J Autism Dev Disord (2014) 44:1565–1576
DOI 10.1007/s10803-013-2024-3

ORIGINAL PAPER

Psychometric Properties of the Mandarin Version of the Childhood Autism Spectrum Test (CAST): An Exploratory Study

Xiang Sun · Carrie Allison · Bonnie Auyeung · Fiona E. Matthews · Samuel Norton · Simon Baron-Cohen · Carol Brayne

Published online: 13 February 2014

Abstract Limited studies have investigated the latent autistic traits in the mainland Chinese population for autism spectrum conditions (ASC). This study explored the psychometric properties of a Mandarin Chinese version of the CAST in a sample consisting of 737 children in mainstream schools and 50 autistic cases. A combination of categorical data factor analysis and item response theory suggested a good-fit model of a two-factor solution for 28 items on the Mandarin CAST including social and communication, and inflexible/stereotyped language and behaviours (Goodness-of-fit indices: RMSEA $= 0.029$, CFI $= 0.957$, TLI $= 0.950$, SRMR $= 0.064$). The correlation between the two factors was moderate (GFC $= 0.425$). This study provided evidence for the CAST as a multidimensional measure for ASC screening in a Chinese population and also showed that the symptom manifestation of ASC in Chinese children shares similarity with western populations.

Keywords Autism spectrum conditions · Categorical data factor analysis · Item response theory · CAST · China

X. Sun (✉) · C. Brayne
Department of Public Health and Primary Care, Institute of Public Health, University of Cambridge, Forvie Site, Robinson Way, Cambridge CB2 0SR, UK
e-mail: xs227@medschl.cam.ac.uk

X. Sun · C. Allison · B. Auyeung · S. Baron-Cohen
Department of Psychiatry, Autism Research Centre, University of Cambridge, Douglas House, 18b Trumpington Road, Cambridge CB2 2AH, UK

X. Sun
Jockey Club School of Public Health and Primary Care, The Chinese University of Hong Kong, Prince of Wales Hospital, Shatin, New Territories, Hong Kong

B. Auyeung
Department of Psychology, University of Edinburgh, Edinburgh EH8 9JZ, UK

F. E. Matthews
MRC Biostatistics Unit, Cambridge Institute of Public Health, Forvie Site, Robinson Way, Cambridge CB2 0SR, UK

S. Norton
Psychology Department, Institute of Psychiatry, Kings College London, Bermondsey Wing, Guy's Hospital Campus, London Bridge, London SE1 9RT, UK

Introduction

Autism spectrum conditions (ASC) are characterised by impairments in social interaction and communication, alongside repetitive/restricted interests, behaviour and activity (American Psychiatric Association 2000). Defined by *The International Classification of Disease, 10th revision* (ICD-10) (World Health Organisation 1993), ASC include four subgroups: childhood autism, atypical autism, Asperger's syndrome and pervasive developmental disorder, unspecified. Due to the lack of biomarkers, the diagnosis of ASC still depends on behavioural descriptions (Magyar and Pandolfi 2007). There has been an emerging consensus that ASC are dimensional disorders representing the upper extremes of one or more quantitative traits, and the traits are likely to be continuously distributed in the population (Constantino and Todd 2003; Mandy and Skuse 2008; Spiker et al. 2002). Early studies for this proposed dimensional concept investigated the association between genetic liability for ASC and the milder, non-psychopathological features within the relatives of individuals with ASC (Bolton et al. 1994). Further studies suggested that autistic traits are continuously distributed in the general population (Constantino and Todd 2003; Posserud et al.

2006). Factor analysis has been used to explore the possible dimensions of ASC, with data being collected using screening and diagnostic instruments in different samples (Bolte et al. 2011; Happe and Ronald 2008).

A single autism dimension was proposed based on findings from a factor analysis of the psychometric structure of the autism diagnostic interview-revised (ADI-R) and the Social Responsiveness Scale (SRS) in both clinical and population samples (Constantino et al. 2004; Constantino and Todd 2003). Other factor analysis studies suggested that the impairments of ASC might be multidimensional (Lecavalier et al. 2006). Three studies using data from the ADI-R proposed a three-factor solution for the structure of ASC. Although the structure of these proposed factors differed among previous studies, these factors roughly covered the three core domains of impairments in ASC including Social Interaction, Communication, Repetitive and Stereotyped Behaviour and Interests (Georgiades et al. 2007; Lecavalier et al. 2006; Van Lang et al. 2006). Another two studies based on the ADI-R suggested that a two-factor solution was sufficient to explain the autism continuum, including both the social/communication and restricted/stereotyped behaviours subgroups (Frazier et al. 2008; Snow et al. 2009). A recent study using the SRS suggested two symptom dimensions including a social communication/interaction dimension, and a restricted/repetitive behaviours dimension in a sample of ASC and non-ASC siblings (Frazier et al. 2012).

The lack of consistency in the number and structure of factors identified by previous studies could be partly explained by the heterogeneous research designs (Kuenssberg et al. 2011; Mandy and Skuse 2008). A recent review on factor analysis studies for ASC indicated that the interpretation of psychometric properties should take the heterogeneity within these studies into consideration, including the variety of sample characteristics (individuals with diagnosis of AD or ASC or typically developed general population), the nature of the assessment instruments (content of items, rating scale), method of statistical analysis and by the subjective interpretations used when naming the identified factors in factor analysis (Kuenssberg et al. 2011; Mandy and Skuse 2008; Shuster et al. 2013). Although there has been no agreement on the number or structure of latent traits for ASC, exploring the factor structure can help with evaluating screening and diagnostic instruments for ASC. This process could indicate which possible dimensions of ASC have been reflected by the items on the instrument. Such research has been conducted by using several screening instruments for ASC. A summary of factor analytic studies of ASC screening instruments is given in Appendix.

Most investigations of the psychometric properties of screening instruments for ASC have been conducted among Western populations. Only two studies have investigated the psychometric properties of screening instruments for ASC in the Chinese population. Both studies were based on samples recruited in Taiwan. The Social Communication Questionnaire (SCQ) has been translated into Traditional Chinese and applied to the Chinese population in Taiwan (Gau et al. 2011). This study applied the Chinese SCQ to 317 parents of participants recruited from clinical settings. It proposed a three-factor solution for the Chinese SCQ involving Social Interaction, Repetitive Behaviours and Communication. The psychometric properties of the Chinese version of the autism spectrum quotient (AQ) were examined in a sample of 4,192 children. This consisted of both clinic-based children with ASC and community-based children showing typical development in Taiwan (Lau et al. 2012). It proposed a 35-item, five-factor solution for the Chinese AQ. These five factors were social skills, mindreading, patterns, attention to detail and attention switching. No studies have investigated the psychometric properties of screening instruments for ASC in the Chinese population in mainland China.

