Measuring autistic traits in the general population: a systematic review of the Autism-Spectrum Quotient (AQ) in a nonclinical population sample of 6,900 typical adult males and females

Emily Ruzich1,2*, Carrie Allison2, Paula Smith2, Peter Watson3, Bonnie Auyeung2,4, Howard Ring1,5,6 and Simon Baron-Cohen2,5,7

Abstract

The Autism-Spectrum Quotient (AQ) is a self-report measure of autistic traits. It is frequently cited in diverse fields and has been administered to adults of at least average intelligence with autism and to nonclinical controls, as well as to clinical control groups such as those with schizophrenia, prosopagnosia, anorexia, and depression. However, there has been no empirical systematic review of the AQ since its inception in 2001. The present study reports a comprehensive systematic review of the literature to estimate a reliable mean AQ score in individuals without a diagnosis of an autism spectrum condition (ASC), in order to establish a reference norm for future studies. A systematic search of computerized databases was performed to identify studies that administered the AQ to nonclinical participant samples representing the adult male and female general population. Inclusion was based on a set of formalized criteria that evaluated the quality of the study, the usage of the AQ, and the population being assessed.

After selection, 73 articles, detailing 6,934 nonclinical participants, as well as 1,963 matched clinical cases of ASC (from available cohorts within each individual study), were analyzed. Mean AQ score for the nonclinical population was 16.94 (95% CI 11.6, 20.0), while mean AQ score for the clinical population with ASC was found to be 35.19 (95% CI 27.6, 41.1). In addition, in the nonclinical population, a sex difference in autistic traits was found, although no sex difference in AQ score was seen in the clinical ASC population. These findings have implications for the study of autistic traits in the general population. Here, we confirm previous norms with more rigorous data and for the first time establish average AQ scores based on a systematic review, for populations of adult males and females with and without ASC. Finally, we advise future researchers to avoid risk of bias by carefully considering the recruitment strategy for both clinical and nonclinical groups and to demonstrate transparency by reporting recruitment methods for all participants.

Keywords: Autism, Autism-spectrum quotient, Autistic traits, Mean score, Sex differences, Systematic review
Introduction

Autism was traditionally considered as a clinical condition distinct from the general population, but recent evidence suggests autistic traits are continuously distributed across the population [1-3]. From observed data of measured autistic traits, people with a diagnosis of an autism spectrum condition (ASC) - at least those who have average IQ or above - score at the extreme end of this distribution [4]. It may be that ‘syndromic’ forms of autism, which often entail comorbid learning disability (or below average IQ) and a known genetic mutation, are discontinuous with autistic traits in the general population, but here the focus is on the general population without learning disability. The Autism-Spectrum Quotient (AQ) is widely used in research and clinical practice to quantify autistic traits. The AQ was first developed as a self-report measure for adults [5] and subsequently as a parent-report measure for adolescents (aged 12 to 15 years) [6] and for children (aged 4 to 11 years) [7]. A toddler version also exists (Q-CHAT (Quantitative Checklist for Autism in Toddlers) [8]). The AQ has 50 items, which are divided into five subscales consisting of 10 items each that assess domains of cognitive strengths and difficulties related to ASC: communication, social skills, imagination, attention to detail and attention switching. While the AQ is not the only research tool used to measure autistic traits (for example, see the SRS (Social Responsiveness Scale) [9]), it has several advantages over other measures, including subscales for both social and nonsocial aspects of behavior and cognition and a format that is brief, self-administered, and forced-choice.

The AQ was designed for adults with average IQ or above [5], who comprise at least 50% of the autism spectrum [10]. Individuals are instructed to respond to each of the 50 items with one of four responses: ‘definitely agree’, ‘slightly agree’, ‘slightly disagree’, and ‘definitely disagree’. Responses are scored using a binary system, where an endorsement of the autistic trait (either mildly or strongly) is scored as a +1, while the opposite response is scored as a 0, leading to a maximum score on the AQ of 50. An alternative scoring system has also been employed that uses a 4-point Likert scale [11]. AQ items are counterbalanced to avoid a response bias, so that half of the ‘agree’ responses and half of the ‘disagree’ responses endorse the autistic trait. The AQ includes questions about both ability and preference. The questionnaire is not suitable for individuals with low IQ, low verbal ability, or language impairment, as it relies on receptive understanding of the 50 questions.

