ORIGINAL PAPER

Factor Structure, Reliability and Criterion Validity of the Autism-Spectrum Quotient (AQ): A Study in Dutch Population and Patient Groups

**Rosa A. Hoekstra · Meike Bartels · Danielle C. Cath ·
Dorret I. Boomsma**

Abstract The factor structure of the Dutch translation of the Autism-Spectrum Quotient (AQ; a continuous, quantitative measure of autistic traits) was evaluated with confirmatory factor analyses in a large general population and student sample. The criterion validity of the AQ was examined in three matched patient groups (autism spectrum conditions (ASC), social anxiety disorder, and obsessive–compulsive disorder). A two factor model, consisting of a "Social interaction" factor and "Attention to detail" factor could be identified. The internal consistency and test–retest reliability of the AQ were satisfactory. High total AQ and factor scores were specific to ASC patients. Men scored higher than women and science students higher than non-science students. The Dutch translation of the AQ is a reliable instrument to assess autism spectrum conditions.

Keywords Autism · Factor analysis · Validity · Reliability · Autism phenotype

R. A. Hoekstra · M. Bartels · D. I. Boomsma
Department of Biological Psychology, Vrije Universiteit,
Amsterdam, The Netherlands

R. A. Hoekstra (✉)
Autism Research Centre, Department of Developmental
Psychiatry, University of Cambridge, Douglas House,
18b Trumpington Road, Cambridge CB2 8AH, England, UK
e-mail: rah58@medschl.cam.ac.uk

D. C. Cath
Department of Psychiatry, GGZ Buitenamstel, Amsterdam,
The Netherlands

Introduction

Pervasive developmental disorders, of which the most common are autistic disorder, Asperger Syndrome (AS) and pervasive developmental disorder not otherwise specified (PDD-NOS), are characterized by a triad of impairment: difficulties in reciprocal social interaction, communication, and the presence of stereotyped behavior, interests, and activities (American Psychiatric Association 2000). Together, these conditions are referred to as autism spectrum conditions (ASCs). This term reflects the assumption that (high functioning) autism and AS lie on a continuum reflecting severity of social communication disability, from classical autism at the most severe end of the spectrum, decreasing via high functioning autism (HFA), AS, and PPD-NOS into normal behavior. Twin and family studies have shown that genetic factors play a major role in the risk for ASC (Rutter 2000). A conservative estimate of the concordance rates of autism in monozygotic twins is 60%. Concordance in dizygotic twins, however, is only 0–5% (Folstein and Rutter 1977; Bailey et al. 1995). Moreover, if one MZ twin has autism, not only is the likelihood of an autism diagnosis in the co-twin increased, but also the risk of other neurodevelopmental difficulties affecting language and social interaction (Le Couteur et al. 1996; Bailey et al. 1995). This notion led to the idea that the same genetic variants affecting the risk for autism may lead to a broader phenotype of autistic traits, as reflected in an increased rate of social deficits, impairments in communication and language, a preference for routines and difficulty with change in non-autistic relatives of autistic individuals (Bolton et al. 1994; Bailey et al. 1998). Rather than a distinct disorder, it is now thought that the autism spectrum conditions as defined in the DSM-IV represent the upper extreme of one or more quantitative traits, and these traits may be

continuously distributed in the population (Constantino and Todd 2003; Piven et al. 1997; Spiker et al. 2002). Studies using quantitative measurements of autistic traits found elevated scores in relatives of autistic and pervasive developmental disorder patients (Constantino et al. 2006; Bishop et al. 2004) and high scores in children whose parents showed high (but sub-diagnosis) endorsement on autistic traits (Constantino and Todd 2005).

The quantitative approach to autistic traits has led to the development of the Autism-Spectrum Quotient (AQ; Baron-Cohen et al. 2001). This self-administered questionnaire was developed to quantify autistic traits in individuals with normal intelligence. The AQ consists of 50 items, assessing personal preferences and habits. Subjects rate to what extent they agree or disagree with the statements on a 4-point Likert scale, with answer categories "definitely agree"; "slightly agree"; "slightly disagree" and "definitely disagree". For approximately half the items an "agree" response is in line with autistic traits (e.g. item 23: "I notice patterns in things all the time"); for the other half a "disagree" response is indicative of an autistic trait (e.g. item 11: "I find social situations easy"). All the item scores are summed; a high AQ score indicates a high autistic load, close to the autistic end of the autism spectrum. In the original version of the AQ (Baron-Cohen et al. 2001), the 50 items were divided into five theoretically derived subscales of 10 items each: Social skill; Communication; Imagination; Attention to detail; and Attention switching.

Both the British (Baron-Cohen et al. 2001) and the Japanese version (Kurita et al. 2005) of the AQ reported good test-retest reliability and moderate internal consistency (Cronbach's α varying from .63 to .78). However, some aspects of the AQ merit further study. Firstly, the five domains of the AQ have been derived on a theoretical basis and have undergone little empirical testing. Austin (2005) conducted an exploratory factor analysis of the AQ in a group of 201 undergraduate students and found evidence for three, rather than five factors, with a focus on Social skill, Details/patterns, and Communication/mind reading. In a principal component analysis using data on the child-version of the AQ, Auyeung et al. (2008) found support for 4 rather than 5 factors. Confirmatory factor analyses can further extend our knowledge on the psychometric qualities of the AQ subscales and empirically test the goodness of fit of the 5-domain model. The first aim of the current study was to examine the model fit of the 5 domain structure proposed by Baron-Cohen et al. (2001) in a large student sample and a general population sample and compare the fit with alternative models. Some of the 5 domains assess traits that may show considerable overlap (e.g. satisfactory social skills are needed for good communication). Therefore we expected some of the domains to be substantially

correlated and predicted that a model in which the highly correlated factors are clustered may be preferred.

