

The Construction and Validation of an Abridged Version of the Autism-Spectrum Quotient (AQ-Short)

Rosa A. Hoekstra · Anna A. E. Vinkhuyzen · Sally Wheelwright ·
Meike Bartels · Dorret I. Boomsma · Simon Baron-Cohen ·
Danielle Posthuma · Sophie van der Sluis

Published online: 10 August 2010

© The Author(s) 2010. This article is published with open access at Springerlink.com

Abstract This study reports on the development and validation of an abridged version of the 50-item Autism-Spectrum Quotient (AQ), a self-report measure of autistic traits. We aimed to reduce the number of items whilst retaining high validity and a meaningful factor structure. The item reduction procedure was performed on data from 1,263 Dutch students and general population adults. The resulting 28-item AQ-Short was subsequently validated in 3 independent samples, both clinical and controls, from the Netherlands and the UK. The AQ-Short comprises two higher-order factors assessing ‘social behavioral difficulties’ and ‘a fascination for numbers/patterns’. The clear factor structure of the AQ-Short and its high sensitivity and specificity make the AQ-Short a useful alternative to the full 50-item version.

Keywords Autism · Factor analysis · Validity · Reliability · Autism phenotype · Autism-spectrum quotient

Introduction

Autism spectrum conditions (ASC; including autistic disorder, Asperger syndrome [AS] and pervasive developmental disorder not otherwise specified [PDD-NOS]) are now commonly conceptualized as dimensional, representing the extreme end of one or more continuously distributed traits in the general population (Constantino & Todd 2003). The dimensional approach to autism inspired the development of the Autism-Spectrum Quotient (AQ), a 50-item self-report questionnaire assessing autistic traits in individuals with normal intelligence (Baron-Cohen et al. 2001). Studies using the AQ in the UK (Baron-Cohen et al. 2001), Japan (Wakabayashi et al. 2006) and the Netherlands (Hoekstra et al. 2008) reported significantly elevated AQ scores in participants diagnosed with ASC compared to the general population and a student sample. AQ scores in a small ASC sample were also significantly higher than in people with social anxiety disorder or obsessive compulsive disorder (Hoekstra et al. 2008), suggesting that a high score on the AQ is specific to ASC rather than to psychiatric problems in general. Moreover, the AQ has been found to predict diagnosis of AS in a clinic sample (Woodbury-Smith et al. 2005). Whilst most studies—including the present report—have been conducted using the adult version of the AQ (AQ-Adult), which is self-report, parental-report adolescent (AQ-Adolescent; Baron-Cohen et al. 2006) and child (AQ-Child; Auyeung et al. 2008) versions also exist.

The English (Baron-Cohen et al. 2001), Japanese (Wakabayashi et al. 2006), and Dutch (Hoekstra et al. 2008) versions of the AQ-Adult all show good test-retest

R. A. Hoekstra (✉)

Department of Life Sciences, Faculty of Science, The Open University, Walton Hall, Milton Keynes MK7 6AA, UK
e-mail: R.A.Hoekstra@open.ac.uk

R. A. Hoekstra · S. Wheelwright · S. Baron-Cohen
Autism Research Centre, Department of Psychiatry,
University of Cambridge, Cambridge, UK

A. A. E. Vinkhuyzen · M. Bartels · D. I. Boomsma
Department of Biological Psychology,
VU University, Amsterdam, The Netherlands

A. A. E. Vinkhuyzen · D. Posthuma · S. van der Sluis
CNCr, Department of Functional Genomics,
VU University, Amsterdam, The Netherlands

D. Posthuma
Department of Medical Genomics,
VU Medical Centre, Amsterdam, The Netherlands

reliability and acceptable internal consistency. In the original version of the AQ, the 50 items were divided into 5 empirically derived subdomains (Baron-Cohen et al. 2001) measuring Social skills, Communication, Imagination, Attention to detail, and Attention switching. Subsequent factor analytic studies in non-clinical populations confirmed that the AQ has a multifactorial structure. Austin (2005) and Hurst et al. (2007) studied the structure of the AQ-Adult in student samples using principal component analysis. Both studies suggested three factors, representing Social skills, Details/patterns, and Communication/mindreading. Auyeung et al. (2008) explored the structure of the AQ-Child and reported 4 factors, measuring Social skills, Attention to detail, Mindreading, and Imagination. In a previous report, Hoekstra et al. (2008) explored the factor structure of the AQ-Adult in a general population and a student sample using confirmatory factor analysis and identified 2 higher-order factors measuring Social interaction and Attention to detail. Altogether these different studies suggest that the AQ is multifactorial and encompasses at least one factor pertaining to social behaviors and one factor assessing non-social traits. The discrepancy between the findings from the different studies are most likely due to the use of different study samples (children, students, or general population adults), which could result in slightly different psychometric qualities of the AQ, and to the use of different statistical techniques to evaluate the factor structure (principal component analysis or confirmatory factor analysis).

