Revue européenne de psychologie appliquée 67 (2017) 299–306

Disponible en ligne sur
ScienceDirect
www.sciencedirect.com

Elsevier Masson France
EM|consulte
www.em-consulte.com/en

Original article

The assessment of autistic traits with the Autism Spectrum Quotient: Contribution of the French version to its construct validity

Évaluation des traits autistiques par le Quotient du Spectre Autistique : contribution d'une version française à sa validation structurelle

C. Kempenaers[a,b,*], S. Braun[a,b], N. Delvaux[b], P. Linkowski[a]

[a] *Laboratory of Psychiatric Research, Psychiatric Department, C.U.B. Erasmus Hospital, Université libre de Bruxelles, Brussels, Belgium*
[b] *Psychology Department, C.U.B. Erasmus Hospital, Université libre de Bruxelles, Brussels, Belgium*

ARTICLE INFO

Article history:
Received 14 July 2015
Received in revised form
1st September 2017
Accepted 4 September 2017

Keywords:
Autism Spectrum Quotient
Structural validity
Confirmatory factor analysis
French version

Mots clés :
Quotient de Spectre Autistique
Validité structurelle
Analyse factorielle confirmatoire
Version française

ABSTRACT

Introduction. – The Autism Spectrum Questionnaire (AQ, Baron-Cohen et al., 2001) is a self-report assessment tool aiming at screening autistic traits in normal intelligence adults. While numerous versions in other languages than English now exist, few factorial evidence do sustain the valid use of this instrument as it was conceived, based upon five distinct dimensions (Social skills, Communication, Attention to detail, Attention switching, Imagination); no such study exists with a French version of the AQ. The aim of our study is therefore to present the French version of the scale and to explore its factorial validity with confirmatory factorial analyses and, possibly, its invariance across men and women.
Method. – Several confirmatory factorial analyses, with the robust WLSMV estimator for categorical response format, were run on the questionnaire data from 788 French-speaking students (17–25 years old) at university faculties or schools for higher education in Belgium. The original five-factor measurement model of the AQ was assessed as well as alternative models. An exploratory factorial analysis was also applied to get more insight as to possible sources of misfit.
Results. – No measurement model – neither the original five-factor one nor any of the six other models tested – did produce statistics or fit indices close to significant values: there was no fit to the data. The internal consistency of the subscales was weak; the exploratory factorial analysis further confirmed that as much as ten factors were needed to explain 45% of the data variance.
Conclusion. – Our results, with a French version of the scale, add to many other ones which suggest that the AQ is a too heterogeneous questionnaire with somewhat ill-defined dimensions and non specific/ambiguous items. The questionnaire should probably be shortened and its content realigned to core features of the autism spectrum.

RÉSUMÉ

Introduction. – Le Questionnaire du Spectre Autistique (AQ, Baron-Cohen et al., 2001) est un outil d'auto-évaluation permettant de dépister des traits autistiques chez des adultes d'intelligence normale. Bien que de nombreuses versions de l'AQ existent en d'autres langues que l'anglais, il n'y a que peu de données factorielles soutenant la validité de l'instrument tel qu'il a été conçu, en cinq dimensions distinctes (Habilités sociales, Communication, Attention au détail, Partage de l'attention, Imagination) ; aucune étude de la sorte n'existe pour une version française de l'AQ. Le but de notre étude est donc de présenter une version française du questionnaire et d'explorer sa validité factorielle au moyen d'analyses factorielles confirmatoires, et, si possible, son invariance selon le sexe.

* Corresponding author. C.U.B. Erasmus Hospital, 808, route de Lennik, 1070 Brussels, Belgium.
E-mail address: chantal.kempenaers@erasme.ulb.ac.be (C. Kempenaers).

http://dx.doi.org/10.1016/j.erap.2017.09.001
1162-9088/

Méthode. – Plusieurs analyses factorielles confirmatoires, utilisant l'estimateur WSLMV en raison du format de réponse catégoriel, ont été menées sur les données de questionnaire de 788 étudiants (17–25 ans) parlant français inscrits à l'université ou à de Hautes Écoles en Belgique. Le modèle original de mesure à cinq facteurs a été évalué, ainsi que plusieurs modèles alternatifs. Une analyse factorielle exploratoire a aussi été réalisée afin d'explorer les possibles sources de mauvais ajustement.