Secondly, more research needs to be done on the criterion validity of the AQ. Studies in England (Baron-Cohen et al. 2001), and Japan (Wakabayashi et al. 2006) found significantly higher AQ scores in subjects with an AS or HFA diagnosis, compared to scores in a student sample and a general population sample. A preliminary study (Woodbury-Smith et al. 2005) reported satisfying ability of the AQ to distinguish between subjects with and without an AS/HFA diagnosis, in a group of 100 referrals to a diagnostic clinic for adults suspected of having AS or HFA. However, a Dutch study comparing AQ scores in mild ASC patients with scores in patients referred to a general outpatient clinic found little differences between the two groups (Ketelaars et al. 2007). The subjects in the latter study were not matched on sex, education and age. It thus remains unclear whether high AQ scores are specific to ASCs, or may be common to psychiatric disorders in general. Apart from a general population and a student sample, the current paper includes a small sample of three different patient groups: one group with AS/HFA and PPD-NOS patients, a group of patients diagnosed with obsessive-compulsive disorder (OCD), and a group with social anxiety disorder (SAD). Because SAD patients demonstrate problems with social interaction, and OCD patients show symptoms of repetitive behavior, we predicted that the AQ scores in these patient groups will be higher than the general population mean. Extremely high AQ scores were expected to be specific to ASC patients.

Lastly, this paper is the first to examine the psychometric properties of the AQ in a Dutch population. The original version of the AQ was translated to Dutch using the backward translation procedure. The characteristics of the Dutch AQ, including test–retest reliability and internal consistency, were studied in a large sample of students and subjects from the general population. In line with the empathizing–systemizing theory of autism (Baron-Cohen 2002; Baron-Cohen et al. 2005) and findings from previous studies (Baron-Cohen et al. 2001; Hoekstra et al. 2007; Wakabayashi et al. 2006) we expected: (a) AQ scores to be continuously distributed in the general population; (b) significantly higher mean AQ scores in men compared to women; and (c) higher AQ scores in science students than in students in the field of humanities and social sciences.

Methods

Participants

This study encompassed four different samples. The first group consisted of 961 students from the VU university in

Amsterdam ($n = 809$) and the University of Twente in Enschede ($n = 152$). During the break of one of their classes, the students were asked to complete the AQ. Students were recruited from the fields of humanities (history and law; $n = 128$), the social sciences (psychology, education, and communication science; $n = 594$) and natural and technical sciences (including mathematics, physics and information sciences; $n = 239$). Participation rates varied from 65 to 100%. The mean age of the students was 21.19 years (SD 3.69).

The general population sample ($n = 302$) consisted of parents of twins who visited an information day for parents of multiples. They were asked to either fill out the AQ immediately or to return the questionnaire to our research group by mail. The participation rate for this data collection was 62%; mean age of the participants was 35.68 years ($SD = 6.33$). The student and general population groups were not matched on age and IQ. However, the two groups were included for separate research purposes. The parent group served as a normative sample, whereas the student sample was included to address differences in AQ scores in different fields of study.

To obtain data on test–retest reliability, a group of 18-year old twins and their brothers and sisters filled out the AQ. These twin families participated in an ongoing study on cognitive development in late puberty and completing the AQ was part of the test protocol (see Hoekstra et al. 2007 for more information on this study). The first 117 participants of the study were re-contacted 1–6 months later (the mean time interval between the first and second assessment was 3.9 months), and were asked to fill out the AQ for a second time. Seventy-five participants returned the questionnaire for the second time (64%). AQ scores of the responders in the retest did not significantly differ from the subjects who did not respond in the retest ($F(1, 115) = .066$, $p = .797$).

The last group consisted of three subgroups of psychiatric patients, who were all adult outpatients recruited from the anxiety outpatient services of GGZ Buitenamstel in Amsterdam (see also Cath et al. 2008). All subjects were administered the Structured Clinical Interview on DSM-IV diagnoses (SCID-I; First 1996) to establish in-, and exclusion criteria. Subjects suffering from co-morbid depression, psychosis, substance abuse, mental deficiency or inability to read or speak Dutch were excluded. To exclude any risk of cognitive deficit and/ or below average intelligence, only patients who had successfully completed an educational degree were included in the study, and patient groups were matched on age (range 19–57 years), sex (10 males; 2 females in all groups) and educational level. The SCID-I does not contain a section on autism disorders, and at the time of data collection no validated Dutch version of either the Autism Diagnostic Observation

Schedule-Generic (ADOS-G; Lord et al. 2000) or the ADI-R (Lord et al. 1994) was available for adult subjects. Therefore all subjects were assessed on presence of ASCs according to DSM-IV criteria by two independent experienced clinicians and with the aid of a structured retrospective interview taken from one of the parents of the patients on early infant development in all domains of the spectrum of autistic conditions. The structured interview encompassed the following topics: age at onset of problem behavior, contact and communication skills, stereotyped behavior, development of speech and language, motor and sensory development, particular interests and skills, ability to display imagination, resistance against change and unexpected events, and impulse control. Only subjects who had independently been diagnosed with an ASC by the two clinicians were included in the study; diagnoses were made independent of the AQ responses. Subjects meeting the inclusion criteria completed the AQ at home, after they had given written informed consent, and returned the questionnaire during their next visit to the outpatient service. The patient groups encompassed (a) 12 patients with an autism spectrum condition ($n = 2$ HFA; $n = 4$ AS; $n = 6$ PDD-NOS); (b) 12 patients with a "pure" obsessive-compulsive disorder, 3) 12 patients with a "pure" generalized social anxiety disorder.

The Dutch Autism-Spectrum Quotient

The AQ was translated after permission from prof. Simon Baron-Cohen (SBC). The translation into Dutch was conducted by an official translator. Subsequently, a second translator translated the Dutch version back into English. After comparing the outcome of the retranslated version to the original text, and discussing discrepancies in the retranslation with SBC, a final version was established (the Dutch version of the questionnaire is obtainable from the first author upon request). Total AQ and domain scores were based on the original 4-point Likert scale scores (1 = "definitely agree" up to 4 = "definitely disagree"). For the items in which an "agree" response is characteristic for autism, the scoring was reversed ("definitely agree" scored 4 points; "slightly agree" 3 points, etc.; This was the case in item 2, 4, 5, 6, 7, 9, 12, 13, 16, 18, 19, 20, 21, 22, 23, 26, 33, 35, 39, 41, 42, 43, 45, 46). All item scores were summed, resulting in a minimum total AQ score of 50 (no autistic traits) and a maximum score of 200 (full endorsement on all autistic items). Note that in most reports using the British version of the AQ (Baron-Cohen et al. 2001) the answer categories have been dichotomized into "agree"/disagree" scores. In these studies, all item responses in line with the autism phenotype scored one point, resulting in a maximum total AQ score of 50.

