldenfeld, N., Delaney, J., Fine, D., Smith, R., … Wakabayashi, A. (2006). Predicting Autism Spectrum Quotient (AQ) from the Systemizing Quotient-Revised (SQ-R) and Empathy Quotient (EQ). *Brain Research, 1079,* 47-56. doi: 10.1016/j.brainres.2006.01.012

Williams, J. G., Allison, C., Scott, F. J., Bolton, P. F., Baron-Cohen, S., Matthews, F., & Brayne, C. (2008). The Childhood Autism Spectrum Test (CAST): Sex differences. *Journal of Autism and Developmental Disorders, 38,* 1731-1739. doi: 10.1007/s10803-008-0556-6

Wing, L., & Gould, J. (1979). Severe impairments of social interaction and associated abnormalities in children: Epidemiology and classification. *Journal of Autism and Developmental Disorders, 9,* 11-29

Woodbury-Smith, M. R., Robinson, J., Wheelwright, S., & Baron-Cohen, S. (2005). Screening adults for asperger syndrome using the AQ: A preliminary study of its diagnostic validity in clinical practice. *Journal of Autism and Developmental Disorders, 55,* 331-335. doi: 10.1007/s10803-005-3300-7

Zucker, N. L., Losh, M., Bulik, C. M., LaBar, K. S., Piven, J., & Pelphrey, K. A. (2007). Anorexia nervosa and autism spectrum disorders: Guided investigation of social cognitive endophenotypes. *Psychological Bulletin, 133,* 976-1006. doi: 10.1037/0033-2909.133.6.976

Appendix A: Measures

Demographic Questionnaire

1. Please indicate your gender
    - Male
    - Female
    - Other
2. Please indicate your age in years. (E.g., 30)
3. Please indicate the highest level of education that you have completed.
    - Primary School
    - High School or equivalent
    - TAFE/Trade School/Vocational School
    - Undergraduate University/College (E.g., Bachelor degree)
    - Postgraduate University/College (E.g., Phd or Masters degree)
4. Was your education completed through:
    - Mainstream schooling
    - Special needs delivery
5. Have you ever been told by a professional that you have an intellectual disability?
6. Have you ever been diagnosed or labelled as having:
    - Autism Spectrum Disorder or "Autism"
    - Asperger's or Aspergers Syndrome
    - Pervasive Developmental Disorder
    - No
7. If yes, who provided the diagnosis? *
    - A Psychologist
    - A Psychiatrist
    - A Paediatrician
    - Other
8. What age were you when you received the diagnosis? *
9. Prior to diagnosis, how many medical health practitioners did you see regarding your challenges? *
10. Have you been diagnosed with any other psychiatric of neurological disorder including epilepsy? Please specify: *

Note. * Denotes questions that were only asked to those participants who did **not** answer "No" to question 6.

Autism Spectrum Quotient

This scale has been removed due to copyright restrictions.

The New AQ-Revised items (prior to item removal resulting from EFA results):

Please indicate the extent to which you agree with each statement.

51. When I was young, I spent more time in my own fantasy world that in reality.	definitely agree	slightly agree	slightly disagree	definitely disagree
52. I have been told that I have an unusual tone of voice (childish, high pitched or hoarse).	definitely agree	slightly agree	slightly disagree	definitely disagree
53. In social situations, when I was young, I would ask very few questions of others. *	definitely agree	slightly agree	slightly disagree	definitely disagree
54. In social situations, I wear a mask or take on a persona other than my own.	definitely agree	slightly agree	slightly disagree	definitely disagree
55. I have not learnt how to act in social situations by observing and mimicking other people. *	definitely agree	slightly agree	slightly disagree	definitely disagree
56. Sometimes I feel confused about my identity.	definitely agree	slightly agree	slightly disagree	definitely disagree
57. When I was young, most of my friends were my age. *	definitely agree	slightly agree	slightly disagree	definitely disagree
58. I like animals, nature, the arts and/or classical literature.	definitely agree	slightly agree	slightly disagree	definitely disagree
59. When I feel angry or stressed, I become outwardly aggressive. *	definitely agree	slightly agree	slightly disagree	definitely disagree
60. During school, I was described as passive, shy and/or compliant.	definitely agree	slightly agree	slightly disagree	definitely disagree
61. I typically cope with demands by tackling them head on. *	definitely agree	slightly agree	slightly disagree	definitely disagree
62. I have actively learnt how to socialise through different media sources e.g., magazines, characters on television, books on body language and novels.	definitely agree	slightly agree	slightly disagree	definitely disagree
63. Friendships were easy for me during adolescence. *	definitely agree	slightly agree	slightly disagree	definitely disagree