The AQ was originally validated in 2001 in adult males and females with Asperger Syndrome (AS) and high-functioning autism (HFA), in scientists versus nonscientists in Cambridge University students, in winners of the mathematical Olympiad (because of the finding that autism may be genetically linked to an aptitude for ‘systemizing’ [12-14]), and nonstudent individuals drawn from the general population. This study found that the total AQ score and its five subscale scores are normally distributed and have demonstrated good test-retest reliability, good internal consistency [5], and that the measure has acceptably high sensitivity and specificity: at a cut-off score of 26, 83% of patients were correctly identified (sensitivity 0.95, specificity 0.52, positive predictive value 0.84, negative predictive value 0.78), while a cut-off score of 32 correctly identifies 76% of patients (sensitivity 0.77, specificity 0.74) [11,15] when the AQ is used in a referred clinical sample.

These results indicate that the AQ is a sensitive measure of autistic traits in the general population, implying that traits reaching a clinical level in autism also exist to a lesser degree in nonclinical counterparts [5]. Within families, AQ score has shown heritability, which is in line with genetic evidence suggesting the heritability of autism [16]. Further, some (but not all) parents of children with autism show a subclinical set of characteristics or traits that index familiarity and/or genetic liability to autism [17,18]. This is referred to as the ‘Broader Autism Phenotype’ (BAP). There is a consistent sex difference in mean AQ score, such that typical males score significantly higher than typical females, while people of both sexes with ASC score at the extreme high end of the scale, in line with the extreme male brain (EMB) theory of autism [19,20].

The AQ is also widely referenced: a recent search of Google Scholar indicated that the original publication has been cited over 1,250 times. The present study reports the first large-scale systematic review of published AQ data over the last 13 years from adults with and without a diagnosis of ASC, in order to characterize the distribution of autistic traits in adult males and females and to contrast scores from clinical versus nonclinical samples. The specific goal is to establish a reliable mean AQ score in nonclinical controls, which can then be used as a guideline for researchers to define their control groups in future studies that compare people with and without a clinical diagnosis of ASC, as well as to other specially selected groups.

Review

Methods

Identification of relevant literature

Citation indexing databases Scopus, PubMed (Medline), PsycINFO, and Web of Science were queried for articles utilizing the AQ. Titles, abstracts and keywords were searched for (“autism quotient”) OR (“autis* spectrum quotient”) OR (“AQ” AND “autism”). Exploded MeSH terms were not used because of the narrow target of interest; studies were only considered if they explicitly mentioned the AQ. However, an additional search of
Scopus and Web of Science was performed by which all peer reviewed journal articles citing the 2001 Baron-Cohen paper introducing the measure were retrieved. The two searches were merged; the citation search delivered 837 hits and the keyword search delivered 321 hits, 287 of which were retrieved by both methods.

Titles and abstracts and then full text articles, were reviewed. Inclusion criteria specified that the study had to include peer-reviewed empirical research (excluding all meta-analyses, literature reviews, book chapters, conference proceedings, etcetera.), be published in English, that the AQ had to be the 50-item AQ adult self-report (and not the AQ-Child, AQ-Adolescent or any of the abbreviated versions of the AQ), and that there was evidence that the English-language version of the AQ had been administered rather than any translations. The non-clinical participant sample had to include both males and females recruited from the population, with a mean age of 18 years or older.