Together, the studies so far suggest that the AQ is a reliable instrument to quantify the autism phenotype. The promising results of the AQ have prompted its use in large population studies (Hoekstra et al. 2007a, b) and in studies of the cognitive (Bayliss et al. 2005; Lombardo et al. 2007), neural (Gomot et al. 2008), genetic (Chakrabarti et al. 2009), and hormonal (Auyeung et al. 2009) correlates of autistic traits. However, the full 50-item version of the AQ may often be too lengthy to be included in large comprehensive studies. The aim of the current study was to construct a shortened version of the original 50-item AQ-Adult, whilst retaining as much information as possible. Furthermore, we aimed to retain a scale with a clear factor structure that would allow univocal interpretation.

Methods

Participants

The data included in this study comprised 4 independent samples. The first data set (Dutch controls reference sample, $n = 1,263$) consisted of a Dutch general population sample and a student group (respectively, $n = 302$ and 961) that both have been described in more detail

elsewhere (Hoekstra et al. 2008). The students were registered at either the VU University in Amsterdam or the University of Twente in Enschede and were asked to fill out the Dutch translation of the full-scale AQ-Adult during the tea break of one of their classes. The general population sample was recruited on an information day for parents of multiples. Participants were asked to fill out the full-scale AQ on the same day or return the questionnaire to the research group by mail. Mean age of the students was 21.19 years ($SD = 3.69$), mean age of the general population sample was 35.68 years ($SD = 6.33$). The combined sample included 502 men and 739 women (22 cases sex unknown). Previous analyses of these data showed no significant age effect on AQ-scores (Hoekstra et al. 2008).

The second sample (Dutch controls replication sample, $n = 1,121$; 485 men, 363 women; age: Mean = 45.63 years, $SD = 14.74$) included adults from the Dutch general population who participated in a large extended twin family study on the influences of genes and environment on cognition and behavior. In this study, the AQ-Short was part of a larger questionnaire that was sent to the participants by mail and could be returned by mail or during a test session.

The third sample (English controls sample, $n = 1,838$; 737 men, 1,101 women; age: Mean = 20.90 years, $SD = 2.47$) comprised students from the University of Cambridge. They were recruited via several routes including email, post, newspaper adverts and notices around the university, and invited to complete the full-scale AQ using an online version. Participants who reported a history of psychiatric difficulties (depression, ASC, bipolar illness, psychosis or anorexia) were excluded from the analysis.

The fourth sample encompassed individuals with a formal AS diagnosis (the English AS sample, $n = 274$; 156 men, 117 women, 1 sex unknown; age: Mean = 35.37 years, $SD = 13.05$). These participants, all volunteers registered in the Cambridge Autism Research Centre database (see www.autismresearchcentre.com), filled out the full-scale AQ online. All participants were diagnosed by experienced clinicians according to DSM-IV or ICD-10 criteria. The large majority was diagnosed with AS in adulthood.

Data from the Dutch controls reference sample were used to develop the short version of the AQ. Data from the Dutch controls replication sample and the English controls were used to verify the factor structure, whilst the AS sample was included to examine AQ-Short scores in a clinical sample.

Materials

The full-scale AQ comprises 50 descriptive statements assessing personal preferences and habits. Participants respond to each statement on a 4-point Likert scale, with answer categories “1 = definitely agree”; “2 = slightly agree”; “3 = slightly disagree” and “4 = definitely

disagree”. The scoring is reversed for items in which an “agree” response is characteristic for autism (24 out of the 50 items). Item scores are summed, resulting in a minimum AQ score of 50 (indicating no autistic traits) and a maximum score of 200 (full endorsement of all autistic traits).¹

The Dutch controls reference sample and both English samples filled out the full-scale AQ. The Dutch controls replication sample filled out the AQ-Short. Following the criteria set out in the original AQ paper (Baron-Cohen et al. 2001), if more than 10% of the total number of items were missing (>5 in the full-scale AQ, >3 in the AQ-Short), the questionnaire was considered unreliable and the data were discarded.