64. During my schooling years, I recall desiring the exclusive attention of one particular friend.	definitely agree	slightly agree	slightly disagree	definitely disagree
65. Socialising is easy; I never feel lonely and desire friendship. *	definitely agree	slightly agree	slightly disagree	definitely disagree

Note. * Denotes a reverse scored item

Empathy Quotient – Short

This scale has been removed due to copyright restrictions.

Systemising Quotient – Short

This scale has been removed due to copyright restrictions.

Impact Scale

In the past month, how much trouble have you experienced in the following areas?

If a specific question does not apply, mark "Not at all"

1. Working/Studying	Not at all	Just a little	Quite a bit	Very much
2. Sleeping	Not at all	Just a little	Quite a bit	Very much
3. Attention and concentration	Not at all	Just a little	Quite a bit	Very much
4. Interacting with others	Not at all	Just a little	Quite a bit	Very much
5. Feeling anxious and stressed	Not at all	Just a little	Quite a bit	Very much
6. Relationships	Not at all	Just a little	Quite a bit	Very much
7. Planning my work and/or daily activities	Not at all	Just a little	Quite a bit	Very much

Kessler Psychological Distress Scale (K-6)

The following questions ask about how you have been feeling during the past 30 days. For each question, please choose the option that best describes how often you had this feeling.

During the past 30 days, about how often did you feel …

1.	…nervous?	None of the time	A little of the time	Some of the time	Most of the time	All of the time
2.	…hopeless?	None of the time	A little of the time	Some of the time	Most of the time	All of the time
3.	…restless or fidgety?	None of the time	A little of the time	Some of the time	Most of the time	All of the time
4.	…so depressed that nothing could cheer you up?	None of the time	A little of the time	Some of the time	Most of the time	All of the time
5.	…that everything was an effort?	None of the time	A little of the time	Some of the time	Most of the time	All of the time
6	…worthless?	None of the time	A little of the time	Some of the time	Most of the time	All of the time

Note. This scale is in the public domain. Acknowledgement is made to Ronald C. Kessler, Phd, Harvard Medical School, Department of Health Care Policy, 180 Longwood Ave Boston, MA.

Appendix B: Email Correspondence with Permission to Adapt the Autism Quotient

This correspondence has been removed for reasons pertaining to privacy

Appendix C: Item Selection Process and Relevant Statistics for Phase 1 of the study

Questionnaire Completed by ASD Practitioners

1. Please indicate which category your title falls into:
   - Psychologist
   - Clinical Psychologist
   - Psychiatrist
2. Approximately how many individuals on the Autism Spectrum have you treated and/or assessed?
3. Approximately how many **females** on the Autism Spectrum have you treated and/or assessed?
4. Based on your knowledge and interaction with females on the Autism Spectrum, on a scale of 1 to 5, with on being extremely irrelevant to 5 being extremely relevant, please rate your confidence in each item being relevant to the female ASD presentation.

1.	When I was young, I spent more time in my own fantasy world than in reality	1	2	3	4	5
2.	Preferring practical clothing, I have a carefree attitude towards dress and appearance	1	2	3	4	5
3.	I have been told that I have an unusual tone of voice (Childish, high pitched or hoarse)	1	2	3	4	5
4.	In social situations, when I was young, I would ask many questions of others	1	2	3	4	5
5.	In social situations, I wear a mask or take on a persona other than my own	1	2	3	4	5
6.	I would describe myself as a determined person	1	2	3	4	5
7.	I have learnt how to act in social situations by observing and mimicking other people	1	2	3	4	5
8.	Sometimes I feel confused about my identity	1	2	3	4	5
9.	When I was young, I had difficulty sleeping	1	2	3	4	5
10.	I have a different relationship with food compared to others	1	2	3	4	5
11.	I like animals, nature, the arts and/or classic literature	1	2	3	4	5
12.	When I feel angry or stressed, I become withdrawn	1	2	3	4	5
13.	During school, I was described as passive, shy and/or compliant	1	2	3	4	5
14.	When I was young, I had difficulty sleeping	1	2	3	4	5
15.	I have difficulty completing simple daily activities	1	2	3	4	5
16.	I typically cope by avoiding demands	1	2	3	4	5
17.	I have actively learnt how to socialise through different media sources e.g., magazines, characters on television, books on body language and novels	1	2	3	4	5