The Childhood Autism Spectrum Test (CAST) is a screening instrument developed in the UK for use in detecting potential cases along the whole autism spectrum. As many children with ASC, especially those children who have subtle or milder symptoms, are usually not identified before entering primary school, the CAST is designed to be applied to primary school-aged children (4–11 years) (Kamio 2007; Williams 2003). This study conducted an exploration of potential factors in a Mandarin version of the CAST in a sample drawn from the Chinese population in both clinical settings and mainstream schools. As the CAST measures autistic traits with binary responses, a combination approach was adopted. This analytic approach used an exploratory categorical data factor analysis (CDFA) with model parameters transformed to item response theory (Khalid 2011) (IRT) model parameters to assist with interpretation (Sharp et al. 2006).

Method

Participants

The CAST was developed for the general population in the UK, where children with ASC are integrated into mainstream schools. However, in mainland China, children with ASC are turned away from mainstream schools (McCabe 2003). Most children with an existing diagnosis of ASC cannot enter mainstream schools but must enter various intervention centres (Sun et al. 2012). The sample in the present study was drawn from two samples: (1) one based in mainstream schools in Beijing, involving the parents of

737 typically developing children in school years 1–4 who did not have an ASC diagnosis; (2) the other was a clinic-based sample involving the parents of 50 children with an existing diagnosis of autism. The latter were recruited from the Beijing China Disabled Persons' Federation (BCDPF) and the Elim autism rehabilitation centre in Qingdao. The BCDPF, as a local branch of the China Disabled Persons' Federation (CDPF), takes charge of people with all kinds of disabilities in Beijing. The Elim autism rehabilitation centre is a private centre in Qingdao that provides interventions for children from various regions across mainland China aged 3–6 years with ASC. All the students in mainstream schools did not have a diagnosis of ASC before this study. All the cases recruited from clinical settings had an existing diagnosis of ASC made by Chinese clinicians using the DSM-IV or ICD-10.

The Mandarin CAST

The CAST is a 37-item parental completion questionnaire, of which 31 items can be scored. Each item was measured using a dichotomized approach (yes/no scores 0/1) with the total score ranging from 0 to 31. The items were designed to cover all three domains of the impairments of ASC defined by DSM-IV-TR (Williams et al. 2005). The higher the score on the CAST the higher the probability of having more autistic features (Scott et al. 2002a). The CAST has demonstrated good validity (sensitivity = 100 %, specificity = 97 %) as a screening instrument in Western populations (note: not all the screen negatives were assessed) (Williams et al. 2005). Using the UK as a validation sample, an exploratory factor analysis of the CAST identified four factors: social behaviour and routines, speech and communication, peer relationships, and imaginative play (Williams 2003). One study has adopted the CAST to assess autistic traits which was conducted in the UK and reported moderate to high heritability for autistic traits in the general population (Robinson et al. 2011). Used as a screening instrument, the recommended cut-off is 15 ($\geq$15) for indicating a child is at risk for ASC (Scott et al. 2002b). The CAST was translated from English to Mandarin by the first author, a native Chinese speaker. It was back-translated by two Chinese–English bilingual speakers who are not involved in autism research. The Mandarin CAST was initially piloted in an opportunistic sample with ten Chinese parents who had children aged 5–10 years; they were selected from the outpatients in the Paediatric Department of the Peking University First Hospital (PUFH). The final version was then back-translated and approved by the UK authors. Each item on the Mandarin CAST also requires a binary (yes/no) response. The 6 items that did not contribute to the score (items 3, 4, 12, 22, 26, 33) were removed from the dataset before analysis.

Procedures

Parents of children in both the mainstream schools and clinical settings were informed of the purpose and procedures of this study by their institutions and then invited to participate in this study. Each parent was sent a screening pack which contained an information sheet, a Mandarin CAST, a consent form and an envelope to return the questionnaire. Ethical approval for this research was obtained from the Ethics Committees of the research institutions.

Data Analysis

The Mandarin CAST was distributed to all children in the two samples for screening purposes. A number of 34 questionnaires were completely missing as they were not returned back and another 59 questionnaires had missing items. Fifty-three (7.4 %) had one or two missing items and six (0.8 %) had three to seven items missing. Item 25 had the most missing data (ten were missing). In total, 694 fully completed questionnaires were available for analysis, of which 655 came from the mainstream sample and 39 from the clinical sample. Since only completed CAST questionnaires were used in the analyses, statistical methods were used to examine the differences in the characteristics between children who were included and those excluded. Unpaired t tests and one-way ANOVA were used to compare means, and Chi square test was used to examine differences in proportions. Whenever the numbers were small, a Fisher's exact test was used.

Traditional factor analytical methods use observed data which are continuous to explore an underlying continuous latent variable (Bartholomew et al. 2002). Factor analysis provides a linear factor combination or principal components for observed scores from tests based on continuous data (Kline 2000; Williams 2003). However, these methods are not appropriate for tests with binary item responses (yes/no), such as the Mandarin CAST, due to the fact that factor analysis assumes item responses are on a continuous metric (Angold et al. 2002; Sharp et al. 2006). If categorical data are analysed using methods for continuous data in factor analysis, the true factor structure may be distorted in a multi-dimensional analysis and the factor loading may be biased in uni-dimensional models (Muthen 1989; Sharp et al. 2006). When applying linear models to binary data, the predictions generated may not lie within a plausible range (<0 or >1) (McDonald 1999; Sharp et al. 2006).

In this study, the latent structure of the Mandarin CAST was examined using exploratory CDFA. The analysis was conducted using MPlus 6.0 (Muthen and Muthen 2010). Due to the fact that the item responses

were binary, the correlation matrix used in the CDFA was a tetrachoric correlation. The factor structure of the Mandarin CAST was estimated by a robust weighted least square estimator (WLSMV) (Muthen and Muthen 2010), which has been show to perform well in similarly sized samples (Flora and Curran 2004; Muthén et al. 1997).