Résultats. – Aucun modèle de mesure – ni le modèle d'origine à cinq facteurs ni aucun des six autres modèles – n'a produit de statistiques ou d'indices d'ajustement s'approchant des valeurs-seuils : il n'y avait pas d'ajustement aux données. La consistance interne des sous-échelles était faible ; l'analyse factorielle exploratoire a aussi montré que pas moins de dix facteurs étaient nécessaires pour rendre compte d'environ 45 % de la variance des données.

Conclusion. – Nos résultats, obtenus avec une version française de l'échelle, s'ajoutent à ceux, nombreux, qui suggèrent que l'AQ est un questionnaire trop hétérogène composé de dimensions mal définies et d'items peu spécifiques/ambigus. Le questionnaire devrait sans doute être raccourci et son contenu réajusté pour correspondre mieux aux éléments centraux du spectre autistique.

1. Introduction

The interest for identifying and measuring autistic traits and differentiating them from other potentially disabling aspects of human psychological functioning has widely grown in the last 20 years (Volkmar, State, & Klin, 2009). One main progress in autism has probably been its nosographic evolution as illustrated in the revised DSM-5 diagnostic criteria (American Psychiatric Association, 2013; Lord & Jones, 2012; Lai, Lombardo, Chakrabarti, & Baron-Cohen, 2013); the impact of this recent theoretical shift is however by now largely unknown.

S. Baron-Cohen and his team in Cambridge developed the Autism-Spectrum Quotient based on the assumption of autistic behaviors being manifestations of an underlying continuous autism dimension (AQ; Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001). This self-administered questionnaire aims to provide a quantified evaluation of the degree to which an adult with a normal intelligence quotient shows signs of the broad autism phenotype. The AQ was conceived to sample behaviors/preferences and cognitions most typical of this autistic functioning along five correlated dimensions (social skills, attention switching, attention to detail, communication and imagination). Importantly, special efforts were devoted to the development of the AQ to make items readily understandable, according to their authors.

Despite Baron-Cohen et al. studies originally supporting the validity of the AQ (2001, Woodbury-Smith, Robinson, & Baron-Cohen, 2005), several psychometrical aspects of the questionnaire have raised concern. In particular, questions about its factorial validity and the internal consistency of some of its dimensions emerged ten years ago already (Austin, 2005; Hurst, Mitchell, Kimbrel, Kwapil, & Nelson-Gray, 2007). Most studies using exploratory factor analyses on the English AQ obtain factorial solutions of modest to moderate explanatory value (% explained variance) with a reduced number of factors and items. Some core dimensions are repeatedly found, consistent with the basic dimensions in Baron-Cohen questionnaire, as "socialness", communication, and cognitive dysfunctions (restricted interests/behaviors) (Austin, 2005; Hurst et al., 2007, Stewart & Austin, 2009, Russel-Smith, Mayberry, & Bayliss, 2011). A few, more recent, studies with confirmatory factor analysis (CFA) on the total AQ clearly do not obtain the fit of the five-factor original model to their data in student samples (Hoekstra, Bartels, Cath, & Boomsma, 2008; Kloosterman, Keefer, Kelley, Summerfeldt, & Parker, 2011; Lau, Gau et al., 2013). When a fit can be obtained for a five-factor model, a series of the original items do not significantly participate to the variance (no significant loading) or are not discriminant enough (complex items) (Hoekstra et al., 2011; Lau, Gau et al., 2013; Lau, Kelly, & Peterson, 2013); consequently, the resulting item sets produce a somewhat different questionnaire content. In a very comprehensive study with their Dutch adaptation of the questionnaire, Hoekstra et al. (2008) suggested that the best fitting model for their data was that of a hierarchical factor structure with a higher-order "Social Interaction" factor made of four lower-order domains (social skill, communication, attention switching, imagination) and a second "Attention to detail" factor; a subset of 28 items resulted from their combined exploratory and confirmatory analyses. Among authors using CFA, Kloosterman et al. (2011) also conclude that shortening the AQ allows a better adjustment to the data; maybe, as sketched by Kuennsberg, Murray, Booth, and McKenzie (2012), "a large variability in the factor models across samples" is rather the main lesson of all these studies. For a review of factorial analyses on the original 50-items AQ, see Table 1[1].

