Investigating the structure of the Autism-Spectrum Quotient using Mokken scaling

Mary Elizabeth Stewart[1a], Carrie Allison[2], Simon Baron-Cohen[2, 3], Roger Watson[4]
[1] Psychology, Heriot-Watt University, Riccarton, Edinburgh, EH14 4AS, UK
[2] Autism Research Centre, Department of Psychiatry, Cambridge University, Douglas House,
18b Trumpington Road, Cambridge, CB2 8AH, UK.
[3] CLASS Clinic, Cambridgeshire and Peterborough NHS Foundation Trust (CPFT).
[4] Faculty of Health and Social Care, University of Hull, Hull, HU6 7RX, UK.

[a] corresponding author m.e.stewart@hw.ac.uk +44 131 45

Abstract

Traits similar to those shown in Autism Spectrum Condition (ASC) are apparent in relatives of individuals with ASC, and in the general population without necessarily meeting diagnostic criteria for an ASC. We assess whether the Autism-Spectrum Quotient (AQ), a self-report measure, has hierarchical properties using Mokken scaling. Hierarchical scales allow the presence of a latent trait to be identified by discovering whether and how many specific items form an ordered array along it. Data were collected from two groups: 1) people with ASC (n = 449 [240 males, 209] females, mean age 35.4 yrs, s.d.=12.8) and 2) university students (n = 943 [465 males, 475 females], mean age=23.0 yrs, s.d.=8.4). A single Mokken scale was obtained in the data from university students and three scales were obtained in the data from people with ASC. The scales all showed moderate Mokken scaling properties with the single scale obtained from university students showing weak invariant item ordering and two of the scales from people with ASC showing weak invariant item ordering. The AQ formed reliable Mokken scales. There was a large overlap between the scale from the university student sample and the sample with ASC, with the first scale, relating to social interaction, being almost identical. The present study confirms the utility of the AQ as a single instrument that can dimensionalize autistic traits in both university student and clinical samples of ASC, and confirms that items of the AQ are consistently ordered relative to one another.

Key words: Autism Spectrum Condition, item response theory, Mokken scaling, Autism-Spectrum Quotient

Recent estimates suggest that 1% of children in the UK are on the autism spectrum (Baird, et al., 2006; Baron-Cohen et al, 2009). Autism Spectrum Conditions (ASC) are characterised by impairments in social interaction and social communication, alongside the presence of unusually strong and narrow interests and repetitive behaviour (American Psychiatric Association, 2013). Characteristics similar to those shown in people with ASC also are sometimes seen in relatives of individuals with ASC, such as in reciprocal social interaction, pragmatic language and stereotypic behaviours (Bolton, et al., 1994; Landa, et al., 1992; Piven, et al., 1997; Baron-Cohen & Hammer, 1997). In addition, similar characteristics are found in the general population, such that individuals can report autistic traits without having or even necessarily requiring a diagnosis of ASC. Scales have been developed to quantify autistic character traits: these include the Autism-Spectrum Quotient (AQ) (adult, adolescent, and child versions; Baron-Cohen, Wheelwright, Skinner, et al., 2001; Baron-Cohen et al., 2006; Auyeung et al., 2008); the Broad Autism Phenotype Questionnaire (BAPQ) (Hurley, Losh, Parlier, Reznick, & Piven, 2007); the Social Responsiveness Scale (SRS) (Constantino, Przybeck, Friesen, & Todd, 2000); the Broad Autism Phenotype Symptom Scale (BAPSS) (Dawson, et al., 2007)); the Childhood Autism Screening Test (CAST) (Scott, Baron-Cohen, Bolton, & Brayne, 2002), the Quantitative Checklist for Autism in Toddlers (Q-CHAT) (Allison et al., 2008); and the Children's Communication Checklist (CCC) (Bishop, 1998).

The study of autistic traits in the general population may be useful in several ways. Using quantitative measures, individuals with a diagnosis of ASC can be compared to those without a diagnosis, allowing for more statistically sensitive designs that take advantage of the variability of autistic traits across individuals (Kennedy, 2009; Sung, et

al., 2005). By studying individuals with autistic traits we can gain further insight into, for instance, processing styles and language impairment across the autism spectrum (Almeida, Dickinson, Maybery, Badcock, & Badcock, 2010; Bayliss & Tipper, 2005; Stewart & Ota, 2008; Stewart, Watson, Allcock, & Yaqoob, 2009). In addition, some researchers have proposed that the impairments characterising ASC may not cluster together and should be studied separately (Happé, Ronald, & Plomin, 2006). By assessing traits we can assess, in a general population sample, which traits are most predictive of behaviour and symptoms, and which traits cluster together (Austin, 2005; Hoekstra, Bartels, Cath, & Boomsma, 2008; Stewart & Austin, 2009).

Autistic traits as measured by the AQ show high heritability (Hoekstra, Bartels, Verweij, et al., 2007), are stable cross-culturally, in Dutch, French-Canadian and Japanese samples (Hoekstra, et al., 2008; Kurita, Koyama, & Osada, 2005; Lepage, Lortie, Taschereau-Dumouchel, & Theoret, 2009; Wakabayashi, Baron-Cohen, Wheelwright, & Tojo, 2006), and are normally distributed in the population (Hurst, Mitchell, Kimbrel, Kwapil, & Nelson-Gray, 2007). Several studies have found associations between autistic traits as measured by the AQ and behavioural and cognitive measures. The AQ has shown utility as a screening tool in a clinical sample (Woodbury-Smith, Robinson, Wheelwright, & Baron-Cohen, 2005). The AQ predicts performance on cognitive tasks such as an adapted block design task (Stewart, et al., 2009) and the Embedded Figures Test (Almeida, Dickinson, Maybery, Badcock, & Badcock, 2009; Almeida, et al., 2010; Grinter, et al., 2009). Scores on the AQ are related to performance on tests of social cognition such as the ‘Reading the Mind in the Eyes’ task (Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001), gaze preference to social and non-

social stimuli (Bayliss & Tipper, 2005) and in auditory speech perception (Stewart & Ota, 2008). Reduced spontaneous facial mimicry has also been reported in high scorers (Hermans, van Wingen, Bos, Putman, & van Honk, 2009).