18.	Friendships were difficult for me during adolescence	1	2	3	4	5
19.	During my schooling years, I recall desiring the exclusive attention of one particular friend	1	2	3	4	5
20.	Although socialising is exhausting, sometimes I feel lonely and desire friendship	1	2	3	4	5

5. If you believe that these are other areas that are not covered by the aforementioned items, but are highly relevant to the female ASD presentation, please indicate these in the space below.

Details of the Professional Experience of the ASD Practioners

Of the 30 participating practioners, 12 identified as clinical psychologists and 18 identified as psychologists. No psychiatrists participated. The approximate mean number of individuals on the Autism Spectrum that professionals reported assessing and/or treating was 659.72 (*SD* = 1853.22). The approximate mean number of females on the Autism Spectrum that they reported assessing and/or treating was 162.55 (*SD* = 391.93). These means are approximate, as the participants were asked to record “approximately” how many ASD individuals and females they had assessed/treated and some participants placed an addition (+) sign after their responses. Thus, it is likely that these figures are an underrepresentation of the practioners’ experience.

Table 7
Relevant Statistics for Item Selection

Question number	Mean	Variance	Standard Deviation	Median	Mode	Comparator score for each question
18	4.66	0.29	0.54	5.00	5.00	5.00
7	4.52	0.39	0.62	5.00	5.00	5.00
5	4.41	0.45	0.67	5.00	5.00	5.00
17	4.31	0.35	0.59	4.00	4.00	5.00
8	4.10	0.99	0.99	4.00	5.00	5.00
20	4.03	0.59	0.76	4.00	4.00	5.00
1	3.97	0.52	0.72	4.00	4.00	4.00
19	3.93	0.82	0.91	4.00	4.00	4.00
13	3.79	0.65	0.80	4.00	4.00	4.00
9	3.76	0.60	0.77	4.00	4.00	4.00
4	3.69	1.80	1.34	4.00	5.00	4.00
12	3.66	0.85	0.92	4.00	3.00	3.00
3	3.62	0.93	0.96	4.00	4.00	3.00
11	3.62	1.48	1.22	4.00	4.00	5.00
16	3.41	1.00	1.00	3.00	3.00	4.00
15	**3.31**	**1.46**	**1.21**	**3.00**	**3.00**	**3.00**
6	**3.17**	**1.38**	**1.18**	**3.00**	**3.00**	**3.00**
14	**3.17**	**1.32**	**1.15**	**3.00**	**4.00**	**4.00**
2	**3.14**	**1.02**	**1.01**	**3.00**	**3.00**	**3.00**
10	**3.00**	**1.59**	**1.26**	**3.00**	**4.00**	**3.00**

Note. The average, variance, standard deviation, median and mode were calculated based on the responses of 29 practitioners. As one practitioner indicated that they had assessed/treated 10 000 individuals with ASD and 2000 Females with ASD, the responses from this participant were separated from the group analysis and evaluated independently as a comparator to the larger group. These scores can be viewed in the far-right column.

Process of Elimination

The bolded items in Table 7 are those that would be eliminated. These items were selected for elimination based on their comparatively low mean, mode and median scores and high variance scores across both the 29 respondents and comparator, as well as their request for a judgement about qualities rather than behaviour. Following the elimination process, 15 items from the original 20 remained.