Despite these psychometrical restrictions, many translated versions of the questionnaire were developed, i.e. in Japanese (Wakabayashi, Baron-Cohen, Wheelwright, & Tojo, 2006), Dutch (Hoekstra et al., 2008; Ketelaars et al., 2008), Italian (Ruta, Mazzone, Mazzone, Wheelwright, & Baron-Cohen, 2011), Persian (Mohammadi, Zarafshan, & Ghasempour, 2012), Chinese (Lau, Gau et al., 2013), Polish (Pisula et al., 2013), and Turkish (Kose et al., 2013). In France, Rousselot-Pailley, Fortin, Golse, Falissard, and Robel (2011) and Robel et al. (2014) studied an adapted and shortened version of the AQ rather than preserving and translating the questionnaire, which does not allow assessing the quality of the original AQ in French; Sonié et al. (2011, 2013) present AQ data in French on adolescents but without any factorial validity assessment. This is also the case with Lepage, Lortie, Taschereau-Dumouchel, and Théoret (2009) in Canada who published basic data on a French-Canadian version of the AQ, without any structural analysis of the instrument.

The aim of the present study is therefore to translate the AQ into French and test for its factorial validity with CFA. Aside the five-dimension Baron-Cohen original measurement model, alternative models – as suggested by the aforementioned research studies – will be considered and substantive analyses of the CFA and EFA outputs will provide further theoretical insight into the autism-spectrum assessment tool. In addition, because the AQ was designed with no specificity for gender, measurement invariance across men and women will be considered provided that an ade-

[1] In order to get all published scientific work on factorial studies of the AQ, we proceeded with a systematic search on the following databases, available at the library of the Université libre de Bruxelles: Proquest, ADB, CIBLE+, EBSCO. The search was constrained by the following criteria: the paper should be written in English or French, have been published in a Journal with a lecture committee, must report on a study of the AQ (keyword: "autism spectrum quotient"); the questionnaire data must have been submitted to a factor analysis (at least an exploratory one), with subjects being adults (adolescents excluded).

Table 1
Summary of selected studies on factorial analyses of the AQ.

Authors, years [language][a]	Sample	Type of analysis (EFA/PCA – CFA)	Factors (number of items; % explained variance)
Austin (2005)	Undergraduate students	PCA	Social skills – details – patterns – communication – mindreading (26; 28%)
Hurst, Mitchell, Kimbrel, Kwapil, and Nelson-Gray (2007)	University students	PCA[b] [Promax] – (dichotomic response format; on 26 items)	Social skills – details – patterns – communication – mindreading (26; 29%)
Hoekstra, Bartels, Cath, and Boomsma (2008) [Dutch]	University students General population	CFA	Social interaction (by social skill, attention switching, communication, imagination) – attention to detail[c]
Stewart and Austin (2009)	University students	EFA [Oblimin]	Socialness – patterns – understanding others – communication – imagination (43; 29%)
Russel-Smith, Mayberry, and Bayliss (2011)	Undergraduate students Undergraduate students	EFA [Oblique rotation]	Social skills – details – patterns – communication – mindreading (27; 22%) Social skills – details – patterns – communication – mindreading – imagination (38; 26%)
Kloosterman, Keefer, Kelley, Summerfeldt, and Parker (2011)	Undergraduate students	CFA	Social skills – communication – mindreading – restricted behavior – imagination – attention to detail (28; 45%)
Lau, Gau et al. (2013) [Chinese]	Mixed (parents of ASD children + control parents)	CFA	Socialness – mindreading – patterns – attention to detail – attention switching (35; 42%)
Lau, Kelly, and Peterson (2013)	Mixed (general population + students) AS	CFA	Sociability – social cognition – narrow focus – interest in patterns – resistance to change (39; 48%)

EFA: exploratory factor analysis; PCA: principal component analysis; CFA: confirmatory factor analysis.
[a] When not English.
[b] PCA with number of components fixed to 3.
[c] Hierarchical factor model with 2 main factors and 4 secondary factors is the best fitting model.

quate fit is afforded first by any model tested. Testing the invariance of the measurement model of a questionnaire is indeed a prerequisite for a reliable interpretation of observed differences between subject subgroups (Gregorich, 2006; Byrne, 2008).