Missing Answers and Outliers

If more than five items were left blank (10% of the total number of items), the AQ was considered incomplete and the data were discarded in analyses ($n = 2$ in the student group; $n = 7$ in the general population group; none in the twin family or patient groups). Two social sciences students and one subject from the general population group obtained a score >160. Since these subjects completed the AQ anonymously, it could not be verified whether this reflected a true score. These outliers (deviation >4 SD's of the mean) were therefore left out of the analyses.

Statistical Analyses

To examine the factor structure of the AQ, models were fitted on the student group data using confirmatory factor analyses in LISREL. The items of the AQ were measured on an ordinal scale, therefore the diagonally weighted least square procedure was used. First, a five-factor model, based on the 5 domains put forward by Baron-Cohen et al. (2001) was fitted. The factors were allowed to correlate since it is reasonable to assume that different aspects of autistic traits will be related. The fit of this model was compared with the fit of alternative models. Secondly, the model fitting procedure was repeated on the data from the general population sample. Based on the model fitting results in these two samples, the most parsimonious best fitting model was chosen.

To evaluate model fit, several model fit statistics were inspected. The χ^2 test statistic and the Standardized Root Mean Square Residual (SRMR) were included as a measure of overall goodness of fit. Low χ^2 values compared with the number of degrees of freedom indicate that the model fits the data well. The SRMR is a standardized overall badness-of-fit measure and is based on the fitted residuals. An SRMR value of zero indicates no residuals and thus a perfect fit. As a rule of thumb, an SRMR of less than .05 indicates a good fit, whereas values smaller than .10 can be interpreted as acceptable fit (Schermelleh-Engel et al. 2003). To enable a comparison between different models, the Goodness of Fit Index (GFI), Parsimony Goodness of Fit Index (PGFI); and the Expected Cross Validation Index (ECVI) were included as well. The GFI measures the relative amount of the variances and covariances in the observed matrix that is predicted by the model. The GFI ranges between zero and one with higher values indicating a better fit. The PGFI is a modification of the GFI and takes model complexity into account. High PGFI values indicate good fit in a relatively parsimonious (and thus less complex) model. If choosing between alternative models, the highest PGFI indicates the superior model. The ECVI evaluates how well the model fitted in the sample under study would generalize to other samples and is thus a measure of cross validation. The model with the smallest ECVI indicates the model with the best fit.

Group differences in total AQ score and factor scores, and the validity and internal consistency of the scale were analyzed using SPSS. Group differences were tested using mixed model multivariate analysis of variance (MANOVA). The mixed linear model is an expansion of the general linear model and permits non-constant variability in the data. This way, it was possible to test for group differences in variables with unequal variances and unequal group sizes. In the student sample, overall sex differences and differences between students from different fields were explored. Effects of sex and age were studied in the general population sample. Furthermore, AQ score differences between patient groups and the general population sample were tested. Internal consistency was assessed in the student and general population using Cronbach's alpha.

The test–retest reliability of AQ scores was assessed using AQ-scores collected in 18-year-old twins and their brothers and sisters ($n = 75$). These participants are genetically related. In order to use all available data and to satisfy the independent observations assumption for statistical testing, the test–retest analysis was performed using structural equation modeling in the computer program Mx (Neale et al. 2006).

Results

Factor Analysis

A correlated 5-factor structure, based on the domains suggested by Baron-Cohen et al. (2001) was fitted to the data of the student group using confirmatory factor analysis. The results of this analysis showed that the domains "Social skill"; "Communication"; "Attention switching"; and "Imagination" were highly correlated (correlations varying from $r = .53$ to $r = .84$). These correlations indicate a considerable overlap between the domains, and suggest that a model allowing these domains to cluster together may fit the data better. To examine this, a hierarchical model, encompassing 1 higher order factor, existing of 4 lower order domains ("Social skill"; "Communication"; "Attention switching"; and "Imagination") and 1 separate factor "Attention to detail" was fitted to the data. Lastly, a 1-factor model incorporating all 50 items was fitted, to evaluate whether the AQ really is multifactorial or rather measures 1 construct. The fit statistics of the three models are given in the top of Table 1. The SRMR was <.10 in all three models indicating adequate fit. The χ^2

Table 1 Summary of the fit statistics of the several factor model structures

	5 factor model	Higher order factor model	1 factor model
Student sample ($n = 961$)			
χ^2	11793.341 ($p = 0.0$)	**11755.716 ($p = 0.0$)**	14574.788 ($p = 0.0$)
df	1165	**1170**	1175
SRMR	0.0904	**.0907**	0.0939
GFI	0.732	**.730**	0.694
PGFI	0.668	**.670**	0.640
ECVI	1.443	**1.438**	1.432
General population ($n = 302$)			
χ^2	5365.023 ($p = 0.0$)	**5414.112 ($p = 0.0$)**	5855.664 ($p = 0.0$)
df	1165	**1170**	1175
SRMR	0.107	**.107**	0.112
GFI	0.534	**.525**	0.435
PGFI	0.488	**.482**	0.401
ECVI	4.601	**4.585**	4.568

Best fitting model is shown bold faced

value of the 1 factor model was considerably larger than the value of the two other models, suggesting that the 1 factor model showed relatively poor fit. The GFI and PGFI suggested the same; both values were low for the 1 factor solution. Based on these fit indices, the 1 factor solution was rejected. When comparing the fit indices for the two remaining models, the hierarchical factor model showed the lowest value for χ^2 and ECVI and the highest PGFI, suggesting that this model fitted the data best. The GFI was marginally higher in the 5-domain model, but the difference with the hierarchical model was very small. All in all, the fit indices suggest that the hierarchical model showed the best fit.