Exclusion criteria were applied that assessed the quality of the study, the usage of the AQ, and the population being assessed. See Figure 1 for the selection process. Articles were excluded if they were case reports, studies containing fewer than 10 participants, or if the study specifically recruited participants who were immediate family members of an individual with ASC or patients with a particular mental or physical disorder or condition. In addition, due to findings from within the original AQ publication indicating the potential for academic disciplines to score more highly on the AQ, and in an effort to remove confounding variables such as age and education level, articles were excluded if participants had been recruited exclusively from within a university (though partial university recruitment was acceptable if
authors indicated that an effort was made to recruit from outside the academic community). Articles were also excluded if an AQ cut-off score was imposed when delineating the control or nonclinical group. Where it was unclear whether an article met eligibility criteria, the article was retained. A number of research groups frequently recruit participants from the same database, which may potentially lead to the same individuals’ AQ scores being duplicated in analyses across more than one publication; to guard against the risk of duplication, articles from the same research group were assessed. If authors used similar phrasing in describing the recruitment process or explicitly stated that participants were drawn from the same database, the publication with the largest population group was included in analysis while the rest were excluded. Finally, several articles published in the same year by the same authors contained identical AQ scores and numbers of participants; in these rare cases, the earliest instance was included while the later publications were conservatively excluded.

In a number of instances, authors indicated that participants had completed the AQ but complete data sets were not reported. For 36 papers, authors were contacted for clarification or more information (11 articles were lost due to lack of a response). The deadline for data queries from authors and for literature searching was Monday, 14 July 2014. From the literature search and screening process 73 articles (reporting 78 independent studies) met the inclusion criteria.

**Inter-rater agreement**
The first author (ER) performed 100% of the literature search, quality assessment, and data extraction. In order to assess reliability of this process, approximately 10% of the results returned by the literature search were examined by authors CA, PS, and SBC. Each of the second reviewers received a random sample of 30 articles for evaluation (totaling 90). Where it was unclear whether an article met eligibility criteria, the article was discussed among the research team and if agreement was reached, it was retained for inclusion in the analysis. Initial percentage inter-rater agreement was respectively 97%, 90%, and 90%; after a resolution process, all disagreements between the lead author and the second raters were resolved in favor of the first author.

**Extraction of data from included papers**
The following information was recorded:

1. Number of participants, delineated by sex if reported
2. Mean and standard deviation of AQ score for males, females, and the sexes combined
3. Range of AQ score, if reported
4. Test for normality, if reported
5. Mean and standard deviation of participants’ age
6. Recruitment strategy, if reported
7. A comment on whether the study excluded individuals who were first-degree relatives of someone with a diagnosis of ASC
8. Margin of error and confidence intervals were calculated for each study by ER
9. Mean AQ score was recorded if the study included a matched sample of participants with ASC.

**Data analysis**
This systematic review aims to explore the distribution of a single variable - AQ score - in a large nonclinical population sample; therefore in this case, a meta-analysis (for effect sizes) is not possible. Data were imported into R [21] for systematic analysis. The mean of means was calculated by differentially weighting the reported values by sample size using weighted linear regression. In addition, the range of standard distributions, along with minimum and maximum values, was reported, and confidence intervals for reported average AQ scores were calculated. These values were also calculated for studies reporting separate male and female AQ scores, which were then compared using meta-analytic techniques. Finally, a small subset of studies (N = 9) reported that, in addition to taking a personal medical history, participants were only eligible to be considered a part of the nonclinical population group if they also had no first-degree relatives with ASC. For these studies, a separate mean of means for the AQ was also calculated. The focus of this study concerned average performance on the AQ, but standard deviation was also noted from eligible publications. From these scores, pooled variance was calculated.

While the primary focus of the review was to explore AQ scores in a nonclinical population sample, AQ score from the ASC sample for the selected papers was also noted where relevant. These scores were analyzed in the same method reported above. In addition to the quantitative approach described above, the papers that met criteria were subjected to a qualitative reading of the recruitment strategy for the nonclinical participant sample. This was in an effort to provide a description of the background for the participants included in analysis.

**Results**

**Quantitative characterization of the Autism-Spectrum Quotient in a nonclinical population sample**
From a total of 73 articles reporting 78 studies that met eligibility criteria, data were recorded from 6,934 individual nonclinical participants. Table 1 describes the individual studies reviewed. See also inset plot for study AQ means and standard errors (Figure 2).