The AQ correlates negatively with the Empathy Quotient (Baron-Cohen & Wheelwright, 2004; Wheelwright, et al., 2006), and with scores on measures of interpersonal functioning, such as the Friendship and Relationship Quotient (Baron-Cohen & Wheelwright, 2003) and the UCLA loneliness scale (Jobe & White, 2007; Russell, 1996)). Jobe and Williams (2007) found individuals with higher AQ scores to have fewer and shorter friendships. AQ is positively correlated with length of marriage and is inversely correlated with relationship satisfaction for husbands (but not wives) (Pollmann et al, 2010). AQ is inversely correlated with left hemisphere language dominance, similar to the atypical patterns of hemispheric asymmetry characteristic of individuals with autism (Lindell, Notice, & Withers, 2009). Studies have also assessed the relationship of the AQ with other personality and clinical measures. A moderate relationship has been found between the AQ and the Big Five personality dimensions, in particular Extraversion and Neuroticism (Austin, 2005; Wakabayashi, Baron-Cohen, & Wheelwright, 2006). Scores on the AQ are related to obsessional personality scores and to higher scores on depression and anxiety scales (Kunihira, Senju, Dairoku, Wakabayashi, & Hasegawa, 2006).

Taken together, these findings suggest that AQ serves an important role in our understanding of autistic traits. The AQ has shown good test-retest reliability (Baron-Cohen, Wheelwright, Skinner, et al., 2001; Hoekstra, et al., 2008) and moderate to good

internal consistency (Hurst, et al., 2007; Kurita, et al., 2005; Stewart & Austin, 2009). However, some aspects of the AQ require further study. For instance, the structure originally proposed by Baron-Cohen et al. (2001) has not been consistently replicated. To date following factor analysis of the AQ 50 item questionnaire, one study found a two-factor model, in a Dutch sample, which included a broad factor of social interaction together with a second factor, 'Attention to Detail' (Hoekstra, et al., 2008); two studies have shown a three-factor structure of 'Social Skills', 'Details/Patterns', and 'Communication/Mind Reading' (Austin, 2005; Hurst, et al., 2007); and Stewart and Austin (2009) found a four-factor model of 'Socialness', 'Pattern', 'Understanding Others/Communication', and 'Imagination'. Although these studies do not agree on a factor structure, all of the studies agree on a "social" factor and an "attention to detail" factor.

Baron-Cohen et al.'s (2001) original conceptualisation included items relating to cognitive factors which are not diagnostic and are not included in diagnostic manuals or in other conceptualisations of autistic traits. The Diagnostic and Statistical Manual of Mental Disorders 5th Edition (DSM-V; American Psychiatric Association, 2013), propose a dyad of social and communication impairments. It is interesting and informative in light of this new conceptualisation to assess which items from the AQ carry the most weight.

Researchers have identified the utility of having a short screening scale for frontline health professionals in identifying ASCs. The AQ-Short and the AQ-10 (Hoekstra et al., 2011; Allison et al., 2012) were developed for this purpose. The AQ-10 may have particular utility in primary care settings where for instance the average

appointment time is less than 15 minutes, meaning that a questionnaire can be completed and decisions made in real time. In the case of the AQ-10, two items were taken from each domain with the greatest discriminatory power; and in the case of the AQ-Short, items were selected through a series of steps including inspection of the items, exploratory factor analysis of both the whole scale and the domains, and confirmatory factor analysis per domain and across all the factors.

In this study we assess whether the AQ has *hierarchical* properties, tested using Mokken scaling. Hierarchical scales allow the presence of a latent trait to be identified by discovering whether and how many specific items form an *ordered* array along it. In other words this analysis identifies whether the items of the AQ are consistently ordered relative to one another giving an indication of the relative position of each item on the latent trait assessed by the scale. As an illustration of this concept, if a high jumper were to successfully clear 2 metres they would not be asked to clear 1.95m or 1.9m as these heights are easier. Similarly, if, on a scale an individual endorses an item indicating a certain level of autistic traits, they are likely to also have endorsed all items indicating lower levels of the same latent trait. However, this cannot be taken for granted, and whether items fall into this hierarchy can be tested empirically. As far as we are aware there are no published studies assessing this scale using such analysis.

While the AQ was not developed with deliberate hierarchies of items and not, specifically, developed using Mokken scaling, the method has been retrospectively applied to a range of scales in psychology with some interesting results. For example, hierarchical scales have been useful in assessing constructs such as neuroticism (Watson,

Deary, & Austin, 2007), happiness (Stewart, et al., 2010), psychological distress (Watson, Deary, & Shipley, 2008) and feeding behaviour in dementia (Watson, 1996). One measure, the General Health Questionnaire (GHQ) is often used to assess psychological distress and well-being in the general population. It has several forms with 60, 32, 28 or 12 items. Watson et al. (2008), using Mokken scaling, identified nine items from the GHQ-30 that form a useful and reliable scale. This may be relevant in the development of a shorter scale to reduce the burden on participants to complete long questionnaires. Analysis such as this would give empirical evidence to the weighting of particular traits being used in other scales such as the Autism Diagnostic Observation Schedule (ADOS; Lord, et al., 1989).

In sum, this study firstly will allow us to assess which autistic traits are higher in the hierarchy for individuals in the university sample; which traits are higher in the hierarchy for individuals on the autism spectrum; and whether these traits are similar in both groups. Secondly, this study will not only inform us more about the AQ itself, but could lead to the development of an extremely useful screening tool. In addition this study adds to the body of knowledge regarding the weighting of individual traits and adds empirical evidence regarding weighting to particular traits.