The number of potential factors was determined using three approaches: (1) extraction of as many factors with eigenvalues >1 (the Kaiser criterion) (Bandalos and Boehm-Kaufman 2009); (2) extraction of as many factors that fall below the 'elbow' of a scree plot, and (3) the best fitting model as indicated by standard CDFA fit indices. These indices included the Chi square Test of Model Fit, the root mean square error of approximation (RMSEA), the comparative fit index (CFI) and the Tucker–Lewis index (TLI). Values of indices for indicating a good fit model were RMSEA $\leq$ 0.06, CFI $\geq$ 0.95, TLI $\geq$ 0.95 and SRMR $\leq$ 0.08 (Browne and Cudeck 1992; Hu and Bentler 1999). The criteria for indices of an adequate fit model were RMSEA $\leq$ 0.08, CFI $\geq$ 0.90, TLI $\geq$ 0.90 and SRMR $\leq$ 0.10 (Browne and Cudeck 1992; Hu and Bentler 1999; Muthen and Muthen 2009). Oblique Geomin rotation was applied to the factor solution. The Geomin method is recommended when it is expected that factors will be correlated and there may be cross-loading factors (Allison 2009; Muthen and Muthen 2010). In this study, factor loadings $\geq$0.35 were considered salient (Floyd and Widaman 1995). Once the factor solution was chosen, items that did not load saliently (i.e. <0.35) on to any of the factors or cross-loaded onto more than one factor but loaded only modestly on both factors (i.e. <0.5) were removed from further analysis. In addition, factors where only one or two items saliently loaded onto them were also removed since at least three items were required on each factor to ensure each one was well-measured (DeCoster 1998).

To enable greater insight into the psychometric properties of the Mandarin CAST was represented graphically using two IRT approaches: the item characteristic curve (ICC) and the test information curve (TIC). The graphs for each factor were presented separately. The ICCs provided trace lines for IRT models that were defined by each item's difficulty parameter (related to the intercept/threshold) and discrimination parameter (related to the factor loading). They also showed the probability of responding positively over the full range of the latent trait (Edelen and Reeve 2007).

The TIC provides a graphical representation of the precision of the measurement—related to reliability—over the full range of the latent trait (Edelen and Reeve 2007; Sharp et al. 2006). Since information is equal to the inverse of the squared standard error of the measurement, it is useful for indicating where test has the greatest precision for distinguishing amongst individuals (Lord 1980).

Results

Participants

The mean age of children was 8.2 years old (SD = 1.33, range 4–11.4). After statistical comparison, no differences were observed between the children included in this analysis and those excluded due to missing data. The characteristics of the two samples are shown in Table 1.

The Categorical Data Factor Analysis Model

All the 31 items that could be scored on the Mandarin CAST were included in the CDFA. Based on the Kaiser criterion, the eigenvalues suggested a factor solution of up to seven factors, while the scree plot suggested a two or possibly three-factor solution. Thus, in the following analysis, models extracting up to seven factors were considered. Model fit statistics were in the adequate range for the two-factor solution, and naturally improved with each additional factor extracted (Table 2).

The first factor included items concerning social interaction and communication (factor 1), while the second factor included items concerning inflexible/stereotyped language and behaviours (factor 2). Sixteen items loaded onto the first factor and thirteen items loaded onto the second factor. There were two items, noticing unusual details (item 6) and unusual memory (item 19), that did not have salient loadings on either factor. The correlation between these two factors was moderate (Geomin rotated factor correlations (GFC) = 0.402). The three-factor solution proposed a third factor, Attention to detail (factor 3). The third factor only included the two items (items 6 and 19). In order to ensure that factors are measured well, it is recommended that each factor should have at least three indicators (observed items) (Muthen and Muthen 2009). Thus, only a two factor solution was considered further. Three items (item 6, 14, and 19) were removed from further analysis since they did not saliently load onto either factor. Item 37 was observed to load positively on factor 2 and negatively on factor 1, however, not modestly (>0.35). Thus, item 37 was kept for factor 2 further analysis. The CDFA was re-run with the remaining 29 items for a two-factor solution. The model was still stable and met the goodness of fit indices criteria (RMSEA = 0.029, CFI = 0.957, TLI = 0.950, SRMR = 0.064). The correlation between the two factors was moderate (GFC = 0.425). The factor loadings for the 28 items are shown in Table 3.

Table 1 Characteristics of the sample

Characteristics	Category	Mainstream group		Clinical group	
Age	Mean	8.3 (SD 1.2)		6.3 (SD 1.6)	
		Number	(%)	Number	(%)
Sex	Boys	342	(52.2)	35	(89.7)
	Girls	302	(46.1)	4	(10.3)
	Missing	11	(1.7)	0	(0.0)
Siblings	Only child	503	(76.8)	31	(79.5)
	Having sibling	114	(17.4)	5	(12.8)
	Missing	38	(5.8)	0	(7.7)
Father's occupation	Worker or farmer	110	(16.8)	6	(15.4)
	Clerk	195	(29.8)	5	(12.8)
	Technical staff	142	(21.7)	20	(51.3)
	Manager	30	(4.6)	0	(0.0)
	Own-business	113	(17.3)	5	(12.8)
	Missing	65	(9.9)	3	(7.7)
Mother's occupation	Worker or farmer	162	(24.7)	8	(20.5)
	Clerk	157	(24.0)	7	(18.0)
	Technical staff	153	(23.4)	16	(41.0)
	Manager	11	(1.7)	0	(0.0)
	Own-business	108	(16.5)	4	(10.3)
	Missing	64	(9.8)	4	(10.3)
Father's education	Junior high school	101	(15.4)	2	(5.1)
	High school	170	(26.0)	9	(23.1)
	College	292	(44.6)	16	(41.0)
	Master or higher	47	(7.2)	9	(23.1)
	Missing	45	(6.9)	3	(7.7)
Mother's education	Junior high school	118	(18.0)	1	(2.6)
	High school	180	(27.5)	11	(28.2)
	College	284	(43.4)	18	(46.2)
	Master or higher	32	(4.9)	5	(12.8)
	Missing	471	(6.3)	4	(10.3)
CAST score	Mean	7.8 (SD 3.7)		20.5 (SD 5.0)	

Table 2 Model fit statistics by factor solutions from exploratory CDFA (31 items)

Factors	Chi square			RMSEA	CFI	TLI	SRMR	Negative residual variances	Eigen values
	χ^2	*df*	*p*						
1	1,475.091	434	0.000	0.059	0.783	0.767	0.115	No	9.496
2	684.626	404	0.000	0.032	0.941	0.933	0.068	No	3.601
3	546.210	375	0.000	0.026	0.964	0.956	0.060	No	1.687
4	446.310	347	0.0002	0.020	0.979	0.972	0.052	No	1.434
5	366.505	320	0.0374	0.014	0.990	0.986	0.045	No	1.323
6	320.408	294	0.1389	0.011	0.994	0.991	0.041	No	1.137
7	272.966	269	0.4212	0.005	0.999	0.999	0.037	No	1.082