Descriptive statistics (weighted mean AQ, range of standard deviation, total range, 95% confidence interval,
### Table 1 Articles selected for review

<table>
<thead>
<tr>
<th>Nonclinical sample</th>
<th>Autism spectrum cases</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Males</td>
</tr>
<tr>
<td></td>
<td>AQ m (SD)</td>
</tr>
<tr>
<td>[5]</td>
<td>17.8 (6.8)</td>
</tr>
<tr>
<td>[22]</td>
<td>15.9 (7.4)</td>
</tr>
<tr>
<td>[23]</td>
<td>15 (6)</td>
</tr>
<tr>
<td>[24]</td>
<td>14 (5.4)</td>
</tr>
<tr>
<td>[25]</td>
<td>16.65 (6.81)</td>
</tr>
<tr>
<td>[26]</td>
<td>16.5 (6.38)</td>
</tr>
<tr>
<td>[27]</td>
<td>14.9 (8.58)</td>
</tr>
<tr>
<td>[28]</td>
<td>11.6 (5)</td>
</tr>
<tr>
<td>[29]</td>
<td>12.53 (5.78)</td>
</tr>
<tr>
<td>[30]</td>
<td>14.64 (7.46)</td>
</tr>
<tr>
<td>[31]</td>
<td>17.33 (8.79)</td>
</tr>
<tr>
<td>[32]</td>
<td>18.81 (7.8)</td>
</tr>
<tr>
<td>[33]</td>
<td>15.3 (6.1)</td>
</tr>
<tr>
<td>[34]</td>
<td>14.86 (4.03)</td>
</tr>
<tr>
<td>[35]</td>
<td>13.13 (5.46)</td>
</tr>
<tr>
<td>[17]</td>
<td>17.7 (6.9)</td>
</tr>
<tr>
<td>[36]</td>
<td>14.12 (5.78)</td>
</tr>
<tr>
<td>[37]</td>
<td>15.8 (7.2)</td>
</tr>
<tr>
<td>[38]</td>
<td>19 (6.25)</td>
</tr>
<tr>
<td>[39]</td>
<td>14 (4.74)</td>
</tr>
<tr>
<td>[40]</td>
<td>15.8 (6.35)</td>
</tr>
<tr>
<td>[41]</td>
<td>16 (7)</td>
</tr>
<tr>
<td>[42]</td>
<td>19.64 (7.84)</td>
</tr>
<tr>
<td>[43]</td>
<td>16.72 (7.44)</td>
</tr>
<tr>
<td>[44]</td>
<td>19 (7.87)</td>
</tr>
<tr>
<td>[45]</td>
<td>15.4 (5.9)</td>
</tr>
<tr>
<td>[46]</td>
<td>13.7 (7.43)</td>
</tr>
<tr>
<td>[47]</td>
<td>13.7 (7.7)</td>
</tr>
<tr>
<td>[48]</td>
<td>11.7 (5.4)</td>
</tr>
<tr>
<td>[49]</td>
<td>15 (5.63)</td>
</tr>
<tr>
<td>[50]</td>
<td>13.67 (2.76)</td>
</tr>
<tr>
<td>[51]</td>
<td>14.05 (6.19)</td>
</tr>
<tr>
<td>[52]</td>
<td>14.61 (0.84)</td>
</tr>
<tr>
<td>[53]</td>
<td>13.8 (5.9)</td>
</tr>
<tr>
<td>[54]</td>
<td>14.1 (5.7)</td>
</tr>
<tr>
<td>[55]</td>
<td>16.5 (1)</td>
</tr>
<tr>
<td>[56]</td>
<td>14.55 (5.44)</td>
</tr>
<tr>
<td>[57]</td>
<td>19.1 (7.03)</td>
</tr>
<tr>
<td>[58]</td>
<td>18.6 (6.04)</td>
</tr>
<tr>
<td>[59]</td>
<td>16.8 (7.6)</td>
</tr>
<tr>
<td>[60]</td>
<td>15.1 (5.8)</td>
</tr>
</tbody>
</table>
number of studies, and number of participating individuals) are shown in Table 2. Overall SD reported from included studies ranged from 0.83-9.7. A pooled variance was calculated from the scores ($\sigma^2 = 31.26$), leading to a pooled standard deviation of 5.59.