Subsequently, the fit of the three models was tested in the general population sample (see bottom half of Table 1). Similar to the analyses in the student sample, the 1 factor model showed the poorest fit and was rejected. The SRMR values for the two remaining models were the same and indicated near-acceptable fit. The ECVI value indicated a slightly better fit for the hierarchical model, while the GFI and the PGFI were the highest in the 5-domain model. Altogether the fit indices for the 5-domain model and the hierarchical model were very similar in the general population. Since the hierarchical model is more parsimonious, this model was chosen as the best fitting model.

The Hierarchical Factor Model

According to the analyses in two independent samples, the model that best described the data of the AQ consisted of 1 higher order factor, encompassing the 4 lower order domains Social skill; Communication; Attention switching; and Imagination, and 1 separate factor Attention to detail (see also Fig. 1). The items loading on the first higher order

factor mainly focus on social situations, difficulties in communication with others, and empathic abilities. This factor was called "Social interaction". The second factor is the domain "Attention to detail" and mainly consists of items assessing interests in patterns and details. For both factors, similar to the total AQ score, a high factor sum score implies a high autistic load. Since the same best fitting model was identified in both samples, the best fitting model was subsequently fitted on the combined student and general population sample in order to obtain 1 estimate for the domain and factor loadings (see Table 2 and Fig. 1) and to obtain the correlation between the factors. Apart from item 18 and 30, all domain loadings were positive. The correlation between the 2 factors was estimated at $r = .19$ ($p < .001$).

Internal Consistency and Test–Retest Reliability

The internal consistency was assessed in the student and general population sample. The internal consistency of the total AQ score was satisfactory in both samples (Cronbach's $\alpha = .81$ (student sample) and $\alpha = .71$ (general population sample)). The internal consistency of the higher order factor Social interaction (40 items) was estimated at $\alpha = .84$ and $\alpha = .77$. The internal consistencies for the lower order domains and the Attention to detail factor were somewhat lower and were estimated at: $\alpha = .63/.68$ (Attention to detail); $\alpha = .76/.68$ (Social skill); $\alpha = .63/.62$ (Attention switching); $\alpha = .52/.49$ (Communication); and $\alpha = .63/.52$ (Imagination). The test–retest reliability as assessed in 75 young adults recruited in the twin family study was $r = .78$ for the total AQ score and $r = .79$ for the Social interaction factor. The test–retest reliabilities for the lower order domains and the Attention to detail factor

Fig. 1 Path Diagram of the Best Fitting Model with the Factor Correlation and the Standardized Estimates of the Factor Loadings on the underlying Domains, as estimated in the combined Student and General Population Sample. Att switching = Attention switching; Att detail = Attention to detail
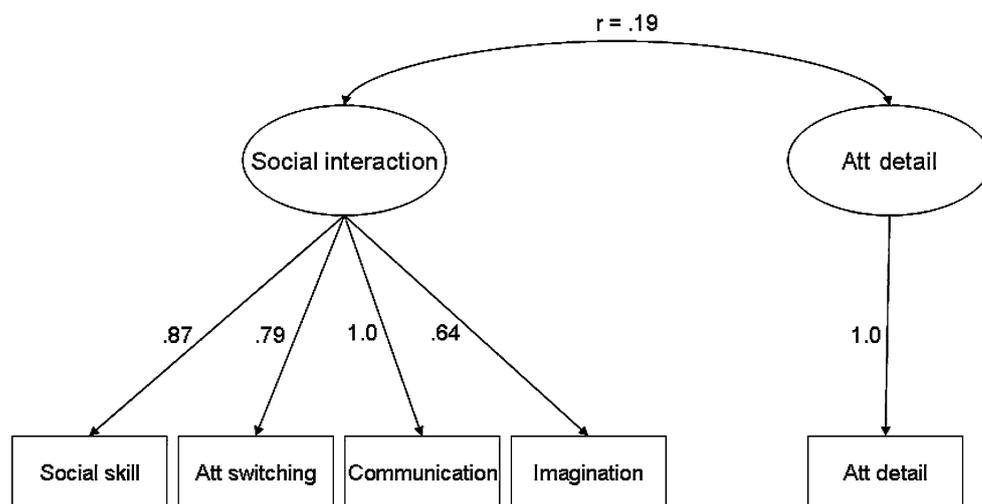
were respectively: $r = .71$(Attention to detail), $r = .69$ (Social skill), $r = .68$ (Attention switching), $r = .60$ (Communication), and $r = .81$ (Imagination).

Students: Effects of Sex and Field of Study

Table 3 shows the mean total AQ and factor scores in the student sample, separated by field of study. Effects of sex and field of study on the factor scores Social interaction and Attention to detail were tested using a mixed model MANOVA. A significant main effect of field of study ($F = 41.41$, $p < .001$) and a significant sex × field of study interaction effect ($F = 5.92$, $p = .003$) was found. Within the different fields of study, the sex effect was not significant ($F = .40$, $p = .53$). Science students scored significantly higher than students engaged in a humanities or social sciences degree ($t = 8.64$, $p < .001$), whereas the social sciences students scored significantly lower than the humanities and science students ($t = 4.48$, $p < .001$). The effect of field of study was significant both in the Social interaction factor ($F = 28.22$, $p < .001$) and in the Attention to detail factor ($F = 5.34$, $p = .005$).

General Population: Effect of Sex and Age

Mean total AQ and factor scores in the general population sample are shown in Table 3, separated by sex. A MANOVA of Social interaction and Attention to detail by sex and age showed a significant sex effect ($F = 5.55$, $p = .02$), with men scoring significantly higher than women. The effect of age was not significant ($F = .46$, $p = .50$). Contrast tests showed that the sex effect was significant in both the total AQ score ($t = 2.01$, $p = .05$) and in the Social interaction factor ($t = 2.46$, $p = .01$), but not in the Attention to detail factor ($t = -.52$, $p = .61$).