Appendix D: Scree Plots of Eigenvalues

Figure 3

EFA Scree Plot for Males and Females Combined on the AQ-Revised with All 15 New Items

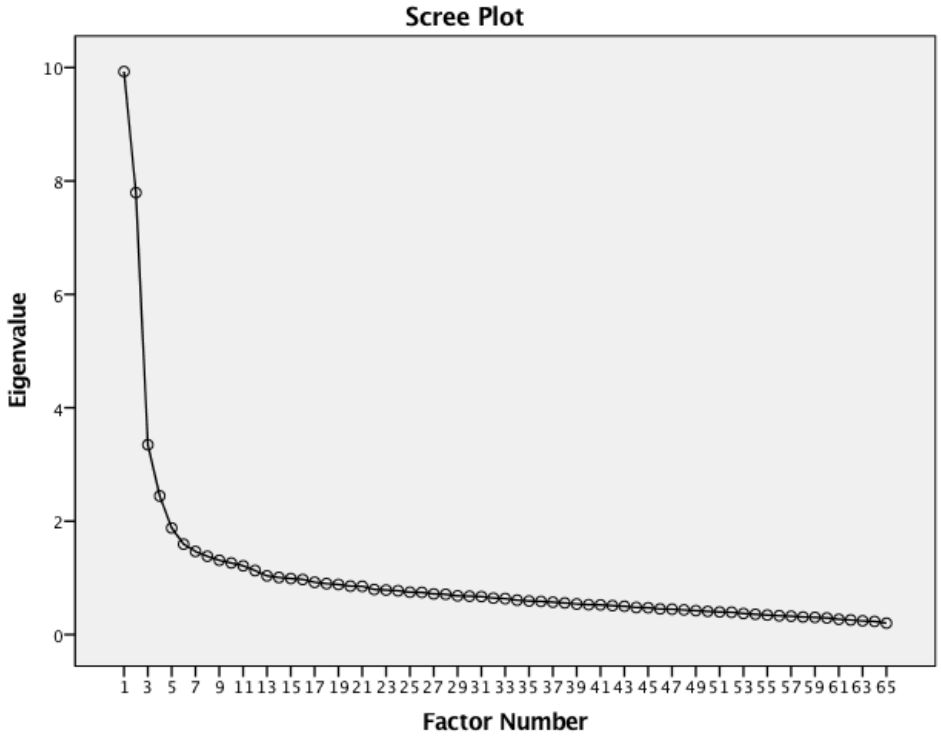


Figure 4

EFA Scree Plot for Females on the AQ-Revised with All 15 New Items