Procedure

The aim was to reduce the number of items of the AQ, whilst retaining as much information as possible, and ultimately obtaining a shortened questionnaire with a clear factor structure. To this end, we used the following 5-step approach on the data of the Dutch controls reference sample:

1. Inspection of the item content of all 50 items. Based on these inspections, some items were flagged as possibly problematic. For instance, items that are very similar in content and phrasing may disrupt the factor structure of a larger number of items.

2. Exploratory factor analysis (EFA) on all 50 items. The fit of different factor structures was evaluated and compared to the 5-domain model originally proposed by Baron-Cohen et al. (2001). If the item loading was high (>.45) on another domain than the one suggested by Baron-Cohen et al. (2001), the item was added to this other scale and omitted from its original domain. Items were allowed to load on multiple factors if cross-loadings were high (>.45).

3. Domain-specific EFAs on all items included in each domain. If the items flagged in step 1 were indeed causing problems (e.g., the exploratory factor structure looked very different when this item was omitted), the item was removed. This step was used to get an idea of whether the empirically derived factors in step 2 were uni-dimensional, or should be further divided into smaller subscales. Reduction of the total number of items was achieved by removing items with low (<.30) factor loadings.

4. Confirmatory factor analysis (CFA) per domain. The factor structure suggested by the EFAs in step 3 was

replicated, and it was examined whether the fit of the resulting model was sufficient or could be improved by introducing minor adjustments, using modification indices² as guidelines. The fit of these models was evaluated using the Comparative Fit Index (CFI) and the Root Mean Square Error of Approximation (RMSEA). The CFI should ideally be above .90, whilst the RMSEA should be below .08 (Schermelleh-Engel et al. 2003). Note that the CFI is calculated as the difference in fit between the independence model (i.e., all items are uncorrelated) and the hypothesized model. If items are ordinal and item intercorrelations are low, the fit of the independence model will not be very bad. In that case, the difference between the fit of this model and the hypothesized model, i.e., the CFI, will thus not become large even if the model describes the data well.³ Several intercorrelations were <.20 in our AQ data. We therefore somewhat relaxed the criterion for CFI, but kept the norm <.08 for the RMSEA.

5. CFA combining all factors identified under step 4. Again, model fit was evaluated using CFI and the RMSEA.

All model fitting was performed using Mplus (Muthén & Muthén 2007) on the raw (ordinal) data using option ‘Categorical’ and weighted least squares estimation with mean and variance adjusted Chi-squares (wlsmv procedure) and Geo-min rotation. Following construction of the AQ-Short using the Dutch reference sample, the same factor structure of the AQ-Short was evaluated in the Dutch controls replication sample⁴ and the English replication sample. The factor structure was not evaluated in the English AS sample due to limited sample size. This latter sample was included to evaluate the correlation between the full-scale AQ and the AQ-Short in a clinical sample, and to test for group differences in mean AQ-Short scores. Lastly, test accuracy was examined by ROC analysis, and a cut-off score is suggested.

Results

Construction AQ-Short

The 5-step procedure described in the “Method” section was followed to arrive at an abridged version of the AQ,

¹ In most previous reports using the English version of the AQ (Baron-Cohen et al. 2001) the answer categories are dichotomized into “agree”/“disagree” scores. In these studies, all item responses in line with the autism phenotype scored one point, resulting in AQ scores ranging between 0 and 50.

² Modification indices express the expected improvement in model fit if constrained parameters are estimated freely, and can be used to identify local misfit in factor models.

³ Please see the website of expert David A Kenny for further information on this topic: <http://davidakenny.net/cm/fit.htm>.

⁴ We corrected for the genetic relatedness of the participants in the Dutch controls replication sample using the option ‘complex’ in Mplus.

using the data from the Dutch controls reference sample. Inspection of the item content (step 1) revealed three pairs of items (items 40 & 50, 29 & 49, and 17 & 38) with similar content or phrasing. An EFA on all 50 items (step 2) showed that a 5-factor model fitted reasonably well and roughly coincided with the 5 domains originally proposed by Baron-Cohen et al. (2001). This EFA also indicated that items 36 and 45 (originally part of the Social skills domain) clustered together with items from the Imagination domain. As both items concern empathic imagination (see Table 1) they fit well in the Imagination domain and were therefore moved to this factor. Similarly, item 41 loaded highly on the Attention to detail domain, and as this domain fitted the item's content (see Table 1), the item was moved from the Imagination scale to the Attention to detail domain. Of the items flagged in step 1, items 17, 29, 40 and 49 were omitted in step 3 (domain-specific EFA) as they either disrupted the factor structure (17 and 40) or showed overall low factor loadings (29 and 49). Items 38 and 50 were retained in the domains Communication (38) and Imagination (50).