Patient Groups vs. General Population; Differences Between Patient Groups

Mean total AQ and factor scores in the different patient groups are shown in Table 3. A mixed model MANOVA of Social interaction and Attention to detail by diagnosis revealed significant differences between the patient and general population samples ($F = 19.17$; $p < .001$). Contrast tests showed that the subjects diagnosed with an ASC ($t = 6.89$, $p < .001$), the subjects diagnosed with OCD ($t = 2.53$, $p = .02$), and the subjects diagnosed with SAD ($t = 2.24$, $p = .04$) all scored significantly higher on the total AQ score than the general population sample. OCD and SAD patients did not differ in total AQ score ($t = -.11$, $p = .91$). The ASC subjects however obtained significantly higher total AQ scores than the OCD ($t = 3.99$, $p < .001$) and SAD patients ($t = 3.99$, $p < .001$). Moreover, the ASC subjects scored the highest on the Social interaction factor. Their scores on the Social interaction factor were significantly higher than the scores of the general population ($t = 6.57$, $p < .001$), the OCD sample ($t = 3.56$, $p = .002$), and the SAD group ($t = 2.88$, $p = .009$). Lastly, the ASC group scored significantly higher on the Attention to detail factor compared to the general population ($t = 2.10$, $p = .04$), the OCD patients ($t = 1.99$, $p = .05$), and the SAD sample ($t = 4.28$, $p < .001$).

The Spectrum of Autistic Traits

To obtain a better insight in the differences in AQ scores within each group, a frequency distribution is tabulated of the total AQ score in the general population and the three patient groups (Table 4). A high total AQ score (>145) was found to be specific to ASC subjects. Neither subjects from

Table 2 Item content and loadings on the 5 domains in the best fitting structure, ordered per higher order factors and lower order domains

Item number		Domain loading
Higher order factor social interaction		
Social skill		
1	I prefer to do things with others rather than on my own	.358
11	I find social situations easy	.718
13*	I would rather go to a library than a party	.356
15	I find myself drawn more strongly to people than to things	.537
22*	I find it hard to make new friends	.643
36	I find it easy to work out what someone is thinking or feeling	.450
44	I enjoy social occasions	.748
45*	I find it difficult to work out people's intentions	.499
47	I enjoy meeting new people	.734
48	I am a good diplomat	.232
Attention switching		
2*	I prefer to do things the same way over and over again	.256
4*	I frequently get strongly absorbed in one thing	.186
10	I can easily keep track of several different people's conversations	.481
16*	I tend to have very strong interests	.172
25	It does not upset me if my daily routine is disturbed	.366
32	I find it easy to do more than one thing at once	.486
34	I enjoy doing things spontaneously	.755
37	If there is an interruption, I can switch back very quickly	.382
43*	I like to plan any activities I participate in carefully	.144
46*	New situations make me anxious	.547
Communication		
7*	Other people frequently tell me that what I have said is impolite	.279
17	I enjoy social chit-chat	.283
18*	When I talk, it isn't always easy for others to get a word in edgeways	−.116
26*	I don't know how to keep a conversation going	.608
27	I find it easy to "read between the lines"	.419
31	I know how to tell if someone listening to me is getting bored	.290
33*	When I talk on the phone, I am not sure when it's my turn to speak	.560
35*	I am often the last to understand the point of a joke	.293
38	I am good at social chit-chat	.463
39*	People tell me that I keep going on and on about the same thing	.209
Imagination		
3	Trying to imagine something, I find it easy to create a picture in my mind	.599
8	Reading a story, I can easily imagine what the characters might look like	.576
14	I find making up stories easy	.441
20*	Reading a story, I find it difficult to work out the characters' intentions	.481
21*	I don't particularly enjoy reading fiction	.241
24	I would rather go to the theatre than a museum	.235
40	When younger, I enjoyed playing games involving pretending with other children	.454
41*	I like to collect information about categories of things	.229
42*	I find it difficult to imagine what it would be like to be someone else	.380
50	I find it easy to play games with children that involve pretending	.602
Factor attention to detail		
5*	I often notice small sounds when others do not	.252
6*	I usually notice car number plates or similar strings of information	.584

Table 2 continued

Item number		Domain loading
9*	I am fascinated by dates	.754
12*	I tend to notice details that others do not	.148
19*	I am fascinated by numbers	.920
23*	I notice patterns in things all the time	.508
28	I usually concentrate more on the whole picture, rather than the small details	.118
29	I am not very good at remembering phone numbers	.120
30	I don't usually notice small changes in a situation, or a person's appearance	−.135
49	I am not very good at remembering people's date of birth	.120

Note: * Designates a reverse-scored item

Table 3 Mean total AQ and factor scores per group, sex, and field of study

Sex	N	Total AQ score (SD)	Social interaction factor (SD)	Attention to detail factor (SD)
Humanities students				
Total	128[a]	104.46 (12.35)	80.71 (11.88)	23.76 (3.99)
♂	39	105.54 (12.94)	81.77 (12.91)	23.77 (3.44)
♀	88	103.99 (12.13)	80.24 (11.43)	23.75 (4.24)
Social sciences students				
Total	594[b]	99.07 (11.19)	75.41 (10.09)	23.66 (4.26)
♂	123	101.32 (12.16)	77.32 (11.17)	24.00 (4.21)
♀	459	98.47 (10.84)	74.91 (9.73)	23.56 (4.28)
Natural and technical sciences students				
Total	239[c]	109.66 (13.37)	85.41 (12.54)	24.26 (4.53)
♂	203	109.41 (13.73)	85.53 (12.80)	23.89 (4.49)
♀	32	111.28 (10.81)	84.69 (10.87)	26.59 (4.08)
Test study	Main effect: $F = 41.41$***	$F = 28.22$***	$F = 5.34$**
Test sex	Interaction effect: $F = 5.92$**		
Test sex	Effect within field of study: $F = .40$ (ns)		
General population				
Total	302[d]	104.20 (11.29)	79.88 (10.68)	24.32 (4.97)
♂	137	105.66 (10.99)	81.52 (10.98)	24.14 (4.78)
♀	160	102.93 (11.50)	78.49 (10.25)	24.44 (5.18)
Test sex	Main effect $F = 5.55$*	$t = 2.46$*	$t = -.52$ (ns)
Test age	Main effect: $F = .46$ (ns)		
ASC	12	142.25 (22.01)	114.83 (19.12)	27.42 (5.29)
SAD	12	114.17 (16.64)	95.50 (15.01)	18.67 (5.16)
OCD	12	114.83 (12.55)	91.50 (14.07)	23.33 (6.49)
Test diagnosis	Main effect $F = 19.17$***		
ASC vs SAD	$t = 3.99$***	$t = 2.88$**	$t = 4.28$***
ASC vs OCD	$t = 3.99$***	$t = 3.56$**	$t = 1.99$*
ASC vs general population	$t = 6.89$***	$t = 6.57$***	$t = 2.10$*