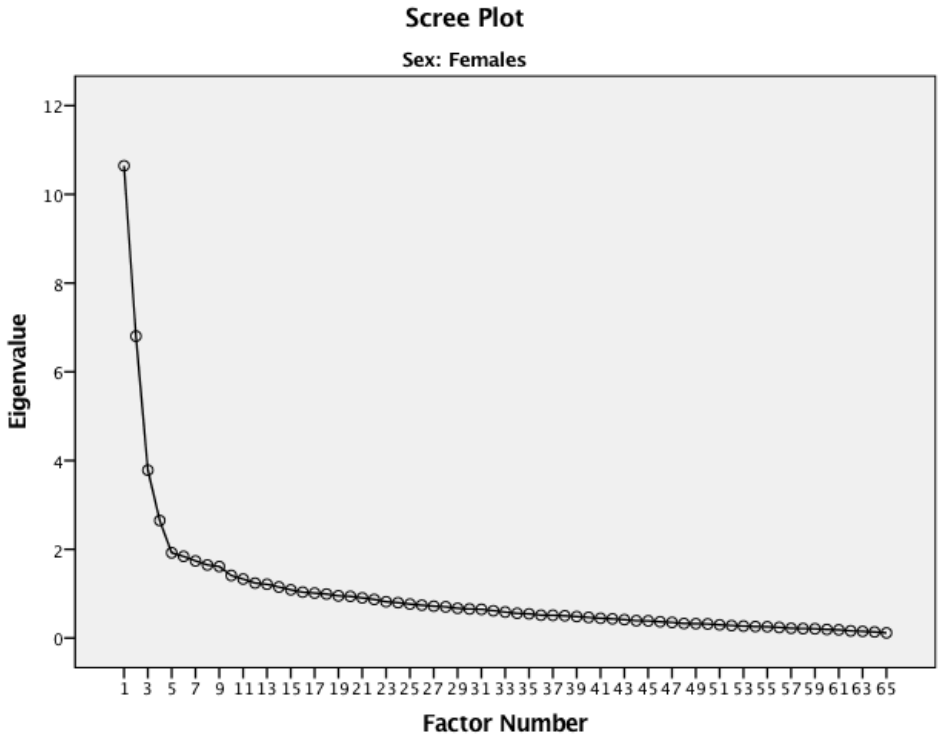


Figure 5

EFA Scree Plot for Males on the AQ-Revised with All 15 New Items

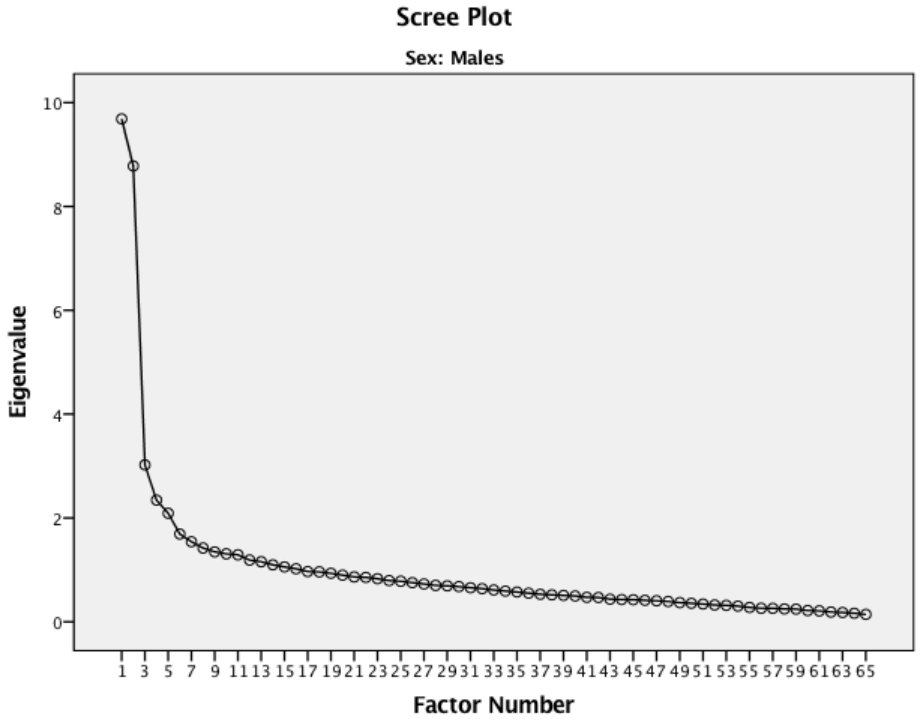


Figure 6