METHODS

Mokken scaling is a non-parametric application of item response theory. Unlike Guttman scaling, from which Mokken scaling is derived, Mokken scaling is stochastic as it allows for a probabilistic relationship between latent traits and item scores (Sijtsma & Molenaar 2002) and can accommodate measurement error. Guttman scaling items, on

the other hand, are deterministic (Katz, 1988) and assume a perfect relationship between the relative scores on items and their relationship to the latent trait. In response to a difference in the level of the latent trait, a Guttman item scores dichotomously (for example '+ or '-). As such, for the scores on pairs of items, for example item i and item j, it is envisaged that the pattern of relative scoring on these items is always the same. If item j represents more of the latent trait than item i, then item i will always be scored '+' before item j and if item j is scored '+' then item i will also be scored '+'. Mokken scaling incorporates Guttman scaling but assumes that there is a probability distribution between the extent to which the latent trait is present and the score on an item and, likewise, in the relative scores of items. A recent comprehensive and relatively non-technical description of Mokken scaling — including invariant item ordering, discussed below — has recently been published (Watson et al. 2011) and readers are referred to this for a fuller understanding of the method and its application. Windows compatible software is available for running Mokken Scaling Analysis and the parameters generated to evaluate Mokken scales include Loevinger's coefficient (H), which is an indicator of unidimensionality in Mokken scales, and values should exceed 0.3 to indicate the presence, at least, of a weak Mokken scale (Molenaar & Sijtsma, 2000). The reliability of Mokken scales can be evaluated using a test-retest type statistic (Rho) that should exceed 0.7 to indicate a reliable scale (Sijtsma & Molenaar, 2002). The probability of obtaining a Mokken scale can be evaluated using a Bonferroni (for multiple iterations) corrected p-value (Molenaar & Sijtsma, 2000) and parameters are generated ('*Crit'*) to indicate whether item response functions (IRF) violate the model of monotone homogeneity (i.e.

that the score on the item increases progressively as the latent trait increases (Sijtsma & Molenaar, 2002).

Invariant item ordering

Provided all of the above parameters are acceptable then an important property to investigate in Mokken scales is invariant item ordering (Sijtsma & Junker, 1996, Sijtsma et al., 2011) and it has only been possible in recent years to investigate this for polytomous items. The software to enable analysis of invariant item ordering is available in the Mokken Scaling Analysis facilities in the public domain statistical software R (http://www.r-project.org/) package 'mokken' (http://cran.r-project.org/web/packages/mokken/mokken.pdf'; van der Ark 2007). Invariant item ordering refers to the non-intersection of item response functions and is a measure of the conceptual 'distance' between items. Items that are well spaced tend to show invariant item ordering and at an acceptable level. A parameter, analogous to H above called Htrans (denoted H^T), can be generated by Mokken Scaling Analysis in R and the minimum value of H^T, indicating weak invariant item ordering, is 0.3 (Ligtvoet et al., 2010).

Participants

Two groups of participants were recruited.

Group 1: 943 participants were recruited from universities. Participants were recruited as part of other ongoing projects, none of this sample reported having a diagnosis of ASC. Participants either completed and returned the questionnaires immediately or returned the

questionnaires to an investigator after completion. Participants were recruited from across universities from a range of schools and departments. Participants included both undergraduates and postgraduates. All participants were volunteers. Participants were invited to take part only if English was their first language. There were 465 males, 475 females, and three people who did not indicate their gender. 15 individuals did not give their age, the mean age of the remaining 928 participants was 23.0 years, standard deviation 8.4 years. All participants gave informed consent and all were included in the analysis. A small proportion (3.5%) omitted one or more items and these individuals were excluded from further analyses, leaving a sample total of 910.

Group 2: 449 participants with ASC were recruited. 402 were diagnosed with Asperger Syndrome and 47 with High Functioning Autism. There were 209 females, and 240 males. The mean age of the group was 35.4 years, standard deviation 12.8. They were recruited via an online portal through a research centre and all had a diagnosis of ASC from an experienced professional using DSM-IV or ICD-10 criteria. All the participants were included in the analysis.

Ethical approval was given prospectively by the local University Ethics Boards.

Materials

Autism-Spectrum Quotient (AQ: Baron-Cohen, Wheelwright, Skinner, et al., 2001): The Autism-Spectrum Quotient is a self-administered questionnaire comprised of 50 items. It consists of five subscales, each containing 10 questions assessing: Social Skills, Communication, Imagination, Attention to Detail and Attention-Switching. Half the questions are worded to elicit an ‘agree’ response and the other half, a ‘disagree’ response. The test was administered as a pen-and-paper task.

Participants were asked to answer each question as quickly as possible by circling their response on a 4-point scale ('strongly disagree', 'disagree', 'agree', 'strongly agree'). The items were scored on a continuous (Likert) scale (1-4) as this retains more information about the participants' responses than the original 0/1 scoring method (e.g. Baron-Cohen, Wheelwright, Skinner, et al., 2001). Use of all the response option choice information also increases the inter-item correlations, scale reliability and validity coefficients (Muniz, Garcia-Cueto, & Lozano, 2005). A total AQ score is calculated by summing scores for each item, with a maximum score of 200. This scoring method has been used previously (Stewart & Ota, 2008; Stewart et al., 2009).

Procedure

Data were entered into SPSS for descriptive analysis and then converted into formats suitable for analysis using Mokken Scaling Analysis in Windows and in R. In both sets of data a search for scales was initiated starting at H=0.05 and then through increments of 0.05 up to H=0.50 to test for the existence of multiple dimensions in the data. For both sets of data the search setting of H=0.30 in Mokken Scaling Analysis for Windows was used to extract Mokken scales. The scales initially obtained were checked for violations of the model of monotone homogeneity and violating items were removed on the basis of *Crit* values > 40 as recommended by Molenaar & Sijtsma (2000). Using the recently described method (Kuikpers, van der Ark & Croon, 2013) the 95% confidence intervals were calculated for H for item pairs, items and the scale. Where 95% confidence intervals for scale and item H include the lowerbound level for a weak scale (0.30) this is reported and for item pairs, the 95% confidence intervals should not

include 0. The resulting scales were entered into R and checked for invariant item ordering using Method Manifest invariant item ordering.