2. Method

2.1. Study subjects

The study was approved by the Ethics Committee of the Medical Faculty of the university the first author belongs to. Seven hundred eighty eight students, affiliated to several schools for higher education and university faculties (engineering [25.3%], medicine [20.1%], economics [14.1%], nursing school [13.8%], law [7.6%], psychology [7.0%], dietetics [3.2%]) participated on a voluntary basis to the study. Note that the heterogeneity of study fields supposedly reduces the risk of a systematic biased response style to the questionnaire. Out of the 788 subjects, 461 were female students (58.5%). The mean age was 19,32 (SD: 1.54) years with a range from 17 to 25 yrs. Race/ethnicity was not recorded; subjects were all French-speaking.

The participants responded to five self-report scales (see below); they recorded their responses on computer scannable answer sheets with a unique yet anonymous identification code. The assessment session was collective; it lasted about one hour and took place during a scheduled class time.

2.2. Questionnaire

The scale was administered in a semi-random order as part of a package of five self-report scales tapping emotional and social attitudes and traits (alexithymia, empathy, self-worth and resilience). The data from these other questionnaires will not be analyzed here. A questionnaire assessing some basic demographic features was also collected.

The AQ (Baron-Cohen et al., 2001) is composed of 50 items assessing five areas (ten items per subscale): Social Skills (Factor 1; e.g.: "I prefer to do things with others than on my own" [negatively keyed]), Attention Switching (F2; e.g.: "I prefer to do things the same way over and over again"), Attention to Detail (F3; e.g.: "I often notice small sounds when others do not"), Communication (F4; e.g.: "I frequently find that I don't know how to keep a conversation going"), and Imagination (F5; e.g.: "I find making up stories easy" [negatively keyed]). A Likert-type scale with four anchor points is used for responding (Definitely agree, slightly agree, slightly disagree and definitely disagree) but the scoring process is secondarily dichotomized so that responses "definitely agree" or "slightly agree" score 1 point for the positive items (those identifying autistic-like behaviors) and responses "definitely disagree" or "slightly disagree" score 0 point; the procedure is reversed for the negative items (those worded as identifying a non-autistic-like behavior). We adhered to the common scoring approach adopted by most researchers consisting in ignoring this transformation of responses and using instead the original 4-point responses completed by the subject. Total scores are obtained for each of the five subscales, with higher scores indicating a greater endorsement of autistic behaviors, attitudes or preferences.

The translation of the scale was a three-step process, as usual in the field: (1) the items were translated into French by the first author with an emphasis on conceptual and cultural rather than linguistic equivalence, (2) a native English psychologist with fluency in French then translated all items back into English, (3) the first author and the English collaborator finally confronted the original and back-translated versions of the questionnaire, and, when needed, adapted the French items to fit the meaning of the original English version better.

2.3. Data analysis

Mplus version 5.21 (Muthén & Muthén, 1998–2009) was used for CFA to assess the factorial validity of the proposed five-factor structure of the scale as well as to test alternative models and, if adequate (see below), measurement invariance across sex. SPSS versions 17 trough 19 (2008–2010) were used for other statistical analyses, like descriptive statistics and the consistency measure (Cronbach alpha), as well as for exploratory factor analyses.

2.4. Confirmatory factor analysis

CFA is a particular form of structural equation modeling (SEM) allowing to test for the factorial validity of a proposed measurement model. In order to accommodate the categorical format of the responses for estimating parameters of the common factor model, we used the robust WLSMV estimator (Muthén & Muthén, 1998–2004), which is one of the best options with categorical outcomes or severely non-normal data (Brown, 2006). WLSMV provides weighted least square parameter estimates using a diagonal weight matrix and robust standards errors and a mean- and covariance-adjusted χ^2 test statistic. Raw data were used as input for the analyses.[2] Scaling of the latent variables was set by fixing the loading of the first item of each subscale to 1.

The basic hypothesized measurement model (Model 1) assumed the following: (1) responses are explained by five factors: social skills, attention switching, attention to detail, communication, and imagination; (2) each subscale item has a nonzero loading on the factor that it is supposed to measure and a zero loading on the four other factors (congeneric model); (3) the five factors are correlated; (4) error variances are uncorrelated.