Indices criteria for a model of good fit: RMSEA $\leq$ 0.06, CFI $\geq$ 0.95, TLI $\geq$ 0.95, SRMR $\leq$ 0.08

Graphical Presentation of IRT

The ICC and TIC of the IRT model for the two factors are shown in Fig. 1. With regard to factor 1, all 16 item difficulties are roughly at the same level for the social and communication trait and all items are located to the right of the figures. This indicates that they are located towards the more severe end of the continuum. The ICCs of factor 1

Table 3 Geomin rotated factor loadings for the two-factor exploratory CDFA solution (28 items)

Item	Social and communication	Inflexible/ stereotyped language and behaviours
1. Play game with others	**0.816**	0.172
2. Spontaneous chating	**0.605**	0.222
5. Fit in peer group	**0.721**	0.056
7. Takes things literally	0.001	**0.514**
8. Pretends to play	**0.726**	−0.167
9. Does the same thing over and over	−0.062	**0.404**
10. Easy to interact	**0.768**	0.210
11. Keeps two-way conversation	**0.757**	0.207
13. Same interests as peers	**0.652**	−0.017
15. Has friends	**0.558**	0.015
16. Show others things of interest	**0.694**	−0.110
17. Enjoys joking around	**0.566**	−0.039
18. Difficulty in understanding the rule of polite behaviours	0.008	**0.370**
20. Unusual voice	−0.095	**0.513**
21. Considers people important	**0.493**	0.092
23. Turn-taking conversation	**0.538**	0.230
24. Engages in role-play	**0.842**	−0.021
25. Tactless language and socially inappropriate behaviours	0.112	**0.637**
27. Eye contact	**0.628**	0.064
28. Unusual and repetitive movements	−0.001	**0.621**
29. One-sided social behaviours	0.279	**0.626**
30. Pronominal reversal	0.015	**0.543**
31. Prefers imaginative activities	**0.644**	−0.264
32. Loses listeners	0.082	**0.643**
34. Imposes routines	−0.143	**0.632**
35. Cares about the perception by others	**0.415**	−0.113
36. Turns conversation to his/her own interests	0.097	**0.668**
37. Odd or usual phrases	**−0.393**	**0.627**

The values in bold are salient loadings

indicates that a child located between one and two standard deviations (SD) above the population mean on the latent trait would have a 50 % probability of endorsing the Mandarin CAST items. Children at the mean latent trait value (0) have a low probability of endorsing any factor 1 item on the Mandarin CAST. The shapes of the ICC curves of 16 items are similar with sharp slopes, indicating a high discriminating power for these items with respect to the social and communication trait. The TIC of factor 1 shows that measurement precision, and thus reliability, is highest around 1.7 SD above the mean of the latent trait. At this point reliability is high at around 0.89, and is also good for people above the mean of latent trait (Table 4). This has implications for the interpretation of low and very high should have atleast trait.

In terms of factor 2, all 12 item difficulties are located towards the more severe end of the continuum (the right of the figure). The ICCs of factor 2 indicate that a child located at around 1 SD above the population mean (0) on this factor would have a 50 % probability of endorsing the Mandarin CAST items. The TIC of factor 2 indicates that measurement precision is highest around 1 SD above the mean of the latent trait. Reliability is 0.83 at this point (Table 4). The two-factor structure of the 28-item CAST is shown in Fig. 2.

Discussion

The exploratory CDFA suggested the Mandarin CAST measures two latent autistic traits, social and communication, and inflexible/stereotyped language and behaviours. It can be considered as a useful screening instrument of ASC with good reliability to discriminate individuals with ASC from those without ASC. It is notable that we also identified a third factor (attention to details) in this sample, however, the CAST was found to measure well for the two factors. The third factor suggested by the analysis focused on the unusual ability to remember details that children with ASC sometimes have. This ability could be considered as one of the unusual talents that have been observed in children with Asperger's Syndrome (Conson et al. 2011; Glanzman 2010; James 2010; Sevik et al. 2010). The two-factor model was also supported by other studies conducted in Western populations (Shuster et al. 2013). For example, this potential factor has been proposed by two studies based on data from the AQ (Auyeung et al. 2008; Hoekstra et al. 2008). However, there has been a move towards combining the social and communication deficits into a social-communication factor, and considering repetitive behaviours, interests and activities (RBIA) as another separate factor (Kuenssberg and McKenzie 2011). One previous study investigated the factor structure of the ADI-R algorithm, which suggested a two-factor solution: Stereotyped Language and RBIA, and Impairments in Social Interaction and Communication (Frazier et al. 2008). Another study on the ADI-R algorithm also identified two factors: social/communication and restricted/repetitive behaviours (Snow et al. 2009). A similar factor solution