RESULTS

Group 1: The results of the Mokken Scaling Analysis are shown in Table 1. A moderately strong Mokken scale (H>0.40) was obtained which was reliable (Rho>0.70) and statistically significant (p<0.001). The scale included 10 items ranging from 'New situations make me anxious' (mean = 2.58), items relating to difficulty in communicating with others, through to social skills such as 'I would rather go to a library than a party' (mean = 1.53). Taking reverse scoring into account such that high scores on these items indicate a greater level of the latent trait (i.e. items that are not reverse scored, e.g. 'I am good at social chit-chat' should be seen as indicating that the respondent does not enjoy social chit-chat) then the items are arranged such that those with a higher mean score (and thereby a greater level of social inhibition) are more readily endorsed than items with a lower mean score. Therefore, the AQ Mokken scaled items are arranged in a hierarchy from the least level of difficulty, one of being anxious in social situations, to one where the respondent would avoid social situations. In between, the arrangement of items is entirely sensible indicating a greater level of social inhibition as the respondents move from situations where they find it difficult to communicate to ones where they really do not seek or enjoy social situations. The AQ Mokken scale shows an acceptable, but weak, level of invariant item ordering at $H^T = 0.32$. For items 13 and 15 the 95% confidence intervals for H included 0.30; none of the item pair H included 0.

No items were included from the Attention to Detail or the Imagination scale and only one from the Attention Switching scale. The majority of the items were made up of Social Skill and Communication items.

Group 2: The results of the Mokken Scaling Analysis are shown in Tables 2a, 2b and 2c. Three scales were extracted, all with moderately strong Mokken scales that are reliable and statistically significant ($p<0.001$). The first, described in Table 2a, is very congruent to that identified in Group 1. Ten items were retained in the scale with nine of the ten items in common with the scale obtained from Group 1. While inclusion and ordering of items is not the same the scale runs—taking the reverse scoring of items into account—from 'I find social situations easy' (reverse scored) (mean = 3.76) to 'I enjoy doing things spontaneously' (reverse scored) (mean = 2.96) thus showing a hierarchical scale in terms of difficulty in social situations through communication difficulty ('I frequently find that I don't know how to keep a conversation going') to greater difficulty meeting people and attention switching. The main difference between Group 1 (university students) and group 2 (people with ASC) is that the mean scores are generally much higher and especially for items in common in the scale from Group 2. The scale shows invariant item ordering ($H^T = 0.24$) but not at a sufficient level of accuracy. For items 22, 26 and 34 the 95% confidence intervals for H included 0.30; none of the item pair H included 0.

Two further Mokken scales were derived from Group 2. Table 2b shows a scale with six items composed entirely of items related to imagination and the scale in Table 2c is composed entirely of four items related to attention to detail. Both the scales show acceptable but low invariant item ordering ($H^T > 0.30$). A hierarchy of items can be

envisaged in Table 2b whereby finding it ‘difficult to imagine what it would be like to be someone else’ (mean = 3.26) is easier for someone with ASC to score than—taking reverse scoring into account—not finding it ‘easy to create a picture in my mind’ (mean = 2.17) represents a greater level of the latent trait of imagination. Likewise, in Table 2c the two items that are easier to score (mean > 3.00) are concerned with noticing things while the more difficult items (mean < 3.00) are concerned with fascination. For items 23 and 42 the 95% confidence intervals for H included 0.30; none of the item pair H included 0.

DISCUSSION

In this study we tested whether hierarchical scales could be formed from the AQ in both a sample of typical adult students and a sample of individuals with ASC. The AQ was not deliberately designed with hierarchies of items and there is no reason *a priori* why hierarchies of items should be found. Nevertheless, beyond some discoveries in other psychological instruments, it is possible that some items in the AQ are more representative of the autism condition, and by endorsing those particular items other items become redundant. If this is consistent in sufficiently large groups then this can be detected using Mokken scaling; then this is inherently interesting and potentially useful. It provides further insight into the structure and functioning of the AQ and also further insight into the underlying traits.

Mokken scales were formed in both samples. In the university sample one Mokken scale was obtained from 10 items while in the sample of individuals with ASC three scales were obtained from 20 items; the remaining items in both groups were rejected on the basis that they did not fit the criteria for Mokken scaling. All four scales would be considered moderately strong (H>0.40) (Molenaar & Sijtsma, 2000), and are highly reliable. According to criteria for sample size adequacy in Mokken scaling (Straat, 2010) the sample sizes here are likely to be adequate due to the range of scale H values; the 95% confidence intervals for some items includes the lowerbound value of 0.30 but the items are included here in the scales; in future work, these could be omitted to see if this improves scale properties.

In both samples the first scale comprises ten items, nine of which overlap, although these overlapping items are not anchored in the same order. The majority of these items are related to social skills, with some items relating to communication. In the Diagnostic and Statistical Manual of Mental Disorders 5th Edition (DSM-V; American Psychiatric Association, 2013) there is an emphasis on social interaction and communication. In addition, in scoring for the Autism Diagnostic Observation Schedule (ADOS; Lord, et al., 1989), a measure which aids diagnosis) there are more social interaction items, and a greater weighting on social interaction in the diagnostic algorithm than communication; imagination/creativity are not included within the algorithm. Both Mokken scales have a dominance of social interaction items; however, communication is also important with some communication items being amongst the most difficult.

In the individuals with ASC, three latent traits were found. The first was made up of items relating to social skills. The additional two latent traits relate to Imagination and to Attention to Detail. The similarity in the first latent trait between the university sample and the sample in individuals with ASC is striking. However, two additional latent traits emerge in the clinical sample, but not in the university sample. Interestingly in the university sample, there were no items relating to the original AQ's domains of Attention to Detail or Imagination and only one item relating to Attention Switching, however, in the participants with ASC, Imagination items were included in the second Mokken scale and Attention to Detail in the third.