As alternative models, we tested the following, based on face content: (a) a four-factor model with the Social skills and Communication factors being collapsed, the Imagination factor, and the Attention Switching and Attention to Detail factors (Model 2), (b) a three-factor model with the two collapsed dimensions as in Model 2 and with the Attention switching and Attention to detail factors (two "cognitive" factors) being also collapsed, and the Imagination factor (Model 3), and lastly (c) a one-factor model (where all factors are collapsed as one global "autistic" dimension; Model 4). All other aspects of the hypothesized measurement models were kept unchanged (as in Model 1). Following the work of Hoekstra et al. (2008, 2011; Model 5), Kloosterman et al. (2011; Model 6), and Lau et al. (2013; Model 7), we also tested three other competing measurement models based on abridged AQ versions. Model 5 is a hierarchical model with a Social behavior higher-order factor, defined as overarching four primary factors (Social skills [7 items], Routine [4 items], Switching [4 items] and Imagination [8 items]), and a second principal factor labeled Number/Patterns (5 items). Model 6 posits the five following factors: social skills (8 items), communication/mindreading (5 items), restricted/repetitive behavior (5 items), imagination (5 items) and Attention to detail (5 items). Model 7 is also based on five factors: Sociability (13 items), Social cognition (11 items), Narrow focus (7 items), Interest in patterns (4 items), and Resistance to change (4 items). Readers are referred to the original papers to learn more about the rationale of the proposed models and their specific factor content.

Before evaluating invariance of the measurement model across sex, the best fitting model must first be tested separately in the two groups at hand – males and females. Only when the data are considered as fitting the specified model sufficiently well in men and women separately, may the gender invariance testing proceed as a progressively constraining assessment in a simultaneous CFA with both men and women (Brown, 2006).

[2] The matrix of variances-covariances can be obtained from the first author on request.

2.5. Chi-square test statistic and fit indices

The basic test of goodness-of-fit is the Chi-square, which is used for evaluating the fit of a prespecified model to the actual covariance/correlation data matrix. Because of its inherent limitations (due to its sensitivity to sample size, the Chi-square test may be statistically but not substantively significant in large samples), the Chi-square test statistic is usually supplemented by a series of additional indices (Brown, 2006), among which the following are the most commonly used ones (Jackson, Gillaspy, & Purc-Stephenson, 2009). The first index is the Comparative Fit Index (CFI, Bentler, 1990), for which a value of 0.90 or higher indicates a reasonable model fit and a value of 0.95 indicates a good fit. The second index is the Tucker-Lewis index (TLI, Tucker & Lewis, 1973) for which values approaching 1.0 (greater than 0.95) are interpreted as indicating a good fit. The Root Mean Square Error of Approximation (RMSEA, Browne & Cudeck, 1993) for which a value of 0.06 or lower indicates a good fit is a third widely used index of fit. Other guides to model evaluation (extent and sources of [mis]fit) rely on the examination of model estimated parameters and inspection of normalized residuals and modification indices provided by Mplus.

The Akaike Information Criterion (Brown, 2006; Hu, 2007), commonly presented when comparing competing non-nested models, cannot be used here because the databases implied by these alternative models are different (number of items).

2.6. Invariance testing

The reader is referred to Kempenaers, De Boeck, Dehon, Braun, and Linkowski (2014) for a summary of the invariance testing procedure. However, equivalence of the measurement model across sex can only be assessed when the a priori model has basically proven sufficiently adequate in both sexes first. Because, as our data will show, this is not the case here, the invariance testing was eventually not conducted.

2.7. Exploratory factor analysis

Additionally, an exploratory factor analysis (EFA) will be applied in case of bad fit to explore the possible causes of illfit. The unweighted least square extraction method (SPSS, 2008–2010) was preferred because of the ordinal quality of the Likert response variables (four-level category); next, the Oblimin rotation method was applied, supposing correlated dimensions. As to the selection of significant factors, a parallel analysis (PA, Horn, 1965) is considered as the best method (Hayton, Allen, & Scarpello, 2004); it was realized based on the Monte-Carlo program of Watkins (2005). Ledesma and Valero-Mora (2007) recommend using as cut-off to assess the observed eigenvalues the 95th percentile eigenvalue obtained through the simulation process instead of the mean. For the presentation of results, only loadings superior to 0.30 will be reported.