To compare the weighted mean AQ scores between males and females in studies that reported this information, a continuous random-effects model was used to find standard difference in means, SMD. There was a significant difference in scores between males and females: Hedges’ $g = 0.40, P <0.001, z = 3.36$. This holds true even if simple unweighted means are compared, though individual mean values are slightly reduced (Figure 3). A suggestion of bimodality was observed for males and females. However, previous observations of AQ scores indicate that there is a normal distribution within the population; likely

<table>
<thead>
<tr>
<th>Study</th>
<th>Overall SD</th>
<th>Study</th>
<th>Overall SD</th>
<th>Study</th>
<th>Overall SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>[61]</td>
<td>14.05 (5.8)</td>
<td>[62]</td>
<td>16.4 (6.11)</td>
<td>[63]</td>
<td>17.88 (8.21)</td>
</tr>
<tr>
<td>[64]</td>
<td>13.72 (5.67)</td>
<td>[65]</td>
<td>19 (7.1)</td>
<td>[66]</td>
<td>13.6 (7.86)</td>
</tr>
<tr>
<td>[67]</td>
<td>16.69 (5.69)</td>
<td>[68]</td>
<td>14.86 (6.22)</td>
<td>[69]</td>
<td>16.3 (7.7)</td>
</tr>
<tr>
<td>[70]</td>
<td>15.6 (6.9)</td>
<td>[71]</td>
<td>12.11 (5.03)</td>
<td>[72]</td>
<td>16 (5.1)</td>
</tr>
<tr>
<td>[73]</td>
<td>13 (5)</td>
<td>[74]</td>
<td>18.1 (7.5)</td>
<td>[75]</td>
<td>15.6 (6)</td>
</tr>
<tr>
<td>[76]</td>
<td>14.6 (5)</td>
<td>[77]</td>
<td>13.09 (5.79)</td>
<td>[78]</td>
<td>20 (9.7)</td>
</tr>
<tr>
<td>[79]</td>
<td>13.03 (5.85)</td>
<td>[80]</td>
<td>17 (7)</td>
<td>[81]</td>
<td>17.41 (6.89)</td>
</tr>
<tr>
<td>[82]</td>
<td>13.2 (6.1)</td>
<td>[83]</td>
<td>18.26 (4.49)</td>
<td>[84]</td>
<td>12.52 (5.41)</td>
</tr>
<tr>
<td>[85]</td>
<td>12.26 (5.45)</td>
<td>[86]</td>
<td>13.8 (5.7)</td>
<td>[87]</td>
<td>19.57 (7.48)</td>
</tr>
<tr>
<td>[88]</td>
<td>15.65 (1.59)</td>
<td>[89]</td>
<td>13.85 (1.26)</td>
<td>[90]</td>
<td>15 (6)</td>
</tr>
<tr>
<td>[91]</td>
<td>17.2 (5.2)</td>
<td>[92]</td>
<td>15.24 (6.37)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[93]</td>
<td>11.43 (4.2)</td>
<td>[94]</td>
<td>14.2 (4.5)</td>
<td>[95]</td>
<td>12.58 (4.36)</td>
</tr>
<tr>
<td>[96]</td>
<td>17.4 (4.36)</td>
<td>[97]</td>
<td>7 to 20</td>
<td>[98]</td>
<td>12 (5)</td>
</tr>
<tr>
<td>[99]</td>
<td>12 (4)</td>
<td>[100]</td>
<td>13 (5)</td>
<td>[101]</td>
<td>18 (5)</td>
</tr>
</tbody>
</table>

**Table 1 Articles selected for review (Continued)**

Studies ordered chronologically; blank cells indicate data not reported (or for patients, AQ not administered to individuals with an autism spectrum condition). Overall AQ mean refers to combined male and female score.
this observation stems from the comparatively small number of data points used in this calculation (10 studies per group) or from the internal differences in study recruitment paradigms.