This raises some questions. How much do these second and third latent traits add to the characterisation of individuals with ASC? Is there something specific about this

sample, or about the measure itself, the AQ? Do these items help define ASC? It may be that the additional latent traits are more specific to ASC, and that although some aspects of ASC lie on a continuum, others are categorical and only emerge in individuals who meet diagnostic criteria for ASC. It may be that individuals with ASC respond differently to people in the general population on AQ items or that the AQ (50 items) does not accurately represent autistic traits, and that a shorter version may be more appropriate. It would be useful to include ratings by others such as friends and carers, and to include a range of autistic trait measures in order to identify if any latent traits are unique to the questionnaire. In addition, there may be something specific about this particular sample of individuals with ASC or with the university sample, which may not be representative of the general population. The participants included in this study were mainly individuals with Asperger Syndrome, and it may be that, for instance, the attention to detail latent trait, which has been identified from cognitive studies, is particular to this group and not to individuals with ASC per se. One further difference between the samples is that the individuals with ASC were older than the university student sample. Age could have an effect on response patterns on the AQ, although effects of age on autistic traits has not previously been found. Age cannot be covaried in Mokken scaling.

Nevertheless, while the items in the first latent trait are in the main consistent with diagnostic criteria, there is very little emphasis on restricted repetitive and stereotyped patterns of behaviour. The items in the second and third latent traits relate to imagination and attention to detail, which are not considered diagnostic features of ASC (although sensory hyper-sensitivity, which relates to attention to detail, is now part of the DSM-5 criteria for autism). These items do not appear to be high in the hierarchy when

measuring autistic traits, nor are they weighted in diagnostic assessment instruments, such as the ADOS. Whether this study has identified differences in the autistic trait profile between these groups still needs to be established as it remains to be tested how important the items from the Mokken scales are clinically and how important they are in predicting behaviour. There may be differences between the groups in traits but these may not be related to diagnostic criteria, for instance, a lack of relationship has been found between cognitive features which are present in individuals with ASC and indices of autistic symptomatology (Pellicano et al., 2006; Teunisse et al., 2001). Given that mainly Social Skills and Communication items were included in the Mokken Scale for the university sample, this raises questions of construct underrepresentation concerning the excluded items and whether important domains have been excluded. The analysis shows that their item response functions overlap with the other items, and that they do not add to the model, however, it remains to be tested whether this will influence the utility of the shortened scale either as a screening tool or as a predictive measure of behaviour.

The 'value-added' nature of Mokken scaling is demonstrated in the present study. Partly, existing knowledge about the structure of that AQ has been confirmed but new information has also been gained. The main Mokken scale related to social skills is evident in both samples and shares many items in common in both samples. In addition, the hierarchical nature of these items is demonstrated and this enables the overall score on the latent trait to be more accurately related to specific items in the scale. This is not possible using factor analysis, where a score could in theory be composed of any set of items. A difference, as yet to be explained fully, is observed between the two samples in terms of the sets of items that are extracted into Mokken scales with two additional scales

being observed in the sample of peoples with ASC. In both cases a hierarchy of items was evident. These items may be useful in developing shorter scales which may have greater utility as they are less time consuming (Allison, Auyeung, & Baron-Cohen, 2012). The Mokken scales must go through rigorous validation studies to test whether they are related to clinical variables and whether they are predictive of autistic behaviours. Short scales which are self-report would be of great clinical utility if they are indeed useful as screening tools. We therefore recommend that studies are completed to validate the Mokken scales found in this study, and to test whether the Mokken scales show sensitivity and specificity in screening for ASC.

Assessing the hierarchical structure of the autism spectrum may help inform whether a dimensional approach adds more utility to assessing change and development of autistic traits and characteristics across time, over and above a categorical approach (Russo, Levine, Demjaha et al., 2014). Little is known about lifespan changes in autistic traits or autistic characteristics, and whether throughout development particular characteristics or traits are more or less prevalent or more or less severe in an older group than a younger one. Nor is anything known about interaction between traits or characteristics across the lifespan. This hierarchical approach may help inform regarding developmental changes in the weighting of particular traits or characteristics. It would therefore be interesting to assess whether these Mokken scales hold across the lifespan.

The current study has limitations in the ability to generalise beyond the recruited samples. The university sample all had English as their first language and is in the main comprised of UK nationals. No data were recorded regarding ethnicity. In order to test

whether the findings can be generalised to the general population a community sample would need to be recruited. The ASC sample is limited in that the majority were diagnosed with Asperger Syndrome (AS), therefore the study would need replication in a group which is more representative of ASC. However, given that in this case we were testing a self-report instrument rather than another-report instrument it would only be applicable to those with High Functioning Autism. In addition, DSM-V makes no distinction between High Functioning Autism and Asperger Syndrome.

This study shows that the AQ has hierarchical properties both in a general population student sample and in individuals with ASC. There are some differences in the derived scales between the two groups; however, it remains to be tested what these differences in the latent traits found are due to. The AQ is known to be a useful screening measure (Woodbury-Smith et al., 2005). This study raises the question of whether we can reduce this screening questionnaire down to 10 items, relating to social interaction and communication. It is perhaps intuitive that there are some behaviours which can be labelled as being very characteristic of an individual with ASC. For both those in the typically developing population and individuals with ASC, we can conclude that a resistance to being around other people, chatting with them and enjoying social situations are especially important indicators of autistic traits. The present study confirms the utility of the AQ as a single instrument that can dimensionalize autistic traits in both the general population and clinical samples of ASC, and identifies that the items of the AQ are consistently ordered relative to one another, giving an indication of the relative position of each item on the latent trait assessed by the scale.

References

Allison, C., Auyeung, B., & Baron-Cohen, S. (2012). Toward Brief "Red Flags" for Autism Screening: The Short Autism Spectrum Quotient and the Short Quantitative Checklist in 1,000 Cases and 3,000 Controls. *Journal of the American Academy of Child & Adolescent Psychiatry, 51*, 202-212.