Note: ASC = autism spectrum condition; SAD = social anxiety disorder; OCD = obsessive-compulsive disorder. [a] 1 subject sex unknown; [b] 12 subjects sex unknown; [c] 4 subjects sex unknown; [d] 5 subjects sex unknown

In general population sample: outlier score = 167 left out of the analysis; in student sample: 2 outliers score = 165 and 161 left out of the analysis

* $p < .05$; ** $p < .01$; *** $p < .001$; ns = non significant

Table 4 Frequency distribution of the total AQ score in the general population sample and three patient groups

Total AQ score	General population	ASC	SAD	OCD
50–75	100	100	100	100
76–80	99.7	100	100	100
81–85	97.7	100	100	100
86–90	95.0	100	91.7	100
91–95	89.4	100	91.7	100
96–100	76.8	100	83.3	100
101–105	61.9	100	83.3	83.3
106–110	45.4	100	75	75
111–115	30.1	*100 PDD-NOS*	50	58.6
116–120	15.9	*91.7 PDD-NOS*	50	58.3
121–125	7.9	*83.3 PDD-NOS*	41.7	41.7
126–130	3.3	*66.7 PDD-NOS*	33.3	25
131–135	1.0	*58.3 PDD-NOS*	8.3	8.3
136–140	0	50	8.3	8.3
141–145	0	50	8.3	0
146–150	0	*50 HFA*	0	0
151–155	0	*41.7 AS*	0	0
156–160	0	25	0	0
161–165	0	25	0	0
166–170	0	*25 AS/HFA*	0	0
171–175	0	*8.3*	0	0
176–180	0	*8.3*	0	0
181–185	0	*8.3 AS*	0	0
186–200	0	0	0	0

The ASC group is separated for AS, HFA and PDD-NOS diagnoses

Note: ASC = autism spectrum condition; SAD = social anxiety disorder; OCD = obsessive-compulsive disorder; AS = Asperger syndrome; HFA = high functioning autism; PDD-NOS = pervasive developmental disorder not otherwise specified

other patient groups, nor subjects from the general population obtained scores in this segment, while all subjects with an AS or HFA diagnosis scored in this range. Subjects with a PDD-NOS diagnosis obtained lower scores than the AS/HFA patients (total AQ scores between 114 and 134), but scored higher than the majority of the other psychiatric patients and the general population. Figure 2 illustrates that AQ scores followed a continuous distribution in the general population. The AQ scores of ASC patients all fell in the right end of this distribution.

Discussion

The results of our study indicate that the Autism-Spectrum Quotient is a valid and reliable instrument to assess individual differences in autistic traits. The Dutch AQ was found to have satisfactory internal consistency and test–retest reliability. Moreover, high AQ scores were specific to ASC patients. Previously, an exploratory factor analysis performed in a British student population suggested three underlying factors within the AQ, encompassing Social skill, Details/patterns, and Communication/mind reading (Austin 2005). A recent study using the child version of the AQ found evidence for 4 of the 5 domains (Auyeung et al. 2008). Our confirmatory factor analyses in both a general population and a student sample indicated that 4 of the 5 domains of the AQ (Social skill, Communication, Attention switching, and Imagination) were highly correlated. Rather than four separate domains, we propose a hierarchical model allowing these domains to cluster together. This way, one broad band "Social interaction" factor (incorporating the 4 highly correlated domains) could be identified, together with a small second factor, consisting

Fig. 2 AQ Scores in the General Population and the three Patient Groups. All Subjects scoring >145 were diagnosed with Asperger Syndrome or High Functioning Autism. ASC = Autism Spectrum Condition; SAD = Social Anxiety Disorder; OCD = Obsessive-Compulsive Disorder
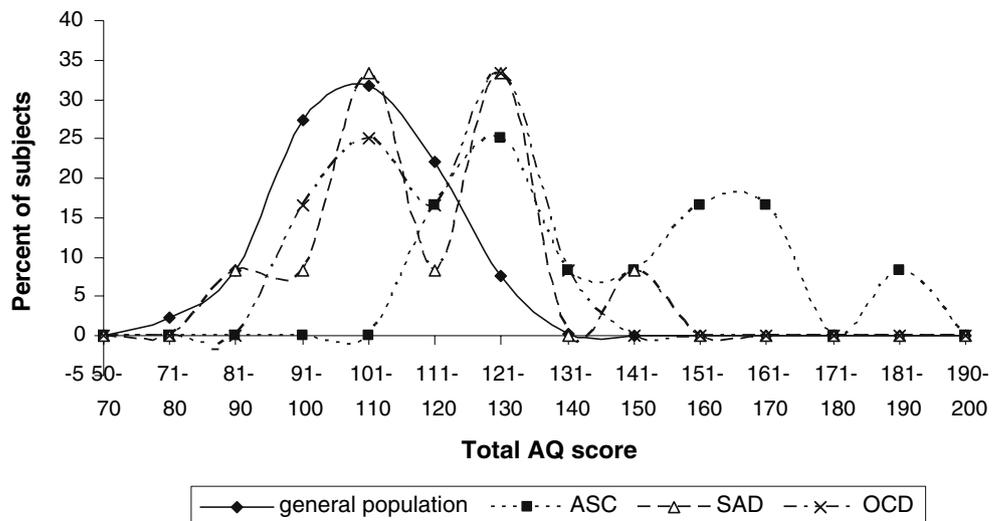
of items focusing on a preference for details and patterns (the domain "Attention to detail").