After initial selection criteria were applied, \( N = 9; \{35,40,48,52,59,62,70,83,91\} \) studies were identified that excluded any individuals who might have the BAP from the typical group [17]. Table 2 presents the descriptive statistics for this set.

Quantitative characterization of the Autism-Spectrum Quotient in a clinical sample

The 78 included studies were also examined for the presence of a matched clinical sample of individuals with a
formal diagnosis of ASC. Of these, 43 studies contained data from 1,963 individuals with ASC (Table 1). Descriptive statistics for matched clinical cases are shown in Table 2. Overall SD reported from included studies ranged from 4.6 to 10.09. A pooled variance was calculated from the scores ($\sigma^2 = 39.27$), leading to a pooled standard deviation of 6.27.

To compare the weighted mean AQ scores between clinical and nonclinical groups, a continuous random-effects model was used to find SMD. There was a significant difference in scores between these groups: Hedges’ $g = 2.86$, $P < 0.0001$, $z = 26.42$, confirming that AQ scores are elevated in individuals with ASC. Contrasting with the findings reported for nonclinical controls, the SMD for males and females with ASC only reached a value of 0.33, which, while significant, indicates that males and females with ASC do not effectively differ in autistic traits as measured by the AQ; in fact, if anything, in this sample, the trend is reversed so that females self-report higher levels of traits.

Table 2 Descriptive statistics for selected articles

<table>
<thead>
<tr>
<th></th>
<th>Nonclinical Samplea</th>
<th>Nonclinical Sample - BAP excludedb</th>
<th>Matched ASC Casesc</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Malesd</td>
<td>Femalese</td>
<td>Overallf</td>
</tr>
<tr>
<td>Mean AQ</td>
<td>17.89</td>
<td>14.88</td>
<td>16.94</td>
</tr>
<tr>
<td>Range</td>
<td>11.4 to 19.0</td>
<td>10.4 to 17.4</td>
<td>11.6 to 20.0</td>
</tr>
<tr>
<td>SD Range</td>
<td>4.2 to 7.9</td>
<td>4.2 to 8.0</td>
<td>0.8 to 9.7</td>
</tr>
<tr>
<td>CI</td>
<td>16.7 to 19.1</td>
<td>13.3 to 16.5</td>
<td>16.4 to 17.4</td>
</tr>
<tr>
<td>N (studies)</td>
<td>10</td>
<td>10</td>
<td>72</td>
</tr>
<tr>
<td>N (participants)</td>
<td>872</td>
<td>1378</td>
<td>4931</td>
</tr>
</tbody>
</table>

Nonclinical Samplea describes reports from nonclinical samples. Nonclinical Sample – BAP excludedb describes a subset of the previous sample, where it was specified that individuals from the BAP had been excluded. Matched ASC Casesc describes reports from available matched autism spectrum cases. Statistics calculated for all available reported data (see also Table 1). Malesd and Femalese refer to available data reported by sex, while Overallf refers to samples from studies reporting undifferentiated or combined male and female score. AQ: Autism-Spectrum Quotient; ASC: Autism spectrum conditions; BAP: broader autism phenotype; CI: confidence interval; N: number; SD: standard deviation.

Trends in use of the Autism-Spectrum Quotient

In addition to reviewing the reported AQ scores, an effort was made to qualitatively assess AQ usage for included studies. Several trends were noted in administration and reporting of the full-scale AQ for adults. The majority of studies included in this review had recruited via newspaper adverts, employment agencies, email, post, and flyers. In many cases, the participants were also partially drawn from continuously maintained participant databases and research pools. There was also evidence of partial recruitment through hospitals and universities (though, as stated, studies were excluded where recruitment was exclusively within an academic community). In a number of instances, participants were recruited using publicly available online survey tools such as Amazon Mechanical Turk (M-Turk) and surveymonkey.com. Finally, several large studies were made possible through the use of birth cohorts, including the Raine Cohort (in Western Australia).