Domain-specific EFA and subsequent CFA on the data from the Dutch controls reference sample (step 4) showed that for the Social skills domain, a 1-factor model encompassing 7 items fitted reasonably well (CFI = .86, RMSEA = .06) (see Table 1 for item content). EFA and CFA showed that the original Attention switching domain was not uni-dimensional. A 2-factor solution proved more appropriate (CFI = .96, RMSEA = .05), with one factor assessing 'Routine' (4 items) and another factor measuring 'Switching' ability (4 items). For the Imagination domain, a 1-factor model comprising 8 items fitted well (CFI = .97, RMSEA = .05). Similarly, for the Communication domain, a 1-factor model including 8 items provided a good fit (CFI = .94, RMSEA = .05). EFA on the original Attention to detail domain showed low factor loadings for some of the items. These items were omitted and CFA on the remaining 5 items showed a well fitting 1-factor model (CFI = .99, RMSEA = .06) specifically assessing a fascination for 'Numbers/patterns'.

Next, the data from the Dutch controls reference sample were used to conduct CFA on all six newly formed subscales (step 5). The resulting model indicated that the scale Communication correlated strongly with the scales Social skills ($r = .79$) and Imagination ($r = .85$), whilst the two latter scales correlated moderately with each other ($r = .47$). Since it was our aim to retain as much information as possible with a minimal number of items, all Communication items were eliminated from the AQ-Short. CFA on the remaining five factors showed that the factors Social skills, Routine, Switching and Imagination correlated substantially with each other (r between .43 and .74). These factors could be subsumed under a higher-order

Table 1 Item content of the 5 factors in the best fitting structure (item number in the full-scale AQ in parentheses)

| Higher-order factor Social Behavior | |
|--|--|
| Social skills | |
| I prefer to do things with others rather than on my own (1) | |
| I find social situations easy (11) | |
| I would rather go to a library than to a party (13) | |
| I find myself drawn more strongly to people than to things (15) | |
| I find it hard to make new friends (22) | |
| I enjoy social occasions (44) | |
| I enjoy meeting new people (47) | |
| <i>English samples only:</i> New situations make me anxious (46) | |
| Routine | |
| I prefer to do things the same way over and over again (2) | |
| It does not upset me if my daily routine is disturbed (25) | |
| I enjoy doing things spontaneously (34) | |
| New situations make me anxious (46) | |
| Switching | |
| I frequently get strongly absorbed in one thing (4) | |
| I can easily keep track of several different people's conversations (10) | |
| I find it easy to do more than one thing at once (32) | |
| If there is an interruption, I can switch back very quickly (37) | |
| Imagination | |
| Trying to imagine something, I find it easy to create a picture in my mind (3) | |
| Reading a story, I can easily imagine what the characters might look like (8) | |
| I find making up stories easy (14) | |
| Reading a story, I find it difficult to work out the character's intentions (20) | |
| I find it easy to work out what someone is thinking or feeling (36) | |
| I find it difficult to imagine what it would be like to be someone else (42) | |
| I find it difficult to work out people's intentions (45) | |
| I find it easy to play games with children that involve pretending (50) | |
| Factor numbers and patterns | |
| I usually notice car number plates or similar strings of information (6) | |
| I am fascinated by dates (9) | |
| I am fascinated by numbers (19) | |
| I notice patterns in things all the time (23) | |
| I like to collect information about categories of things (41) | |
| Items omitted from the AQ-Short | |
| I often notice small sounds when others do not (5) | |
| Other people frequently tell me that what I have said is impolite (7) | |
| I tend to notice details that others do not (12) | |
| I tend to have very strong interests (16) | |

Table 1 continued

| Items omitted from the AQ-Short |
|---|
| I enjoy social chit-chat (17) |
| When I talk, it isn't always easy for others to get a word in edgeways (18) |
| I don't particularly enjoy reading fiction (21) |
| I would rather go to the theatre than a museum (24) |
| I don't know how to keep a conversation going (26) |
| I find it easy to "read between the lines" (27) |
| I usually concentrate more on the whole picture, rather than the small details (28) |
| I am not very good at remembering phone numbers (29) |
| I know how to tell if someone listening to me is getting bored (31) |
| I don't usually notice small changes in a situation, or a person's appearance (30) |
| When I talk on the phone, I am not sure when it's my turn to speak (33) |
| I am often the last to understand the point of a joke (35) |
| I am good at social chit-chat (38) |
| People tell me that I keep going on and on about the same thing (39) |
| When younger, I enjoyed playing games involving pretending with others (40) |
| I like to plan any activities I participate in carefully (43) |
| I am a good diplomat (48) |
| I am not very good at remembering people's date of birth (49) |