3. Results

3.1. Descriptive statistics and consistency of AQ

In contrast to the other statistics, based on the questionnaire responses in four categories (see method), means and standard deviations for the subscales and total scores were computed on dichotomized responses, as in the original work of Baron-Cohen et al. (2001).

Mean scores of the five dimensions of the questionnaire were normally distributed. As can be seen in Table 2, men scored a little higher on four of the subscales, while women scored higher on Attention to detail; however, unless measurement invariance

Table 2
Means (SD) for the five subscales of the French AQ (as dichotomous variables) and internal consistencies (categorical variables).

Subscales	Women[a]	Men[b]	Total[c]	Cronbach's alpha
Social skills	2.54 (1.76)	2.64 (2.06)	2.58 (1.89)	0.629
Attention switching	3.90 (2.03)	4.14 (1.86)	4.00 (1.96)	0.522
Attention to detail	5.49 (2.17)	5.22 (2.01)	5.38 (2.11)	0.635
Communication	2.57 (1.74)	2.64 (1.82)	2.60 (1.78)	0.505
Imagination	2.99 (1.78)	3.47 (1.82)	3.19 (1.81)	0.481
Total	17.23 (5.10)	18.17 (5.65)	17.63 (5.35)	0.710

[a] Sample sizes range from 415 to 437.
[b] Sample sizes range from 299 to 313.
[c] Sample sizes range from 714 to 750.

is ascertained, these (observed) differences cannot be assumed to tell something valid about the "real" differences in (latent) autism spectrum traits in men and women.

Relatively speaking, the cognitive aspects of AQ are the most endorsed ones, with respective means of 5.4 (Attention to detail) and 4 (Attention switching). Comparing our data to those of Baron-Cohen et al. (2001) and Hurst et al. (2007), the few ones who present comparable subscale scores, we see that: (1) our group scores higher for Imagination and at the upper range for Social skills; (2) scores for Attention to detail and communication are grossly in the same range than theirs; (3) Attention switching scores are lower in our subjects. Also in their work are cognitive characteristics more acknowledged than the social (Social skills and Communication) and Imagination ones, as is the case here. As about total AQ scores, comparing both to the original work of Baron-Cohen et al. (2001) and to the recently published review of Ruzich et al. (2015), it can be seen that our student group scores in the highest range of values but clearly below the lowest range of values reported for clinical groups.

Only the internal consistency of the total questionnaire is reasonably good; all other Cronbach alpha's are lower than 0.70 with an especially poor consistency for Imagination (0.481).

3.2. Test of measurement models with confirmatory factor analysis

Table 3 presents the CFA fit statistics and indices for the different measurement models tested for the AQ. First the results are presented that concern the original five-factor 50-item AQ, next modified models with varying number of factors on the 50-item set; finally, selected alternative models on shortened AQ versions were tested.

According to fit statistics and indices, the following can be summarized. The original five-factor model is not supported by our data as shown by bad fit statistics and indices (see Table 3). Fourteen items with (either statistically or quantitatively) not significant estimated parameter estimates (and consequently very low or nil R^2), unexpected negative estimated factor covariances – all implying the Attention to detail factor – as well as a series of not-modeled cross-loadings and correlated errors confirmed a misspecified model (not shown). Weak items mostly belong to Imagination, Communication and Attention switching (four "bad" items each), with two more superfluous items in Attention to detail; only in Social skills are all ten items significant (not shown).

Results remain very bad for all three alternative models, collapsing factors according to apparent content and keeping the total 50-item scale (fit statistics and indices remain at very low levels). Results are far from satisfying for Hoekstra's hierarchical model. The two remaining models on a five-factor shortened AQ score relatively better in terms of CFI, TLI and RMSEA – which get higher values for CFI and TLI and a lower value for RMSEA – than the complete questionnaire, but fit indices still stay at insufficient absolute levels.