Few articles specified the precise inclusion and exclusion criteria for control participants, instead focusing primarily...
on the characterisation of the clinical group. While authors did routinely specify that control participants did not have a history of psychological or neurodevelopmental conditions, articles rarely reported having also excluded participants there was a family history of these. Studies also rarely reported testing the psychometric properties of the AQ or the normality of the distribution of AQ score. For instance, the mean was only reported alongside median in one instance [54], and there was also only one instance of a test for normality (Kolmogorov-Smirnov test [83]). There were occasional reports of other psychometric properties of the AQ, such as Cronbach’s alpha, establishing good internal consistency of the AQ.

Conclusions
This is the first systematic review of the AQ, with several findings emerging. First, the mean AQ score in a typical sample drawn from a nonclinical population is approximately 17 (CI 16.4 to 17.4) (for those explicitly excluding BAP, the mean is approximately 15 (CI 13.0-17.1)), with a narrow confidence interval of one to two points. In addition, the mean AQ score in individuals with ASC is approximately 35, nearly 20 points above that found in the general population. Second, control males and females have significantly different average AQ scores, with males scoring higher, confirming earlier reports. Third, from 2001 onwards, there is considerable fluctuation in reported mean AQ scores, but scores have not appreciably drifted in one direction or another within the general population.

Several rationales were employed by researchers for using the AQ. Many of the studies administered the AQ not as a central variable correlated with the outcome measure, but as an accessory measure for characterizing the population or defining the experimental groups. Further, a number of articles used the AQ as a proxy for diagnosis, using the cut-off scores of either 32 or 26 to exclude individuals either from the clinical or from the nonclinical control group. (These articles were not included in the final analysis). However, caution is recommended when using the AQ in this way, as the AQ was designed to be a descriptive, rather than a diagnostic, measure of autistic traits. While, perhaps due to it being freely available, easy to administer, and widely precedent in the literature, the AQ is used as a screening instrument (such as for patients referred to a diagnostic clinic for a detailed assessment for ASC [93]), it has been argued that the AQ does not have the sensitivity and specificity for population screening with an eye to diagnosis [94-96]. This follows logically from the fact that the AQ is a brief self-report, reliant upon the individuals’ own self-awareness, and from the self-imposed limitations of age (16+) and IQ (85+). As discussed in the original publication, the AQ was developed from a theoretical understanding of autism, and therefore has not necessarily undergone the rigorous psychometric evaluation procedure that diagnostic screening tools must pass for inclusion in clinical practice. A more conservative use for the AQ is to segment the population into bands of autism phenotypes (broad, medium and narrow) as in the method of Wheelwright and colleagues [17], or as a descriptive quantitative measure of autistic traits.

Strengths of the current review include the exhaustive search criteria, especially the citation search for relevant papers, followed by the rigorous selection process. In addition, the total number of individuals (N = 8,897 clinical cases and nonclinical controls) examined by this review lends weight to the findings. The study was limited by a risk of bias, at the outcome level, the selection level and at the level of the review, though an effort was made to mitigate possible disproportionate effect of means from studies of varying samples through weighting by group size. Limitations also exist in the review procedure, in that each study included in the review was not judged for methodological rigor, rather a holistic evaluation was made of study methodology in an effort to report trends in items such as recruitment strategies, participant inclusion, and AQ data psychometric properties. Second, the number of participants from each study was relatively small (minimum N was set at 10); this is balanced by the large overall sample size derived from summing all studies together. More broadly, an ideal investigation of AQ score distribution would evaluate the raw data from each of the included studies in order to also measure data spread and subscale scores. However, this was not feasible for the current study. Finally, not every included article verified that the control group did not have ASC. Therefore there may have been incomplete information on how representative the demographic distribution of the nonclinical sample that make up this analysis may be.