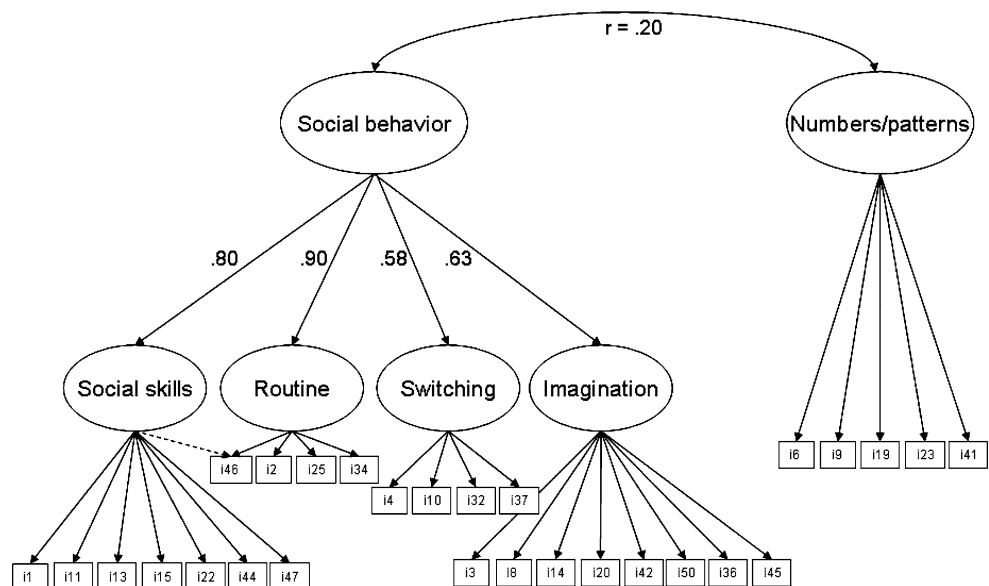
factor 'Social behavior'. The factor Numbers/patterns only correlated modestly with the other factors (higher-order factorial correlation = .20). The final factor model, depicted in Fig. 1, fitted the data of the Dutch controls reference sample well (CFI = .87, RMSEA = .06).

In sum, the 50-item AQ was shortened to 28 items. These 28 items can be assigned to five clearly defined factors, assessing difficulties with social skills ('Social skills'), a preference for routine ('Routine'), attention switching difficulties ('Switching'), difficulties with imagination ('Imagination'), and a fascination for numbers/patterns ('Numbers/patterns') (see Table 1 for item content).

Factor Structure Replication and Validation AQ-Short

Subsequently, the best fitting model identified using the data from the Dutch controls reference sample was tested in respectively the Dutch controls replication sample and the English controls. Application of the hierarchical factor model to the data of the Dutch controls replication sample resulted in a reasonable model fit (CFI = .87, RMSEA = .07). Inspection of the modification indices (MI) suggested that the fit could not be appreciably improved by changing one or more of the parameters (all MI < 28). Next, the same model was fitted to the English controls sample. The MI showed that the fit would improve considerably if one item (item 46, "New situations make me anxious") was allowed to load on the factor Social skills in addition to its loading on the factor Routine. The model including this cross-loading showed acceptable fit (CFI = .86, RMSEA = .07). The differential functioning of item 46 in the Dutch and the English samples might be due to a slight interpretational difference of the word 'anxious' in Dutch. The connotation of the Dutch translation ('angstig') is closer to 'fearful' than to 'worried'. This relatively strong expression of fear may have resulted in the item loading more strongly on the Routine factor and not on the Social factor in the Dutch samples. Note that both scales are part

Fig. 1 Factor structure of the AQ-Short, including factor correlation and factor loadings as estimated in the Dutch controls reference sample



of the higher-order Social behavior factor. As in the Dutch controls reference sample, the factorial correlation between Social behavior and Numbers/patterns was only modest in the Dutch controls replication sample ($r = .10$) and in the English controls sample ($r = .16$).

The distribution of the AQ-Short scores was approximately normal in all three control samples, and slightly skewed to the left in the English AS sample (see Fig. 2). In the three samples for which full-scale AQ data were available, Pearson's correlations were calculated between sumscores based on the full-scale AQ and the AQ-Short. The correlations were very high and significant in all samples (Dutch controls reference sample: $r = .93$, $p < .001$; English controls sample: $r = .94$, $p < .001$; English AS sample: $r = .95$, $p < .001$). Table 2 displays the internal consistency of the five scales and the higher-order factor Social behavior in each of the independent samples. The Cronbach's alpha values indicate acceptable to good internal consistency for the total AQ-Short (α between .77 and .86), the broad Social behavior factor

(α between .79 and .86) and the Numbers/patterns factor (α between .67 and .73). The internal consistency for the scales Routine and Switching are somewhat low, but this is probably due to the small number of items in both scales (α values usually increase with more items).