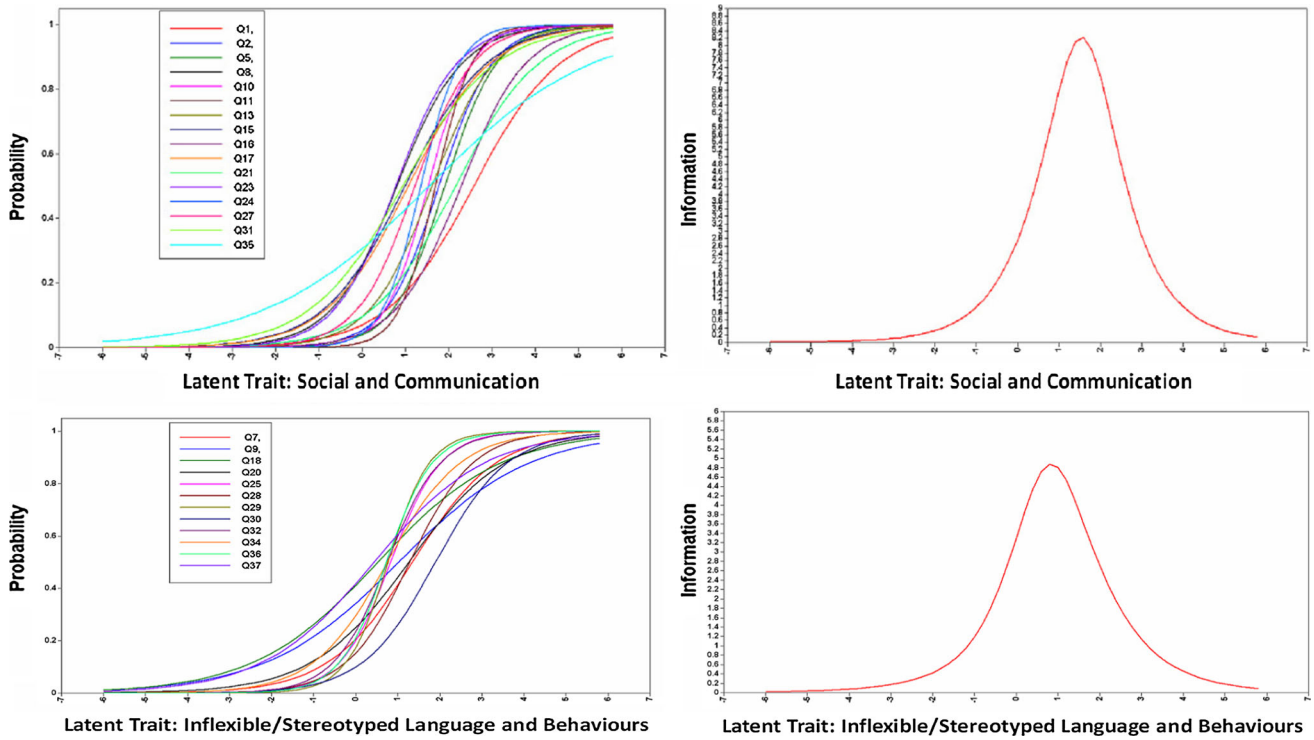

Fig. 1 ICC and TIC for two latent traits of the 28-item Mandarin CAST

Table 4 Reliability of the Mandarin CAST measuring two factors across the range of the latent trait (θ)

θ	Factor 1 Reliability	Factor 2 Reliability
−3	0.09	0.17
−2	0.29	0.36
−1	0.50	0.55
0	0.74	0.76
1	0.87	0.83
2	0.88	0.74
3	0.74	0.55

was further proposed by one study based on another screening questionnaire, the Adult Asperger Assessment (Baron-Cohen et al. 2005; Kuenssberg and McKenzie 2011) and the SRS (Frazier et al. 2012). In this study, the correlation between the two factors was found to be modest which suggested that the two factors may be largely independent of each other. Thus, this study provides further evidence for a two-factor model regarding autistic traits which was also suggested by a recent review (Shuster et al. 2013). Although the CAST was designed according to the criteria in DSM-IV, the latent traits identified in this Chinese population are in line with the DSM-V (Ghaziuddin 2010). It also provides evidence for the adaptability of the CAST in a different culture.

Regarding the performance of the CAST items, except the two items (item 6 and 19) on the third factor, item 14 did not perform well for either factor. These three items were removed. In addition, item 37 cross-loaded on both factors. However, as items are expected to cross-load due the correlation between latent factors, especially when both cross-loadings are salient, item 37 was not removed. Thus, the final CAST contained 28 items.

Regarding the utility of the CAST, this study found the CAST measured autistic traits with high precision in people around 1–2 SD's above the mean of the latent trait. This suggests that the CAST can be used as a screening instrument to discriminate between ASC and non-ASC individuals for population-based studies. However, the variation of reliability within the full range of the latent autistic trait (Table 4) suggested that caution needs to be adopted when using the current CAST to assess autistic symptoms. A previous twin study used the CAST and identified genetic heterogeneity across symptom domains (Robinson et al. 2011). Since the reliability of the CAST decreased when measuring the low/moderate levels of autistic traits, if the factor structure identified from this study is applicable to a western sample, it is possible that the decreased reliability of the CAST could potentially