Allison, C., Baron-Cohen, S., Wheelwright, S., Charman, T., Richler, J., Pasco, G., et al. (2008). The Q-CHAT (Quantitative CHecklist for Autism in Toddlers): A Normally Distributed Quantitative Measure of Autistic Traits at 18-24 Months of Age: Preliminary Report. *Journal of Autism and Developmental Disorders 38*, 1414-1425.

Almeida, R. A., Dickinson, J. E., Maybery, M. T., Badcock, J. C., & Badcock, D. R. (2009). A new step towards understanding embedded figures test performance in the autism spectrum: The radial frequency search task. *Neuropsychologia*, *48*, 374-381.

Almeida, R. A., Dickinson, J. E., Maybery, M. T., Badcock, J. C., & Badcock, D. R. (2010). Visual search performance in the autism spectrum II: the radial frequency search task with additional segmentation cues. *Neuropsychologia, 48*, 4117-4124.

American Psychiatric Association (2013). *Diagnostic and Statisitical Manual of Mental Disorders* (5th edition ed.). Washington D.C.

American Psychiatric Association (1994). *Diagnostic and Statisitical Manual of Mental Disorders* (4th edition ed.). Washington D.C.

Austin, E. J. (2005). Personality correlates of the broader autism phenotype as assessed by the Autism Spectrum Quotient (AQ). *Personality and Individual Differences, 38*(2), 451-460.

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.

Baird, G., Simonoff, E., Pickles, A., Chandler, S., Loucas, T., Meldrum, D., et al. (2006). Prevalence of disorders of the autism spectrum in a population cohort of children in South Thames: the Special Needs and Autism Project (SNAP). *Lancet, 368*(9531), 210-215.

Baron-Cohen, S., & Hammer, J. (1997). Parents of children with Asperger Syndrome: what is the cognitive phenotype? *Journal of Cognitive Neuroscience, 9*, 548-554.

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*(3), 343-350.

Baron-Cohen, S., Scott, F. J., Allison, C., Williams, J., Bolton, P., Matthews, F. E., et al. (2009). Prevalence of autism-spectrum conditions: UK school-based population study. *British Journal of Psychiatry, 194*(6), 500-509.

Baron-Cohen, S., & Wheelwright, S. (2003). The Friendship Questionnaire: An investigation of adults with Asperger syndrome or high-functioning autism, and normal sex differences. *Journal of Autism and Developmental Disorders, 33*(5), 509-517.

Baron-Cohen, S., & Wheelwright, S. (2004). The Empathy Quotient: An Investigation of Adults with Asperger Syndrome or High Functioning Autism, and Normal Sex Differences. *Journal of Autism and Developmental Disorders, 34*(2), 163-175.

Baron-Cohen, S., Wheelwright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 42*(2), 241-251.

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*(1), 5-17.

Bayliss, A. P., & Tipper, S. P. (2005). Gaze and arrow cueing of attention reveals individual differences along the autism spectrum as a function of target context. *British Journal of Psychology, 96*(Pt 1), 95-114.

Bishop, D. V. M. (1998). Development of the Children's Communication Checklist (CCC): A method for assessing qualitative aspects of communicative impairment in children. *Journal of Child Psychology and Psychiatry, 39*(6), 879-891.

Bolton, P., Macdonald, H., Pickles, A., & Rios, P. (1994). A case-control family history study of autism. *Journal of Child Psychology and Psychiatry, 35*(5), 877-900.

Constantino, J. N., Przybeck, T., Friesen, D., & Todd, R. D. (2000). Reciprocal social behavior in children with and without pervasive developmental disorders. *Journal of Developmental and Behavioral Pediatrics, 21*(1), 2-11.

Dawson, G., Estes, A., Munson, J., Schellenberg, G., Bernier, R., & Abbott, R. (2007). Quantitative assessment of autism symptom-related traits in probands and parents: Broader phenotype autism symptom scale. *Journal of Autism and Developmental Disorders, 37*(3), 523-536.

Grinter, E. J., Maybery, M. T., Van Beek, P. L., Pellicano, E., Badcock, J. C., & Badcock, D. R. (2009). Global visual processing and self-rated autistic-like traits. *Journal of Autism and Developmental Disorders, 39*(9), 1278-1290.

Happé, F., Ronald, A., & Plomin, R. (2006). Time to give up on a single explanation for autism. *Nature Neuroscience, 9*(10), 1218-1220.

Hermans, E. J., van Wingen, G., Bos, P. A., Putman, P., & van Honk, J. (2009). Reduced spontaneous facial mimicry in women with autistic traits. *Biological Psychology, 80*(3), 348-353.

Hoekstra, R. A., Bartels, M., Cath, D. C., & Boomsma, D. I. (2008). Factor structure of the broader autism phenotype: a study using the Dutch translation of the autism-spectrum quotient (AQ). *Journal of Autism and Developmental Disorders, 38*, 1555-1566.

Hoekstra, R. A., Bartels, M., Verweij, C. J., & Boomsma, D. I. (2007). Heritability of autistic traits in the general population. *Archives of Pediatrics & Adolescent Medicine, 161*(4), 372-377.

Hoekstra, R. A., Vinkhuyzen, A. A. E.; Wheelwright, S., Bartels, M., Boomsma, D. I., Baron-Cohen, S., Posthuma, D., & van der Sluis, (2011). The Construction and

Validation of an Abridged Version of the Autism-Spectrum Quotient (AQ-Short). *Journal of Autism and Developmental Disorders, 41*, 589-596.

Hurley, R. S. E., Losh, M., Parlier, M., Reznick, J. S., & Piven, J. (2007). The broad autism phenotype questionnaire. *Journal of Autism and Developmental Disorders, 37*, 1679-1690.

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

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

Katz, E. (1988) In memoriam: Louis Guttman, 1916–1987. *Public Opinion Quarterly* 52, 40–242.