Table 3 shows the mean AQ-Short scores and the scores on the factors Social Behavior and Numbers/patterns in the two Dutch control samples, the English control sample, and the English AS sample. In all three control samples, men obtained significantly higher scores than women on the total AQ-Short, and on the Numbers/patterns factor. The sex difference was also significant for the Social behavior factor in the Dutch controls reference sample and the English controls sample. Against expectation, this effect failed to be significant in the Dutch controls replication sample. Analyses of the English AS sample showed that women with AS scored significantly higher on the AQ-Short than men with AS. This sex difference was mainly reflected in the Social behavior factor and was not significant for Numbers/patterns. The sex differences in this

Fig. 2 Distribution of mean AQ-Short scores in all samples

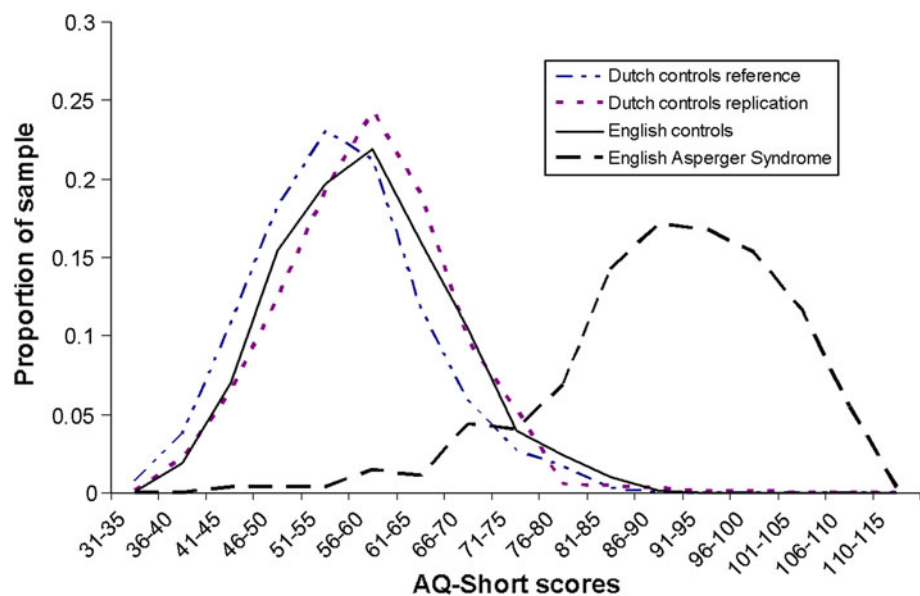


Table 2 Internal consistency coefficients (Cronbach's alpha) of the factors making up the AQ-Short in all samples (number of items in each scale in parentheses)

| Factor (# items in scale) | Dutch controls reference | Dutch controls replication | English controls | English AS |
|-----------------------------------|--------------------------|----------------------------|------------------|------------|
| AQ-Short (28) | .79 | .77 | .78 | .86 |
| Broad factor | .80 | .79 | .80 | .86 |
| Social behaviour (23) | | | | |
| Social skills (7) | .76 | .72 | .79 | .80 |
| Social skills English version (8) | | | .80 | .80 |
| Routine (4) | .55 | .54 | .57 | .62 |
| Switching (4) | .56 | .54 | .47 | .59 |
| Imagination (8) | .68 | .69 | .68 | .75 |
| Numbers/patterns (5) | .70 | .71 | .67 | .73 |

Table 3 Mean AQ-Short and factor scores in all samples (SD in parentheses)