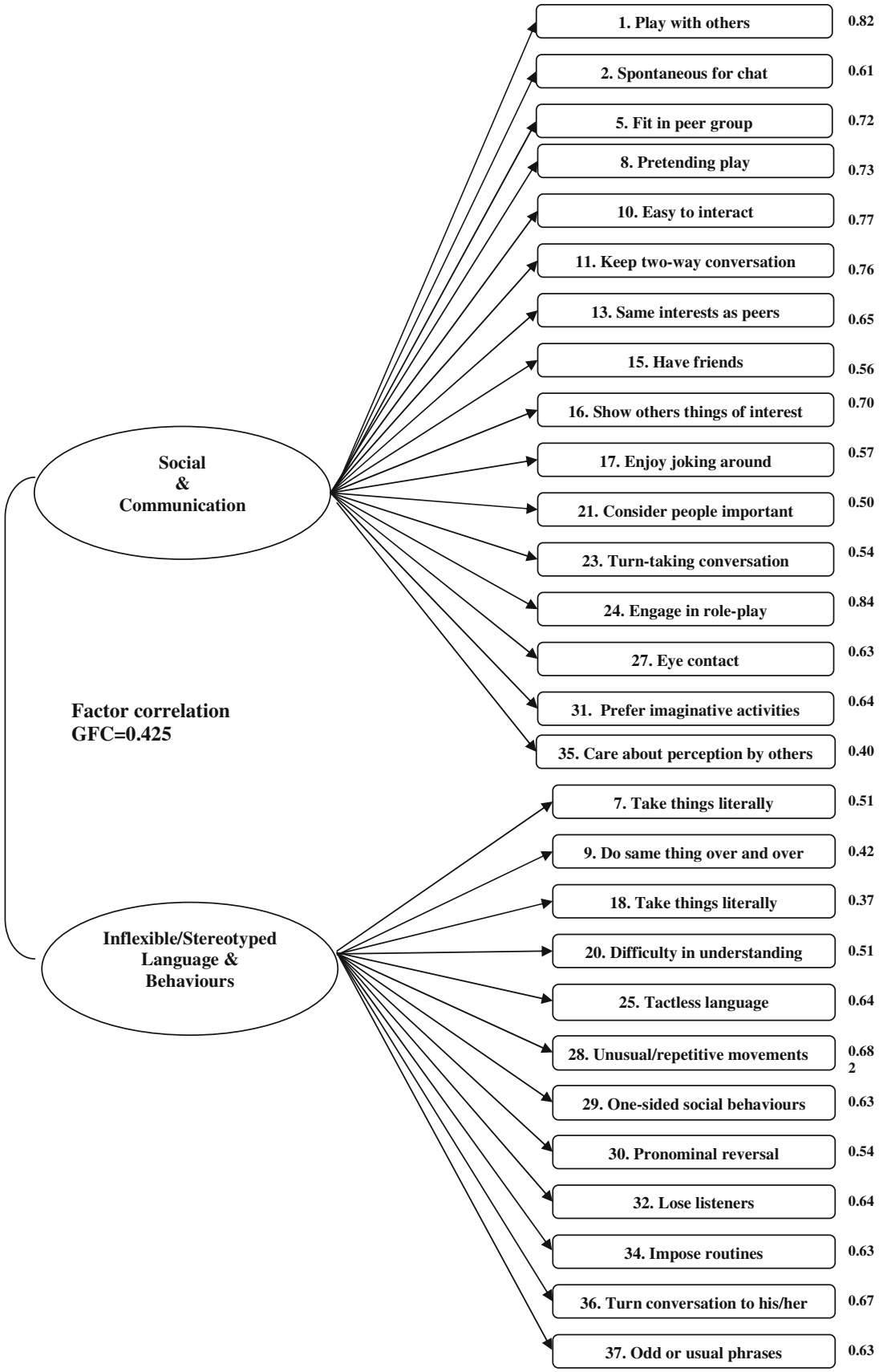

Fig. 2 Two-factor model of the Mandarin CAST (28 items)

underestimate the correlation across different traits. However, there is a clear difference between the samples of the twin study and the current study. In the former study, autistic traits were measured in a representative twin sample. In the current study, autistic traits were measured in a combined sample from a clinical setting and mainstream schools. This may have lead to the differences in the factor model generated from the CAST data.

Regarding the clinical implications of the CAST, due to the inconsistent reliability of the instrument, it is important to remember that a low score on the CAST does not necessarily indicate the absence of autistic traits. Similarly, a very high CAST score may not necessarily indicate more severe symptoms. This is a common issue with screening tools that is accentuated by the dichotomous scaling approach of the instrument. The current scaling approach may have limited the measurement ability of the CAST for prognosis and treatment tracking. Further research is needed to develop more dimensional scaling approach to instrument development and investigate whether other ASC screening instruments have the same problems.

Limitations

A limitation of this study is the representativeness of the study sample. Both children with an existing diagnosis of ASC and children in the general population were recruited. Most of the children with an existing diagnosis (childhood autism) were children at the more severe end of the spectrum. This sampling approach was also adopted in a previous study in the UK (Auyeung et al. 2008). The combined sample will not be representative of the general population in mainland China. However, the CAST was developed among the general population in the UK, where children with ASC are integrated into mainstream schools. In mainland China, children with ASC are turned away from mainstream schools (Sun et al. 2012). Thus, the mainstream school population in mainland China may not be comparable to the mainstream school population in the UK. The sampling strategy in this study aimed to produce a more heterogeneous sample with an appropriate proportion of individuals with ASC and typically developing children.

Although most factor analysis studies have been conducted among clinical samples, it has been suggested by a previous factor analysis study using the ADI-R that studies examining measurement factor structures should also be performed in more heterogeneous samples with a greater proportion of individuals who do not meet the diagnostic criteria for the ASC (Frazier et al. 2008). In addition, due to missing values, not all the collected questionnaires were used for this analysis. Another approach using data imputation for those missing values could be conducted. However, 91 % of the questionnaires were fully completed and it is unlikely that the excluded questionnaires would affect the results of the factor structure.

Personal judgments were required in the factor analysis during the reduction of items, which may have influenced the results (Allison et al. 2008). These judgments included choosing the criteria for a salient loading, the factor extraction and the criteria for indices of model fit. However, the inclusion criteria choices adopted in this study were consistent with previous studies (Frazier et al. 2008; Kanne et al. 2011; Pandolfi et al. 2009). In addition, this study is only an exploratory study. Due to the limited sample size, we only performed an exploratory factor analysis and removed items with poor fit to get a fit model. In the future, a confirmatory factor analysis should be performed to confirm the two-factor structure in larger Chinese sample.

Conclusion

A two-factor model was identified for the CAST: social and communication, and inflexible/stereotyped language and behaviours, which is in line with the DSM-V criteria of ASC. The Mandarin CAST measured the two latent traits adequately at the point of discrimination, which suggested the CAST can be considered as a screening instrument for discrimination between children with and without ASC for population-based studies. However, the interpretation of the scores needs to be made with caution since there is variation of reliability when measuring the latent traits. The two factors measured by the Mandarin CAST provided some evidence to support the idea that the symptom manifestations of ASC in Chinese children shared some similarities with Western populations.

Acknowledgments We are grateful to the families who participated in this study. This study was funded by the Waterloo Foundation, the Peking University First Hospital, the Cambridge Institute of Public Health and the Autism Research Centre in the University of Cambridge. XS was partly funded by Cambridge Commonwealth Trust and Clare Hall of the University of Cambridge during the data collection and then funded by the International Development Fund - Cambridge-CUHK Collaboration on Autism Research in Hong Kong and China during the writing up of this paper. SBC, CA and BA were funded by the Medical Research Council UK and the Wellcome Trust, and the team were funded by the NIHR CLAHRC for the Cambridgeshire and Peterborough NHS Foundation Trust during the period of this work. FM was funded by MRC UK.

Appendix

See Table 5.

Table 5 Description of studies on factor analysis of screening instruments for ASC in Western populations

Screening instrument	Author and year	Sample size	Sample source	Age (SD)	Analytical methods	No. of factors	Proposed factors
Autism Spectrum Screening Questionnaire (ASSQ)	Posserud et al. (2008)	6,229	General population	7–9	PCA and EFA	3	Social function Autism-associated problems Cognitive style associated with HFA/AS
Social and Communication Questionnaire (SCQ)	Gau et al. (2011)	736	Clinical ASC and relatives	2–18	CFA	3	Social interaction Communication Repetitive behaviour
Social Responsiveness Scale (SRS)	Constantino et al. (2004)	226	Clinical PDD and other disorders	4–18	CA and PCA	1	Single underlying 'autism' factor
	Frazier et al. (2012)	14,744	Clinical ASC and non ASC	2–18	EFA, LCA and FM	2	Social communication/ interaction Restrcted/repetitive behaviours
Autism quotient (AQ)	Auyeung et al. (2008)	1,765	General population and clinical ASC	4–9	PCA	4	Mind-reading Attention to detail Social skill Imagination
	Hoekstra et al. (2008)	1,299	University students and general population	Mean 21.9 (3.69) 35.68 (6.33)	CFA	2	Social interaction Attention to detail
	Stewart and Austin (2010)	536	University students	Mean 24.3 (10.5)	EFA and CFA	4	Socialness Pattern Understanding others/ communication Imagination
Gilliam Autism Rating Scale (GARS)	Lecavalier (2005)	284	ASC students	Mean 9.3 (3.9)	PCA and EFA	3	Stereotyped and repetitive behaviours Social interaction Communication
Adult Asperger Assessment (AAA)	Kuenssberg and McKenzie (2011)	153	Clinical AS and HFA	Mean 33 (11)	CFA	2	Social communication RIBA
Childhood Autism Rating Scale (CARS)	Magyar and Pandolfi (2007)	164	ASC students	Mean 43.27 months (19.78)	PCA and PAF	4	Social communication Relating to people and visual response Stereotypy and sensory abnormalities Emotional regulation