We recommend that future researchers should think carefully when planning a recruitment strategy, both for nonclinical and clinical participants in order to be able to clearly define participants in each group. Furthermore, the field would greatly benefit if researchers better described the control participants, stating the method of recruitment in the methods. While healthy, typically-developing participants are often taken for granted, the considerable variability found in this review indicates that the method of recruitment of a ‘true’ representative sample - either of the general population or a specific patient group - may significantly impact results. This could have implications when examining group differences on dependent variables if the groups have not been carefully defined, potentially leading to attenuation of real groups differences.

We hope the current review holds value in the light of the considerable range of research types under which the AQ has been used. This dimensional approach to
quantifying autistic traits has been found to correlate with a range of biological measures, including instances of brain activity [97], brain structure [98], social perception using gaze-tracking [42], prenatal testosterone [99], candidate genes and epigenetics [100]; clinical screening [44] and autism genetic risk [17]. Thus, although it is a self-report instrument, it correlates with a large number of more objective measures, suggesting that autistic traits are measurable aspect of personality, independent of the Big 5 [101].

Future research might consider a similar investigation of other versions of the AQ. Aside from the AQ-Adolescent and -Child, widely-used cross-cultural and foreign-language versions of the AQ exist, including translations into Chinese [102], Dutch [103], French [104], Italian [105], Japanese [106], Persian [107] and Polish [108], among others. On the whole, the results from studies that utilize these versions demonstrate analogous findings to those found using English-language versions of the AQ; however, validation by systematic review has not been done. In addition, a future study might attempt to undertake a whole population survey of autistic traits using the AQ, with more detailed information about the respondents collected in order to make stronger claims about generalizability. The underlying structure of taxa leading to AQ score distribution could be assessed using a number of modelling solutions, including latent class, taxometric, or factor mixture modelling. Perhaps, using these techniques in a population sample of individuals along the spectrum might help elucidate the apparent gap between clinical and nonclinical scores, despite the apparent continuity of autistic traits.

Summary

The AQ continues to be a useful brief assessment instrument for measuring autistic traits in adults of normal intelligence. By determining the distribution of the AQ in the nonclinical population, the AQ can now be used more definitively to assess the extent to which other specialist populations exhibit autistic traits.

Abbreviations


Competing interests

The authors declare that they have no competing interests.

Authors’ contributions

ER developed the study; performed the data search, assessment, extraction, and analysis; and wrote the manuscript under the supervision of CA, PS, HR, and SBC. PW provided advice on statistical analysis approaches. BA provided additional supervisory support. All authors have read and approved the final version of this manuscript.

Acknowledgements

This research was supported by the Medical Research Council UK; the Wellcome Trust, and the National Institute for Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care East of England at Cambridgeshire and Peterborough NHS Foundation Trust. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

Author details

1Cambridge Intellectual and Developmental Disabilities Research Group, Department of Psychiatry, University of Cambridge, Douglas House, 188 Trumpington Road, CB2 8AH Cambridge, UK. 2Autism Research Centre, Department of Psychiatry, University of Cambridge, Douglas House, 188 Trumpington Road, Cambridge CB2 8AH, UK. 3MRC Cognition and Brain Sciences Unit, 15 Chaucer Road, Cambridge CB2 7EF, UK. 4Psychology Department, University of Edinburgh, 3 Charles Street, Edinburgh EH8 9AD, UK. 5NIHR CLAHRC for the East of England, Douglas House, 188 Trumpington Road, Cambridge CB2 8AH, England, UK. 6Cambridgeshire and Peterborough NHS Foundation Trust, Peterborough CB21 1EF, UK. 7CLASS Clinic, Cambridgeshire and Peterborough NHS Foundation Trust, Peterborough CB21 1EF, UK.

Received: 18 August 2014 Accepted: 17 December 2014
Published: 14 January 2015

References


Cite this article as: Ruzich et al.: Measuring autistic traits in the general population: a systematic review of the Autism-Spectrum Quotient (AQ) in a nonclinical population sample of 6,900 typical adult males and females. Molecular Autism 2015. 6:2.