| Sample | Sex | N | AQ-Short | Social behavior | Numbers/patterns |
|---------------------------------------|-------|-------|-----------------------------|-----------------------------|----------------------------|
| Dutch controls reference ^a | Men | 502 | 56.91 (9.32) | 46.16 (8.31) | 10.75 (3.18) |
| | Women | 739 | 52.79 (8.06)*** | 43.70 (7.27)*** | 9.10 (2.90)*** |
| Dutch controls replication | Men | 485 | 58.40 (8.70) | 47.02 (7.94) | 11.39 (3.25) |
| | Women | 636 | 56.61 (8.63)** | 46.86 (7.73) ^{ns} | 9.75 (3.18)*** |
| English controls | Men | 737 | 59.73 (8.90) | 47.95 (8.30) | 11.79 (2.96) |
| | Women | 1,101 | 56.00 (8.88)*** | 45.56 (7.99)*** | 10.44 (3.15)*** |
| English AS ^b | Men | 156 | 87.76 (12.06) | 72.51 (10.43) | 15.26 (3.45) |
| | Women | 117 | 91.49 (11.62)* | 76.07 (9.98)** | 15.42 (3.46) ^{ns} |
| English controls versus AS | | | $F(1, 2,107) = 2695.17$ *** | $F(1, 2,107) = 2483.38$ *** | $F(1, 2,107) = 423.69$ *** |

AS Asperger syndrome; ^a sex unknown in 22 participants; ^b sex unknown in 1 participant

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

clinical sample should be interpreted with care, as selection bias may have played a role. The ratio of women to men participating in this study (1:1.3) is much higher than the typical sex ratio reported for AS/high-functioning autism (1:5.5; Fombonne 2006). As expected, people with an AS diagnosis scored significantly higher on the AQ-Short than controls (see also Fig. 2).

The accuracy of the AQ-Short in distinguishing individuals with AS from controls was evaluated using ROC analysis. The area under the curve was .97, indicating excellent test accuracy. A score >65 had a sensitivity of .97 and a specificity of .82. With a more stringent cut-off of ≥ 70 , the sensitivity and specificity were .94 and .91, respectively.

Discussion

This paper described the development and validation of an abridged version of the AQ, using data from four independent samples from two different countries. We found that the AQ can be shortened from 50 to 28 items; the resulting AQ-Short correlates very highly (r between .93 and .95) with the full-scale AQ. The AQ-Short has a clear factor structure, comprising two higher-order factors assessing broad difficulties in social functioning and a fascination for numbers and patterns. The Social behavior factor can be further decomposed into four lower-order factors assessing Social skills, Routine, Switching and Imagination. We suspect that Routine and Switching are tapping into Social Behaviour simply because of the cognitive demands of social interaction. Social Behaviour invariably requires rapid attentional switching between people (especially in a group and in conversation), and is invariably novel, rather than scripted or routine. The factor structure of the AQ-Short is in line with results from previous factor analytic studies (see for reviews: Mandy and Skuse 2008; Happé and Ronald 2008) in which the

majority identified at least one factor measuring autistic traits in the social domain, and one or more factors assessing non-social traits. Similar to previous studies, the correlation between the social factor (Social behavior) and the non-social factor (Numbers/patterns) was only modest (r between .16 and .20). This suggests they are considerably—but not totally—independent.

We stress that the AQ-Short is not a diagnostic instrument. For a clinical diagnosis and detailed clinical studies, established diagnostic assessments and interviews (in which different autistic behaviors including communication impairments are assessed) remain essential. The aim of this study was not to replace the full-scale AQ, but rather to develop a short version of the instrument that could be more easily implemented in large-scale studies. Large epidemiological studies often do not allow for inclusion of a 50-item measure. The AQ-Short may also be valuable as a quick screening for autistic traits in a clinical setting when filling out a 50-item questionnaire is too demanding. Our study suggests that a cut-off of >65 may be useful in this setting. When more time and resources are available, the full-scale AQ still has merits, particularly if one wants to examine communication difficulties.

A limitation of the current study is that the diagnosis of individuals in the English AS sample was based on DSM-IV or ICD-10 criteria and not on standardized diagnostic assessments. There is currently no universally agreed method for making research diagnoses in high functioning adults (the Autism Diagnostic Interview (Lord et al. 1994) and the Autism Diagnostic Observation Schedule (Lord et al. 2000) were not designed for this purpose), but this limitation means that the AS diagnoses could not be validated. A subset of the English sample obtained their diagnosis in the Cambridge clinic where self-report diagnosis was confirmed using DSM-IV criteria, and others from the English sample have been tested in person using the ADI-R and ADOS, again confirming their self-report.

This study did not include participants with other ASC diagnoses, such as PDD-NOS and classic autism. These limitations need to be kept in mind when interpreting our findings. Future studies should examine the strength of the AQ-Short in clinical settings, including different clinical (both ASC and non-ASC) groups. Women were slightly overrepresented in all four samples included in our study, and this might have affected the factor structure of the AQ-Short. Future studies could specifically examine sex differences in the factor structure of autistic traits.