Table 5 continued

Screening instrument	Author and year	Sample size	Sample source	Age (SD)	Analytical methods	No. of factors	Proposed factors
Sub-threshold Autism Trait Questionnaire (SATQ)	Kanne et al. (2011)	1,709	University students	Mean:18.4 (0.99)	EFA and CFA	5	Social interaction and enjoyment Oddness Reading facial expressions Expressive language Rigidity

PCA principle component analysis, *HFA* high functional autism, *AS* Asperger syndrome, *EFA* exploratory factor analysis, *CFA* confirmatory factor analysis, *RIBA* repetitive/restricted interests behaviours and activities, *LCA* latent class analysis, *FM* factor mature model analysis

References

Allison, C. (2009). *The quantitative checklist for autism in toddlers (Q-CHAT).* (PhD) University of Cambridge, Cambridge.

Allison, C., Baron-Cohen, S., Wheelwright, S., Charman, T., Richler, J., Pasco, G. (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*(8), 1414–1425.

American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders, DSM-IV-TR*. Washington, DC: American Psychiatric Association.

Angold, A., Erkanli, A., Silberg, J., Eaves, L., & Costello, E. J. (2002). Depression scale scores in 8–17-year-old: Effects of age and gender. *Journal of Child Psychology and Psychiatry, 43*(8), 1052–1063.

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*(7), 1230–1240.

Bandalos, D. L., & Boehm-Kaufman, M. R. (2009). *Four common misconceptions in exploratory factor analysis In Statistical and methodological myths and urban legends: Doctrine, verity and fable in the organizational and social sciences*. New York: Routledge.

Baron-Cohen, S., Wheelwright, S., Robinson, J., & Woodbury-Smith, M. (2005). The Adult Asperger Assessment (AAA): A diagnostic method. *Journal of Autism and Developmental Disorders, 35*(6), 807–819.

Bartholomew, D. J., Steel, F., Moustaki, I., & Galbraith, J. (2002). *The analysis and interpretation of multivariate data for social scientists*. London: Chapman & Hall/CRC.

Bolte, S., Westerwald, E., Holtmann, M., Freitag, C., & Poustka, F. (2011). Autistic traits and autism spectrum disorders: The clinical validity of two measures presuming a continuum of social communication skills. *Journal of Autism and Developmental Disorders, 41*(1), 66–72.

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

Browne, M. W., & Cudeck, R. (1992). Alternative ways of assessing model fit. *Sociological Methods & Research, 21*(2), 230–258.

Conson, M., Salzano, S., & Grossi, D. (2011). Neuropsychological functioning of an Asperger child with exceptional skill in arranging picture stories. *Neurocase, 17*(4), 353–359.

Constantino, J. N., Gruber, C. P., Davis, S., Hayes, S., Passanante, N., & Przybeck, T. (2004). The factor structure of autistic traits. *Journal of Child Psychology and Psychiatry, 45*(4), 719–726.

Constantino, J. N., & Todd, R. D. (2003). Autistic traits in the general population: A twin study. *Archives of General Psychiatry, 60*(5), 524–530.

DeCoster, J. (1998). Overview of factor analysis. Retrieved 6 Aug 2012, from http://www.stat-help.com/notes.html.

Edelen, M. O., & Reeve, B. B. (2007). Applying item response theory (IRT) modeling to questionnaire development, evaluation, and refinement. *Quality of Life Research, 16*, 5–18.

Flora, D. B., & Curran, P. J. (2004). An empirical evaluation of alternative methods of estimation for confirmatory factor analysis with ordinal data. *Psychological Methods, 9*(4), 466–491.

Floyd, F. J., & Widaman, K. F. (1995). Factor analysis in the development and refinement of clinical assessment instruments. *Psychological Assessment, 7*(3), 286–299.

Frazier, T. W., Youngstrom, E. A., Kubu, C. S., Sinclair, L., & Rezai, A. (2008). Exploratory and confirmatory factor analysis of the autism diagnostic interview-revised. *Journal of Autism and Developmental Disorders, 38*(3), 474–480.

Frazier, T. W., Youngstrom, E. A., Speer, L., Embacher, R., Law, P., & Constantino, J. (2012). Validation of proposed DSM-5 criteria for autism spectrum disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 51*(1), 28–40.

Gau, S. S. F., Lee, C. M., Lai, M. C., Chiu, Y. N., Huang, Y. F., & Kao, J. D. (2011). Psychometric properties of the Chinese version of the Social Communication Questionnaire. *Research in Autism Spectrum Disorders, 5*(2), 809–818.

Georgiades, S., Szatmari, P., Zwaigenbaum, L., Duku, E., Bryson, S., & Roberts, W. (2007). Structure of the autism symptom phenotype: A proposed multidimensional model. *Journal of the American Academy of Child and Adolescent Psychiatry, 46*(2), 188–196.

Ghaziuddin, M. (2010). Should the DSM V drop Asperger syndrome? *Journal of Autism and Developmental Disorders, 40*(9), 1146–1148.

Glanzman, M. (2010). Developing talents: Careers for individuals with asperger syndrome and high-functioning autism, 2nd edition. *Journal of Autism and Developmental Disorders, 40*(2), 266–267.

Happe, F., & Ronald, A. (2008). The 'fractionable autism triad': A review of evidence from behavioural, genetic, cognitive and neural research. *Neuropsychology Review, 18*(4), 287–304.

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*(8), 1555–1566.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal, 6*(1), 1–55.

James, I. (2010). Autism and mathematical talent. *Mathematical Intelligencer, 32*(1), 56–58.

Kamio, Y. (2007). Early detection of and diagnostic tools for Asperger's disorder. *Nippon Rinsho/Japanese Journal of Clinical Medicine, 65*(3), 477–480.

Kanne, S. M., Wang, J., & Christ, S. E. (2011). The Subthreshold Autism Trait Questionnaire (SATQ): Development of a brief self-report measure of subthreshold autism traits. *Journal of Autism and Developmental Disorders, 42*(5), 769–780.