Kennedy, D. P. (2009). Neural correlates of autistic traits in the general population: Insights into autism. *The American Journal of Psychiatry, 166*(8), 849-857.

Kuijpers, R.E., van der Ark, L.A., Croon, M.A. (2013). Standard errors and confidence intervals for scalability coefficients in Mokken scaling analysis using marginal models. *Sociological Methodology* 43, 42-69.

Kunihira, Y., Senju, A., Dairoku, H., Wakabayashi, A., & Hasegawa, T. (2006). "Autistic" Traits in Non-Autistic Japanese Populations: Relationships with

Personality Traits and Cognitive Ability. *Journal of Autism and Developmental Disorders, 36*(4), 553-566.

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 and Clinical Neurosciences, 59*(4), 490-496.

Landa, R., Piven, J., Wzorek, M. M., & Gayle, J. O. (1992). Social language use in parents of autistic individuals. *Psychological Medicine: A Journal of Research in Psychiatry and the Allied Sciences, 22*(1), 245-254.

Lepage, J.-F., Lortie, M., Taschereau-Dumouchel, V., & Theoret, H. (2009). Validation of French-Canadian versions of the Empathy Quotient and Autism Spectrum Quotient. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement, 41*(4), 272-276.

Ligtvoet, R., van der Ark, L.A., te Marvelde, J.M. & Sijtsma, K. (2010) Investigating an Invariant Item Ordering for Polytomously Scored Items. *Educational and Psychological Measurement, 70(4),* 578–595

Lindell, A. K., Notice, K., & Withers, K. (2009). Reduced language processing asymmetry in non-autistic individuals with high levels of autism traits. *Laterality: Asymmetries of Body, Brain and Cognition, 14*(5), 457-472.

Lord, C., Rutter, M., Goode, S., Heemsbergen, J., Jordam, H., Mawhood, L., et al. (1989). Autism Diagnostic Observation Schedule: A standardised observation of

communicative and social behaviour. *Journal of Autism and Developmental Disorders, 19*, 185-212.

Molenaar, I. W., & Sijtsma, K. (2000). MSP5 for Windows. Groningen: iec ProGAMMA.

Muniz, J., Garcia-Cueto, E., & Lozano, L. M. (2005). Item format and the psychometric properties of the Eysenck Personality Questionnaire. *Personality and Individual Differences, 38*(1), 61-69.

Pellicano, E., Maybery, M., Durkin, K., & Maley, A. (2006). Multiple cognitive capabilities/deficits in children with an autism spectrum disorder: "Weak" central coherence and its relationship to theory of mind and executive control. *Development and Psychopathology, 18, 77–98.*

Piven, J., Palmer, P., Jacobi, D., & Childress, D. (1997). Broader autism phenotype: Evidence from a family history study of multiple-incidence autism families. *The American Journal of Psychiatry, 154*(2), 185-190.

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

Russell, D. W. (1996). UCLA Loneliness Scale (Version 3): Reliability, validity, and factor structure. *Journal of Personality Assessment, 66*(1), 20-40.

Russo, M. Levine, S.Z., Demjaha, A., Di Forti, M., Bonaccorso, S., Fearon, P., Dazzan, P., Pariante, C.M., David, A.S., Morgan, C., Murray, R.M., & Reichenberg, A. (2014) Association Between Symptom Dimensions and Categorical Diagnoses of

Psychosis: A Cross-sectional and Longitudinal Investigation. *Schizophrenia Bulletin, 40*, 111-119.

Scott, F. J., Baron-Cohen, S., Bolton, P., & Brayne, C. (2002). The CAST (Childhood Asperger Syndrome Test): Preliminary development of a UK screen for mainstream primary-school-age children. *Autism, 6*(1), 9-31.

Sijtsma K & Junker BW (1996) A survey of theory and methods of invariant item ordering. *British Journal of Mathematical and Statistical Psychology, 49*, 79–105.

Sijtsma K, Meijer RR & van der Ark LA (2011) Mokken scale analysis as time goes by: an update for scaling practitioners. *Personality and Individual Differences, 50*, 31–37.

Sijtsma K & Molenaar IW (2002) *An Introduction to Nonparametric Item Response Theory*. Sage Publications, Thousand Oaks.

Stewart, M. E., & Austin, E. J. (2009). The structure of the Autism-Spectrum Quotient (AQ): Evidence from a student sample in Scotland. *Personality and Individual Differences, 47*(3), 224-228.

Stewart, M. E., & Ota, M. (2008). Lexical effects on speech perception in individuals with "autistic" traits. *Cognition, 109*(1), 157-162.

Stewart, M. E., Watson, J., Allcock, A.-J., & Yaqoob, T. (2009). Autistic traits predict performance on the block design. *Autism: The international journal of research and practice, 13*(2), 133-142.

Stewart, M. E., Watson, R., Clark, A., Ebmeier, K., & Deary, I. J. (2010). A hierarchy of happiness? Mokken scaling analysis of the Oxford Happiness Inventory. *Personality and Individual Differences, 48*(7), 845-848.

Straat, J.H. (2012). *Using Scalability Coefficients and Conditional Association to Assess Monotone Homogeneity* Ridderprint BV, Ridderkerk.

Sung, Y. J., Dawson, G., Munson, J., Estes, A., Schellenberg, G. D., & Wijsman, E. M. (2005). Genetic investigation of quantitative traits related to autism: Use of multivariate polygenic models with ascertainment adjustment. *American Journal of Human Genetics, 76*(1), 68-81.

Teunisse, J., Cools, A. R., van Spaendonck, K. P. M., Aerts, F. H. T. M., & Berger, H. J. C. (2001). Cognitive styles in high-functioning adolescents with autistic disorder. *Journal of Autism and Developmental Disorders, 31*, 55– 66.

van der Ark LA (2007) Mokken scale analysis in R. *Journal of Statistical Software* 20, 1–19.