In conclusion, like the full-scale AQ, the AQ-Short shows a continuous distribution, the total AQ-Short score and its two higher-order factors showed acceptable to good internal consistency in all samples under study, and sex and group differences were in the expected direction. These results suggest that the AQ-Short is a reliable instrument for a quick assessment of quantitative autistic traits. This abridged version of the AQ could be particularly useful in large scale population-based studies and in clinical settings when filling out the full 50-item version is too demanding.

Acknowledgments This research in the Dutch datasets was funded by grants from the Netherlands Organization for Scientific Research (NWO 575-25-006; NWO/SPI 56-464-14192; NWO/MaGW VIDI-016-065-318 and NWO/MaGW VENI-451-08-025). The English data collection was funded by the MRC (UK), the Three Guineas Trust, and was conducted in association with the NIHR CLAHRC for Cambridgeshire and Peterborough NHS Foundation Trust. Dr. Hoekstra is financially supported by NWO Rubicon. We gratefully acknowledge the participants of this study for their time and effort.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

- 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, 38*, 1230–1240.
- Auyeung, B., Baron-Cohen, S., Ashwin, E., Knickmeyer, R., Taylor, K., & Hackett, G. (2009). Fetal testosterone and autistic traits. *British Journal of Psychology, 100*, 1–22.
- 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.
- Baron-Cohen, S., Hoekstra, R. A., Knickmeyer, R., & Wheelwright, S. (2006). The Autism-Spectrum Quotient (AQ)—Adolescent version. *Journal of Autism and Developmental Disorders, 36*, 343–350.
- Bayliss, A. P., di Pellegrino, G., & Tipper, S. P. (2005). Sex differences in eye gaze and symbolic cueing of attention. *Quarterly Journal of Experimental Psychology A, 58*, 631–650.
- Chakrabarti, B., Dudbridge, F., Kent, L., Wheelwright, S., Hill-Cawthorne, G., Allison, C., et al. (2009). Genes related to sex steroids, neural growth, and social-emotional behavior are associated with autistic traits, empathy, and asperger syndrome. *Autism Research, 2*, 157–177.
- Constantino, J. N., & Todd, R. D. (2003). Autistic traits in the general population: A twin study. *Archives of General Psychiatry, 60*, 524–530.
- Fombonne, E. (2006). Past and future perspectives on autism epidemiology. In *Understanding autism, from basic neuroscience to treatment* (pp. 25–48). Boca Raton, FL: Taylor & Francis.
- Gomot, M., Belmonte, M. K., Bullmore, E. T., Bernard, F. A., & Baron-Cohen, S. (2008). Brain hyper-reactivity to auditory novel targets in children with high-functioning autism. *Brain, 131*, 2479–2488.
- Happé, F., & Ronald, A. (2008). The 'fractionable autism triad': a review of evidence from behavioural, genetic, cognitive and neural research. *Neuropsychology Review, 18*, 287–304.
- Hoekstra, R. A., Bartels, M., Hudziak, J. J., Van Beijsterveldt, T. C. E. M., & Boomsma, D. I. (2007a). Genetic and environmental covariation between autistic traits and behavioral problems. *Twin Research and Human Genetics, 10*, 853–860.
- Hoekstra, R. A., Bartels, M., Verweij, C. J. H., & Boomsma, D. I. (2007b). Heritability of autistic traits in the general population. *Archives of Pediatrics and Adolescent Medicine, 161*, 372–377.
- Hoekstra, R. A., Bartels, M., Cath, D. C., & Boomsma, D. I. (2008). Factor structure, reliability and criterion validity of the Autism-Spectrum Quotient (AQ): A study in Dutch population and patient groups. *Journal of Autism and Developmental Disorders, 38*, 1555–1566.
- Hurst, R. M., Mitchell, J. T., Kimbrel, N. A., Kwapil, T. K., & Nelson-Gray, R. O. (2007). Examination of the reliability and factor structure of the Autism Spectrum Quotient (AQ) in a non-clinical sample. *Personality and Individual Differences, 43*, 1938–1949.
- Lombardo, M. V., Barnes, J. L., Wheelwright, S., & Baron-Cohen, S. (2007). Self-referential cognition and empathy in autism. *PLoS One, 12*, e883.
- 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.
- 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.
- Mandy, W. P. L., & Skuse, D. H. (2008). Research review: What is the association between the social-communication element of autism and repetitive interests, behaviours and activities? *Journal of Child Psychology and Psychiatry, 49*, 795–808.
- Muthén, L. K., & Muthén, B. O. (2007). *Mplus user's guide* (5th ed.). Los Angeles, CA: Muthén & Muthén.
- Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online, 8*, 23–74.
- 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.