It may seem surprising that the broad band scale also includes the domain Attention switching. As put forward by Courchesne and colleagues (1994), difficulties in Attention switching make it harder to keep track of social information. Social interaction usually involves frequent and rapid changes in the source of information (visually or auditory information, change in objects or actions, etc.), and requires the ability to follow the flow of social cues (words, gestures, postures, background context, etc). The marked deficit in attention switching in ASC subjects may directly harm their social and communication abilities, which explains why this domain is included within the Social interaction factor.

The internal consistency and test-retest reliability of the total AQ score ($\alpha = .81/.71$ and $r = .78$) and the broad band social interaction factor ($\alpha = .84/.77$ and $r = .79$) were satisfactory. The internal consistency of the Attention to detail factor was moderate ($\alpha = .63/68$) and its test–retest reliability was acceptable ($r = .71$). The internal consistencies of the four domains encompassing the broad band social interaction factor varied but were generally moderate (α between .39 and .76), the test–retest reliability varied between .60 and .81. Our results indicate that future studies using the AQ to assess autistic traits may benefit from a focus on the total AQ score and the 2 factor scores, rather than examining differences and similarities in the five original domains. The power to detect differences will be higher using the 2 factors; this could be of importance especially in linkage or association studies into autistic traits.

A small but significant correlation ($r = .19$) was found between the two Social interaction factor and the Attention to detail factor. This result is in agreement with findings from a British twin project (Ronald et al. 2005) which assessed both social and non-social behaviors characteristic for autism using parent and teacher ratings. Social and non-social behaviors were weakly correlated with each other, both in the teacher ($r = .15$) and the parent data ($r = .29$).

Group differences in AQ scores were in line with previous studies using the AQ (Austin 2005; Baron-Cohen et al. 2001; Hoekstra et al. 2007; Wakabayashi et al. 2006). Males scored higher than females on the total AQ score and the Social interaction factor, although no sex difference on the Attention to detail factor was found. Similar to the findings reported by Baron-Cohen et al. (2001) and Austin (2005), science students obtained significantly higher scores than humanities and social science students. In our study, students enrolled in a social science degree in turn scored significantly lower than humanities students. Rather than a remarkable discrepancy with the British results (no differences between humanities and

social sciences students), these findings are probably due to differences in the student sample. The majority of the social science students included in our study was enrolled in a psychology or education degree. Both these studies have a strong focus on human interaction on an individual level and are likely to attract students who enjoy social interaction. If social sciences focusing on a broader level (such as economics and political science) had been included, the difference in AQ score might not have been significant.

Subjects diagnosed with an ASC scored significantly higher on the total AQ than the general population and the other patient groups. This is a satisfying result, considering the potential symptom overlap between these patient groups and ASC. The overlap between ASC subjects and SAD patients entails problems with social interaction in both groups; the overlap between ASC and OCD subjects encompasses repetitive behaviors that occur in both groups. Relatives of autistic individuals are reported to have up to 10-fold higher rates of social phobia compared to control families (Smalley et al. 1995; Piven and Palmer 1999). Similarly, an increased incidence of OCD is found in autism relatives (Bolton et al. 1998), and the occurrence of obsessive–compulsive traits in parents of an autistic child is significantly more likely if the child displays strong repetitive behavior (Hollander et al. 2003). Recent research suggests that a common genetic pathway, the serotonin transporter gene, could explain a small part of the association between OCD and rigid compulsive behaviors in autism (Ozaki et al. 2003; Sutcliffe et al. 2005). Our study however shows that high AQ scores are specific to the ASC patients. ASC patients scored significantly higher than the other patient groups on both the Social interaction factor and the Attention to detail factor.

Moreover, all subjects with an AS/HFA diagnosis could be distinguished from the other samples, as no subjects without an ASC diagnosis obtained a score >145. The difference is less clear-cut for subjects with PDD-NOS, who obtained intermediate AQ scores. These findings could explain the discrepancy in previous studies using the AQ. Woodbury-Smith et al. (2005) reported satisfying ability of the AQ to distinguish patients with an AS/HFA diagnosis from non-AS/HFA patients in a clinic for adults suspected of having AS or HFA. On the other hand, Ketelaars et al. (2007) found little differences in AQ scores between mild ASC patients and patients referred to the general outpatient clinic. The latter study mainly included PDD-NOS patients and few patients with AS or HFA. PDD-NOS constitutes a broad diagnostic category with criteria less stringent than for autistic disorder (American Psychiatric Association 2000) and is often seen as a milder form of autism. A lower AQ score in PDD-NOS patients compared to AS/HFA patients, yet higher than in most

subjects diagnosed with another psychiatric disorder or subjects from the general population, is compatible with the notion of a continuum of autistic traits.

This study had some limitations. Firstly, the data collection in the student and general population sample was based on anonymous reports. We could therefore not verify whether any subjects in these groups had an ASC diagnosis or should warrant a diagnosis. Furthermore, the sample sizes of the patient groups were small and findings from these groups should be interpreted with care. Future studies in clinical samples should explore the criterion validity of the AQ more extensively, and should especially focus on the differences between severely impaired ASC patients and mildly impaired PDD-NOS subjects.

In conclusion, this study shows that the AQ is a reliable instrument for examining variation in autistic traits. The AQ can be divided into two reliable sub factors, focusing on difficulties in social interaction and on marked interests and attention to details and patterns. Total AQ scores follow a continuous distribution in the general population, and patients with an ASC diagnosis fall in the upper end of this distribution. OCD and SAD patient obtain scores between the general population mean and scores typical for ASC patients. These findings indicate that the AQ is a valuable instrument to assess where an individual lies on the autism spectrum.