EFA Scree Plot for Males and Females Combined on the Original AQ

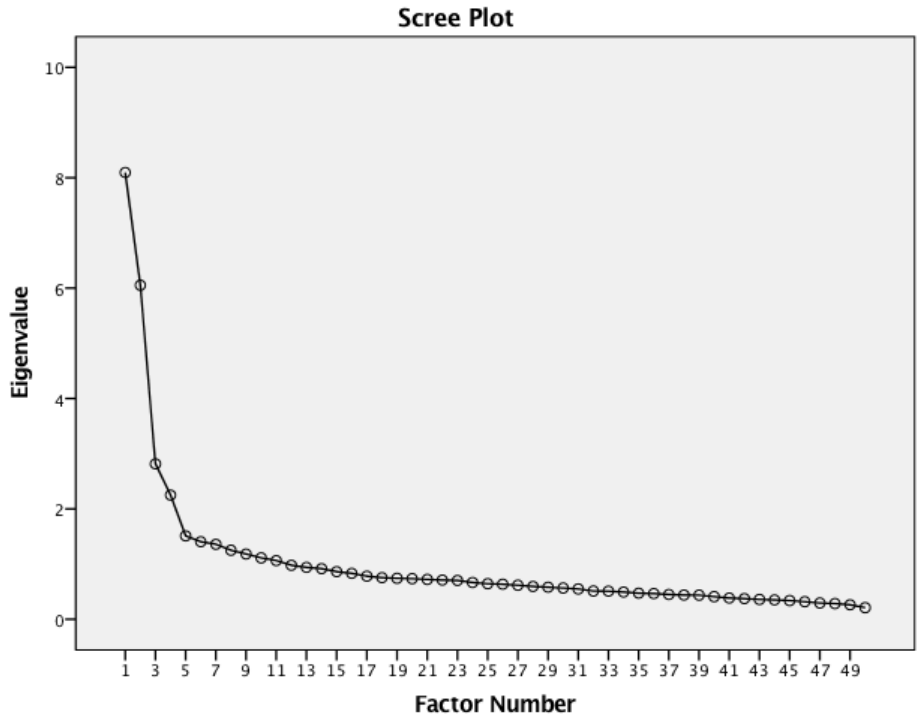


Figure 7

EFA Scree Plot for Females on the Original AQ

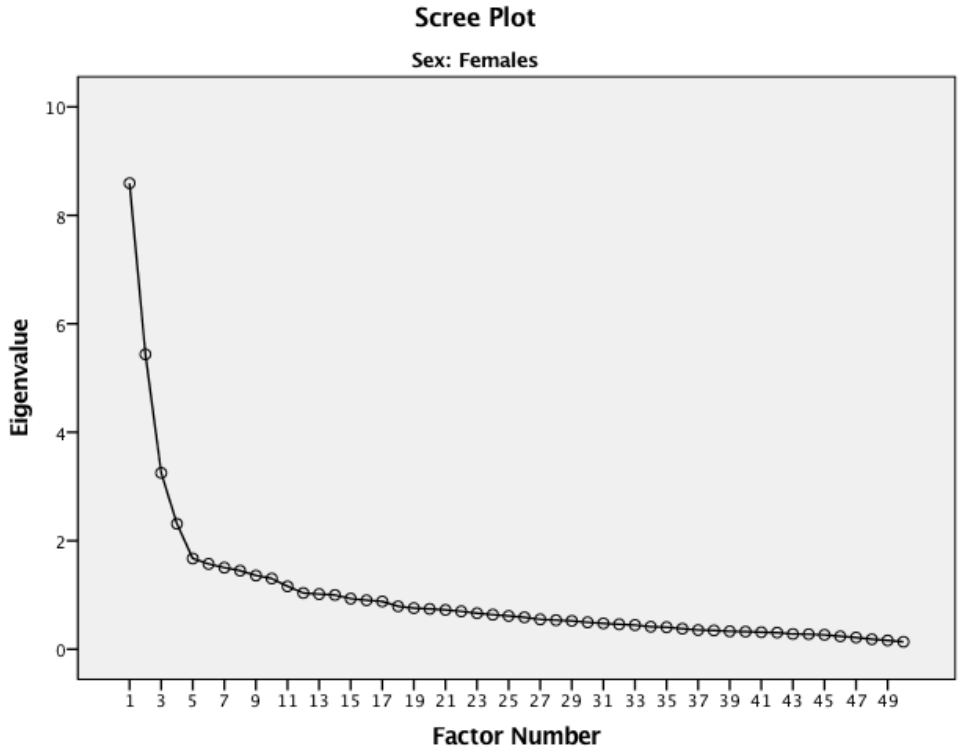
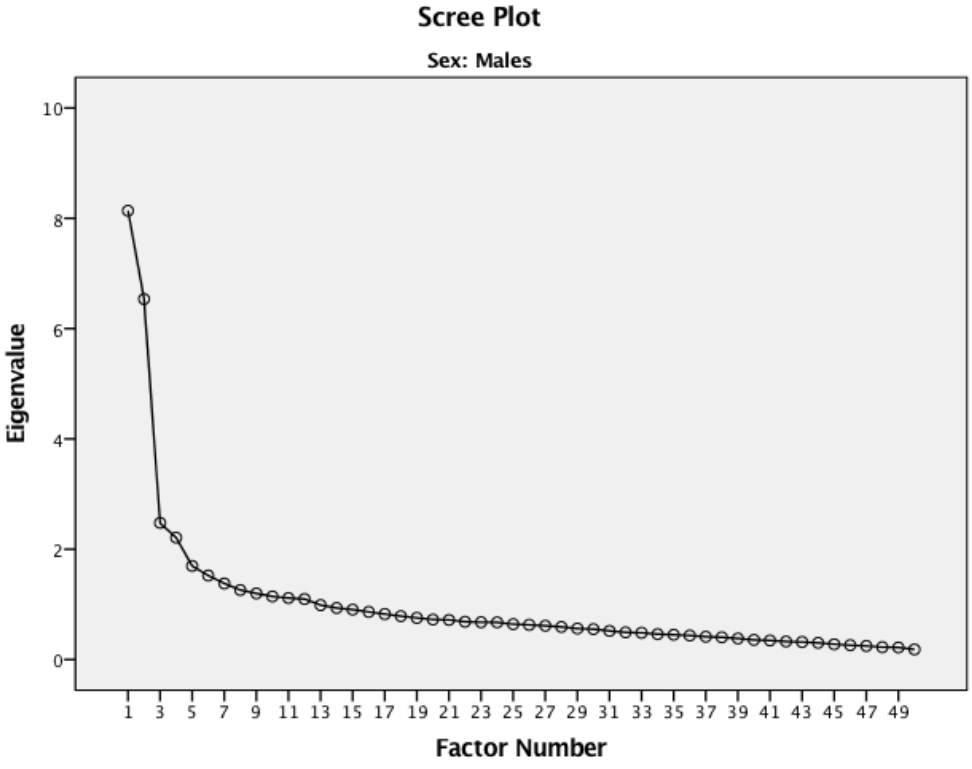