Wakabayashi, A., Baron-Cohen, S., & Wheelwright, S. (2006). Are autistic traits an independent personality dimension? A study of the Autism-Spectrum Quotient (AQ) and the NEO-PI-R. *Personality and Individual Differences, 41*(5), 873-883.

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

Watson, R. (1996). Mokken scaling procedure (MSP) applied to feeding difficulty in elderly people with dementia. *International Journal of Nursing Studies, 33,* 385–393.

Watson, R., Deary, I., & Austin, E. (2007). Are personality trait items reliably more or less 'difficult'? Mokken scaling of the NEO-FFI. *Personality and Individual Differences, 43*, 1460–1469.

Watson, R., Deary, I.J., & Shipley, B (2008) A hierarchy of distress: Mokken scaling of the GHQ-30 *Psychological Medicine*, 38, 575–579.

Watson R, van der Ark LA, Lin L-C, Fieo R, Deary IJ, Meijer, R.R. (2011) Item response theory: How Mokken scaling can be used in clinical practice. *Journal of Clinical Nurs*ing (doi: 10.1111/j.1365-2702.2011.03892.x)

Wheelwright, S., Baron-Cohen, S., Goldenfeld, N., Delaney, J., Fine, D., Smith, R., et al. (2006). Predicting Autism Spectrum Quotient (AQ) from the Systemizing Quotient-Revised (SQ-R) and Empathy Quotient (EQ). *Brain Research, 1079*(1), 47-56.

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 & Developmental Disorders, 35*(3), 331-335.

Table 1:Mokken scaling of the Autism-Spectrum Quotient (n=910) in university students

Item	Mean	H(SE)	Label	Factor
Item13	1.53	0.33(0.029)	I would rather go to a library than a party✝	SS
Item44	1.64	0.41(0.028)	I enjoy social occasions*†	SS
Item47	1.69	0.49(0.022)	I enjoy meeting new people*†	SS
Item17	1.80	0.46(0.022)	I enjoy social chit-chat*†	C
Item22	1.89	0.48(0.020)	I find it hard to make new friends†	SS
Item15	1.95	0.32(0.026)	I find myself drawn more strongly to people than to things*✝	SS
Item11	2.02	0.55(0.017)	I find social situations easy*†	SS
Item38	2.07	0.56(0.017)	I am good at social chit-chat*	C
Item26	2.17	0.44(0.022)	I frequently find that I don't know how to keep a conversation going†	C
Item46	2.58	0.35(0.024)	New situations make me anxious†	AS

H(SE)=0.44(0.017); Rho=0.86; p<0.001; H^T=0.32; *=reverse scored items; †=items showing item ordering; ✝=items where the 95% CI includes 0.30

C: Communication; AS: Attention-Switching; SS: Social Skill

NB: high scores indicate > Autism-Spectrum Quotient

Table 2a: Mokken scaling of the Autism-Spectrum Quotient (n=449) in people with Autism Spectrum Condition

Item	Mean	H(SE)	Label	Factor
Item34	2.96	0.36(0.036)	I enjoy doing things spontaneously*† ✝	AS
Item47	3.12	0.52(0.028)	I enjoy meeting new people*†	SS
Item44	3.31	0.55(0.028)	I enjoy social occasions*†	SS
Item15	3.37	0.44(0.035)	I find myself drawn more strongly to people than to things*†	SS
Item13	3.45	0.41(0.038)	I would rather go to a library than a party†	SS
Item26	3.50	0.33(0.040)	I frequently find that I don't know how to keep a conversation going†✝	C
Item17	3.55	0.51(0.033	I enjoy social chit-chat*	C
Item22	3.58	0.37(0.040)	I find it hard to make new friends†✝	SS
Item38	3.67	0.49(0.044)	I am good at social chit-chat*†	SS
Item11	3.76	0.52(0.040)	I find social situations easy*†	SS

H(SE)=0.45(0.028); Rho=0.87; p<0.001; H^T=0.24; *=reverse scored items; †=items showing invariant item ordering; ✝=items where the 95% CI includes 0.30

AS: Attention-Switching; C: Communication; SS: Social Skill

NB: high scores indicate > Autism-Spectrum Quotient

Table 2b: Mokken scaling of the Autism-Spectrum Quotient (n=449) in people with Autism Spectrum Condition

Item	Mean	H(SE)	Label	Factor
Item3	2.17	0.44(0. 031)	If I try to imagine something, I find it very easy to create a picture in my mind*†	I
Item8	2.58	0.49(0.026)	When I'm reading a story, I can easily imagine what the characters look like*†	I
Item14	2.73	0.46(0.029)	I find making up stories easy*†	I
Item40	3.08	0.41(0.033)	When I was young, I used to enjoy playing games involving pretending with other children*†	I
Item50	3.18	0.43(0.033)	I find it very easy to play games with children that involve pretending*†	I
Item42	3.26	0.31(0.039)	I find it difficult to imagine what it would be like to be someone else†✝	I

H(SE)=0.43(0.025); Rho=0.79; p<0.001; H^T=0.32; *=reverse scored items; †=items showing invariant item ordering; ✝=items where the 95% CI includes 0.30

I: Imagination

NB: high scores indicate > Autism-Spectrum Quotient

Table 2c: Mokken scaling of the Autism-Spectrum Quotient (n=449) in people with Autism Spectrum Condition

Item	Mean	H(SE)	Label	Factor
Item9	2.65	0.44(0.034)	I am fascinated by dates†	AD
Item19	2.83	0.49(0.030)	I am fascinated by numbers†	AD
Item6	3.25	0.41(0.037)	I usually notice car number plates or similar strings of information†	AD
Item23	3.51	0.32(0.043)	I notice patterns in things all the time†✝	AD

H(SE)=0.42(0.03); Rho=0.70; p<0.001; H^T=0.38; *=reverse scored items; †=items showing invariant item ordering; ✝=items where the 95% CI includes 0.30

AD: Attention to Detail

NB: high scores indicate > Autism-Spectrum Quotient