References

American Psychiatric Association (2000). *Diagnostic and Statistical Manual for Mental Disorders*, 4th Ed., Text Revision (DSM-IV-TR). Washington: American Psychiatric Press.

Austin, E. J. (2005). Personality correlates of the broader autism phenotype as assessed by the Autism Spectrum Quotient (AQ). *Personality and Individual Differences*, 38, 451–460.

Auyeung, B., Baron-Cohen, S., Wheelwright, S., & Allison, C. (2008). The Autism Spectrum Quotient – Children's Version (AQ-Child). *Journal of Autism and Developmental Disorders*, In press.

Bailey, A., Le Couteur, A., Gottesman, I., Bolton, P., Simonoff, E., Yuzda, E., et al. (1995). Autism as a strongly genetic disorder: Evidence from a British twin study. *Psychological Medicine*, 25, 63–77.

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

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

Baron-Cohen, S., Knickmeyer, R. C., & Belmonte, M. K. (2005). Sex differences in the brain: Implications for explaining autism. *Science*, 310, 819–823.

Baron-Cohen, S., Wheelwright, S., Skinner, R., & Martin, C. 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.

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

Bolton, P., Macdonald, H., Pickles, A., Rios, P., Goode, S., Crowson, M., et al. (1994). A case-control family history Study of Autism. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 35, 877–900.

Bolton, P. F., Pickles, A., Murphy, M., & Rutter, M. (1998). Autism, affective and other psychiatric disorders: Patterns of familial aggregation. *Psycholical Medicine*, 28, 385–395.

Cath, D. C., Ran, N., Smit, J. H., van Balkom, A. J. L. M., & Comijs, H. C. (2008). The relationship between pervasive developmental disorders, OCD and social anxiety disorder in adults: A preliminary case controlled study. *Psychopathology*, 41, 101–110.

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

Constantino, J. N., & Todd, R. D. (2005). Intergenerational transmission of subthreshold autistic traits in the general population. *Biological Psychiatry*, 57, 655–660.

Constantino, J. N., Lajonchere, C., Lutz, M., Gray, T., Abbacchi, A., McKenna, K., et al. (2006). Autistic social impairment in the siblings of children with pervasive developmental disorders. *American Journal of Psychiatry*, 163, 294–296.

Courchesne, E., Townsend, J., Akshoomoff, N. A., Saitoh, O., Yeung-Courchesne, R., Lincoln, A. J., et al. (1994). Impairment in shifting attention in autistic and cerebellar patients. *Behavioral Neuroscience*, 108, 848–865.

First, M. B. (1996). Structured clinical interview for DSM-IV axis I disorders. New York.

Folstein, S., & Rutter, M. (1977). Genetic influences and infantile autism. *Nature*, 265, 726–728.

Hoekstra, R. A., Bartels, M., Verweij, C. J. H., & Boomsma, D. I. (2007). Heritability of autistic traits in the general population. *Archives of Pediatric and Adolescent Medicine*, 161, 372–377.

Hollander, E., King, A., Delaney, K., Smith, C. J., & Silverman, J. M. (2003). Obsessive-compulsive behaviors in parents of multiplex autism families. *Psychiatry Research*, 117, 11–16.

Ketelaars, C., Horwitz, E., Sytema, S., Bos, J., Wiersma, D., Minderaa, R., & Hartman, C. A., (2007). Brief report: Adults with mild Autism Spectrum Disorders (ASD): Scores on the Autism Spectrum Quotient (AQ) and comorbid psychopathology. *Journal of Autism and Developmental Disorders*, In press.

Kurita, H., Koyama, T., & Osada, H. (2005). Autism-Spectrum Quotient-Japanese version and its short forms for screening normally intelligent persons with pervasive developmental disorders. *Psychiatry Clinical Neurosciences*, 59, 490–496.

Le Couteur, A., Bailey, A., Goode, S., Pickles, A., Robertson, S., Gottesman, I., et al. (1996). A broader phenotype of autism: the clinical spectrum in twins. *Journal of Child Psychology and Psychiatry, 37*, 785–801.

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Jr., Leventhal, B. L., DiLavore, P. C., et al. (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.

Lord, C., Rutter, M., & Le Couteur, A. (1994). Autism diagnostic interview-revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *Journal of Autism and Developmental Disorders, 24*, 659–685.

Neale, M. C., Boker, S. M., Xie, G., & Maes, H. H. (2006). *Mx: Statistical modeling*. 7th Ed. Richmond: VCU.

Ozaki, N., Goldman, D., Kaye, W. H., Plotnicov, K., Greenberg, B. D., Lappalainen, J., et al. (2003). Serotonin transporter missense mutation associated with a complex neuropsychiatric phenotype. *Molecular Psychiatry, 8*(895), 933–936.

Piven, J., & Palmer, P. (1999). Psychiatric disorder and the broad autism phenotype: Evidence from a family study of multiple-incidence autism families. *American Journal of Psychiatry, 156*, 557–563.

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

Ronald, A., Happe, F., & Plomin, R. (2005). The genetic relationship between individual differences in social and nonsocial behaviours characteristic of autism. *Developmental Scence, 8*, 444–458.

Rutter, M. (2000). Genetic studies of autism: From the 1970s into the millennium. *Journal of Abnormal Child Psychology, 28*, 3–14.

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

Smalley, S. L., McCracken, J., & Tanguay, P. (1995). Autism, affective disorders, and social phobia. *American Journal of Medical Genetics, 60*, 19–26.

Spiker, D., Lotspeich, L. J., Dimiceli, S., Myers, R. M., & Risch, N. (2002). Behavioral phenotypic variation in autism multiplex families: evidence for a continuous severity gradient. *American Journal of Medical Genetics, 114*, 129–136.

Sutcliffe, J. S., Delahanty, R. J., Prasad, H. C., McCauley, J. L., Han, Q., Jiang, L., et al. (2005). Allelic heterogeneity at the serotonin transporter locus (SLC6A4) confers susceptibility to autism and rigid-compulsive behaviors. *American Journal of Human Genetics, 77*, 265–279.

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

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, 35*, 331–335.