Figure 8

EFA Scree Plot for Males on the Original AQ



Appendix E: Histograms for the AQ-Revised and 2 Factors

Figure 9

Histogram of the Total AQ-Revised Score for the Entire Sample

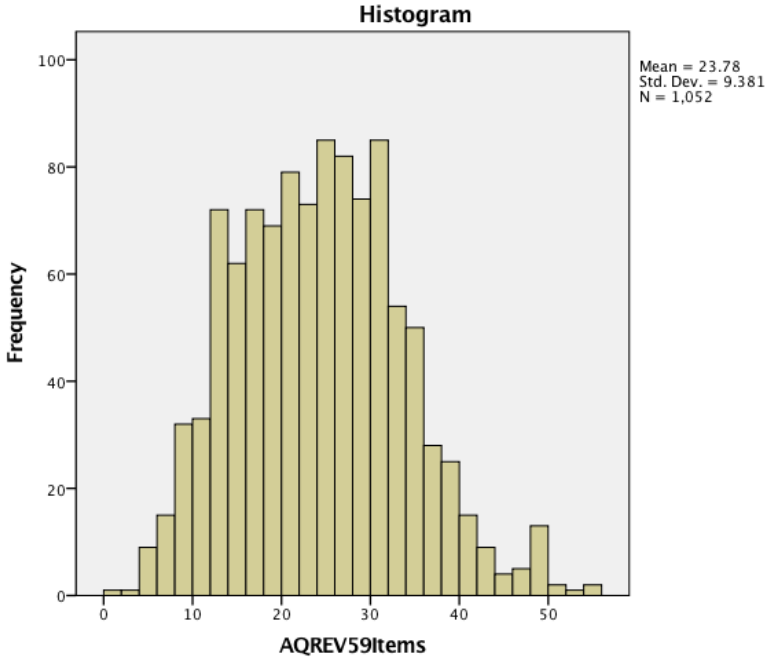


Figure 10

Histogram of the Socialisation Factor for the Entire Sample

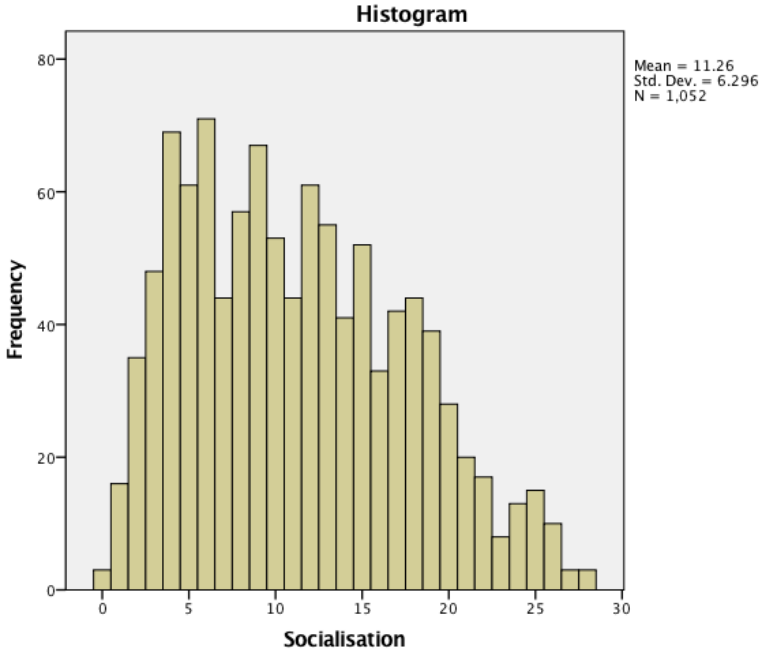
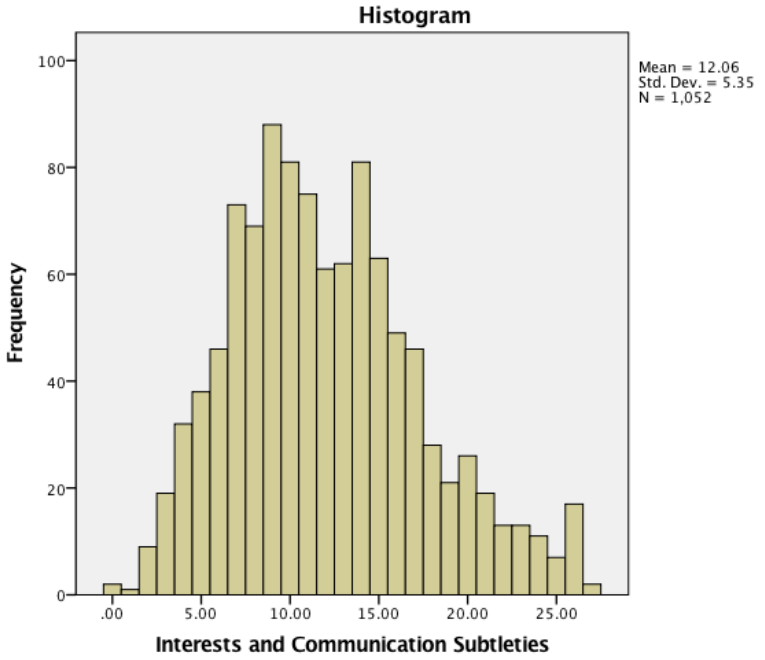


Figure 11

Histogram of the Interests and Communication Subtleties Factor for the Entire Sample



Appendix F: Means and Standard Deviations for ASD and Non-ASD on the Nine New AQ-Revised Items

Table 8

Means and Standard Deviations for ASD and Non-ASD on the 9 New AQ-Revised Items

Item	ASD Female (n = 23)	ASD Male (n = 22)	Non-ASD Female (n = 551)	Non-ASD Male (n = 456)
Q65. Socialising is easy; I never feel lonely and desire friendship. *	.91 (.29)	.50 (.51)	.54 (.50)	.46 (.50)
Q64. During my schooling years, I recall desiring the exclusive attention of one particular friend.	.70 (.47)	.59 (.50)	.44 (.50)	.45 (.50)
Q63. Friendships were easy for me during adolescence. *	.65 (.49)	.64 (.49)	.38 (.49)	.35 (.48)
Q61. I typically cope with demands by tackling them head on. *	.87 (.34)	.45 (.51)	.30 (.46)	.30 (.46)
Q60. During school, I was described as passive, shy and/or compliant.	.70 (.47)	.64 (.49)	.58 (.49)	.54 (.50)
Q56. Sometimes I feel confused about my identity.	.61 (.50)	.50 (.51)	.40 (.49)	.35 (.48)
Q54. In social situations, I wear a mask or take on a persona other than my own.	.78 (.42)	.77 (.43)	.35 (.48)	.36 (.48)
Q52. I have been told that I have an unusual tone of voice (childish, high pitched or hoarse).	.26 (.45)	.45 (.51)	.18 (.39)	.22 (.42)
Q51. When I was young, I spent more time in my own fantasy world that in reality.	.70 (.47)	.73 (.46)	.46 (.50)	.43 (.50)

Note. Standard Deviations are in parentheses.

Author Note

At the time of the study, Alice Macoun was affiliated with the Australian National University in Canberra, Australia. Professor Richard O'Kearney was also affiliated with the Australian National University in Canberra, Australia.

No changes have been made regarding the author affiliation since the time of the study.

We thank Boris Bozumic and Conal Monaghan for their advice regarding the Confirmatory Factor Analysis component of the paper. We also thank the reviewers for improving this paper through feedback.

Correspondence concerning this article should be addressed to Alice Macoun, Research School of Psychology, Australian National University, Acton ACT 2601, Australia. Phone: 0409 172 116; Email: alice.macoun@anu.edu.au.