Khalid, M. N. (2011). An overview of statistical approaches for assessing model fit. *The International Journal of Educational and Psychological Assessment, 8*(2), 69–87.

Kline, P. (2000). *Handbook of psychometric testing*. London: Routledge.

Kuenssberg, R., & McKenzie, K. (2011). Confirmatory factor analysis of the Adult Asperger Assessment: The association of symptom domains within a clinical population. *Research in Developmental Disabilities, 32*(6), 2321–2329.

Kuenssberg, R., McKenzie, K., & Jones, J. (2011). The association between the social and communication elements of autism, and repetitive/restrictive behaviours and activities: A review of the literature. *Research in Developmental Disabilities, 32*(6), 2183–2192.

Lau, W. Y., Gau, S. S., Chiu, Y. N., Wu, Y. Y., Chou, W. J., & Liu, S. K. (2012). Psychometric properties of the Chinese version of the autism spectrum quotient (AQ). *Research in Developmental Disabilities, 34*(1), 294–305.

Lecavalier, L. (2005). An evaluation of the Gilliam Autism Rating Scale. *Journal of Autism and Developmental Disorders, 35*(6), 795–805.

Lecavalier, L., Aman, M. G., Scahill, L., McDougle, C. J., McCracken, J. T., & Vitiello, B. (2006). Validity of the autism diagnostic interview-revised. *American Journal of Mental Retardation, 111*(3), 199–215.

Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Magyar, C. I., & Pandolfi, V. (2007). Factor structure evaluation of the childhood autism rating scale. *Journal of Autism and Developmental Disorders, 37*(9), 1787–1794.

Mandy, W. P., & 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*(8), 795–808.

McCabe, H. (2003). The beginnings of inclusion in the People's Republic of China. *Research and Practice for Persons with Severe Disabilities, 28*(1), 16–22.

McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah, NJ: LEA.

Muthen, B. (1989). Multiple-group structural modeling with non-normal continuous-variables. *British Journal of Mathematical and Statistical Psychology, 42*, 55–62.

Muthén, B., du Toit, S. H. C., & Spisic, D. (1997). Robust inference using weighted least squares and quadratic estimating equations in latent variable modeling with categorical and continuous outcomes. Psychometrika (accepted).

Muthen, L. K., & Muthen, B. (2009). MPlus short courses. http://www.statmodel.com/course_materials.shtml.

Muthen, L. K., & Muthen, B. O. (2010). *Mplus Users' guide version 6*.

Pandolfi, V., Magyar, C. I., & Dill, C. A. (2009). Confirmatory factor analysis of the child behaviour checklist 1.5–5 in a sample of children with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 39*(7), 986–995.

Posserud, M. B., Lundervold, A. J., & Gillberg, C. (2006). Autistic features in a total population of 7–9-year-old children assessed by the ASSQ (Autism Spectrum Screening Questionnaire). *Journal of Child Psychology and Psychiatry, 47*(2), 167–175.

Posserud, B., Lundervold, A. J., Steijnen, M. C., Verhoeven, S., Stormark, K. M., & Gillberg, C. (2008). Factor analysis of the Autism Spectrum Screening Questionnaire. *Autism*, *12*(1), 99–112.

Robinson, E. B., Koenen, K. C., McCormick, M. C., Munir, K., Hallett, V., & Happe, F. (2011). Evidence that autistic traits show the same etiology in the general population and at the quantitative extremes (5, 2.5, and 1 %). *Archives of General Psychiatry, 68*(11), 1113–1121.

Scott, F. J., Baron-Cohen, S., Bolton, P., & Brayne, C. (2002a). Brief report: prevalence of autism spectrum conditions in children aged 5–11 years in Cambridgeshire, UK. *Autism, 6*(3), 231–237.

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

Sevik, A. E., Kultur, E. C., Demirel, H., Oguz, K. K., Akca, O., & Ergun, E. L. (2010). Asperger syndrome with highly exceptional calendar memory: A case report. *Turk Psikiyatri Dergisi, 21*(3), 249–255.

Sharp, C., Goodyer, I. M., & Croudace, T. J. (2006). The Short Mood and Feelings Questionnaire (SMFQ): A unidimensional item response theory and categorical data factor analysis of self-report ratings from a community sample of 7-through 11-year-old children. *Journal of Abnormal Child Psychology, 34*(3), 379–391.

Shuster, J., Perry, A., Bebko, J., & Toplak, M. E. (2013). Review of factor analytic studies examining symptoms of autism spectrum disorders. *Journal of Autism and Developmental Disorders, 44*(1), 90–110.

Snow, A. V., Lecavalier, L., & Houts, C. (2009). The structure of the autism diagnostic interview-revised: Diagnostic and phenotypic implications. *Journal of Child Psychology and Psychiatry, 50*(6), 734–742.

Spiker, D., Lotspeich, L. J., Dimiceli, S., Myers, R. M., & Risch, N. (2002). Behaviorial phenotypic variation in autism multiplex families: Evidence for a continuous severity gradient. *American Journal of Medical Genetics, 114*(2), 129–136.

Stewart, M. E., & Austin, E. J. (2010). The structure of the Autism-Spectrum Quotient (AQ): Evidence from a student sample in Scotland. *Personality and Individual Differences*, *47*(3), 224–228.

Sun, X., Allison, C., Auyeung, B., Baron-Cohen, S., & Brayne, C. (2012a). A review of healthcare service and education provision of Autism Spectrum Condition in mainland China. *Research in Developmental Disabilities, 34*(1), 469–479.

Sun, X., Allison, C., Auyeung, B., Matthews, F. E., Murray, S., & Baron-Cohen, S. (2012b). Service provision for autism in mainland China: A service providers' perspective. *Research in Developmental Disabilities, 34*(1), 440–451.

Van Lang, N. D., Boomsma, A., Sytema, S., de Bildt, A. A., Kraijer, D. W., & Ketelaars, C. (2006). Structural equation analysis of a hypothesised symptom model in the autism spectrum. *Journal of Child Psychology and Psychiatry, 47*(1), 37–44.

Williams, J. (2003). *Screening for autism spectrum disorders*. Cambridge: University of Cambridge.

Williams, J., Scott, F., Stott, C., Allison, C., Bolton, P., & Baron-Cohen, S. (2005). The CAST (Childhood Asperger Syndrome Test): Test accuracy. *Autism, 9*(1), 45–68.

World Health Organisation. (1993). *The ICD-10 classification of mental and behavioural disorder: Diagnosis criteria for research*. Geneva: World Health Organisation.