3.3. Exploratory factor analysis

Two of the usual criteria of quality for an exploratory factorial analysis were satisfying: Bartlett's sphericity test was highly significant (χ^2 = 6,360,257; 1225 df, $p < 0.001$); Kaiser-Mayer-Olkin statistic was 0.768 – which is in the good range (Williams, Brown, & Onsmans, 2012). A third index though – the quality of representation of the solution – was weak in six items out of fifty, showing values inferior to 0.20 (meaning that the proportion of item variance explained by the total item set is low in one in every eight items). According to the parallel analysis, ten factors were retained,

Table 3
CFA fit statistics and indices for several measurement models for the AQ.

Models	Robust χ^2 [a]	CFI	TLI	RMSEA
Model 1 5 factors (50 items) – Baron-Cohen (295 df)	1806.270	0.504	0.590	0.081
Model 2 4 factors (50 items) (296 df)	1806.029	0.504	0.591	0.080
Model 3 3 factors (50 items) (298 df)	1991.259	0.444	0.545	0.085
Model 4 1 factor (50 items) (298 df)	2122.773	0.401	0.510	0.088
Model 5 2 factors (28 items) – Hoekstra hierarchical model[b] (162 df)	885.321	0.638	0.685	0.075
Model 6 5 factors (28 items) – Kloosterman model[b] (164 df)	763.226	0.725	0.764	0.068
Model 7 5 factors (39 items) – Lau model[b] (230 df)	1170.962	0.672	0.729	0.072

CFI: Comparative Fit Index; TLI: Tucker-Lewis Index; RMSEA: root mean square error of approximation; df: degrees of freedom.
[a] All χ^2 significant, $p < 0.00001$.
[b] See method.

Table 4
Ten-factor solution obtained with an exploratory factor analysis (ULS, Oblimin rotation).

Items	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10
47	0.635									
17	0.630									
38	*0.570*		*−0.513*	*0.419*						
44	0.529									
15	0.494									
1	0.488									
13	0.465									
9		0.656								
19		0.632								
41		0.373								
23		0.356								
39			0.548							
18			0.492							
7			0.437							
26	*0.345*			*0.537*	*−0.326*					
46				0.536						
22				0.474						
11				0.441						
35				0.352						
42				0.131						
48				0.309						
49					0.553					
29					0.499					
30					0.438					
25						−0.623				
2						−0.455				
43						−0.449				
34	*0.332*					*−0.430*				
8							−0.588			
20							−0.498			
14							−0.451			
21							−0.443			
3							−0.358			
50							−0.336			
12								0.662		
5								0.575		
28								0.454		
6		*0.352*						*0.426*		
36									0.674	
27									0.450	
45				*0.323*					*0.442*	
31									0.351	
10										0.527
32									*0.323*	*0.507*

Only loadings ≥ 0.30 are indicated. Items with no loading ≥ 0.30 have been omitted. Complex items are italized. ULS: unweighted least square estimation.

explaining 44.65% of the variance (forcing the solution to five factors, as in the measurement model of Baron-Cohen, only 29.8% of the variance was explained, and the concordance with the original item-to-factor structure was weak). When exploring the total correlation matrix of the fifty items (not shown), it is evident that many pairs of items are very weakly correlated. It then comes as no surprise that as much as ten factors are needed to roughly summarize the data (Table 4).

Looking at the factor matrix, several points can be made: (1) all factors are defined by at least three items with loading superior or equal to 0.30, except Factor 10; (2) two factors have only negative loadings: they define an "inverse" autistic dimension; (3) six items are "complex": loading on two or three factors, they are evidently not sufficiently specific of one particular autistic subdimension; (4) two factors comprise one item with a negative loading, but these two items (# 38 and # 26) are complex ones with their greatest loading being positive. Note that because the threshold for a significant loading was set at 0.30, other items are likely to be complex also but with smaller multiple loadings; (5) six items are absent from the solution (loading inferior to 0.30: # 4, # 16, # 24, # 33, # 37, and # 40); (6) as about the tentative content of the ten factors, the following can be proposed. Three new factors appear: Social skills and Communication converge in one main dimension (F1: "Sociophilia"), one ancillary factor (F4: "Social skills/anxiety") and a third very small dimension that seems to capture/subsume an "Empathic ability" (F9). The Imagination factor is restricted to six items. A subsample of four Communication items is also preserved at its own. Attention switching and Attention to detail are both subdivided in respectively two and three sets: for the latter we notice that one dimension is more focused on "Memory of details", the other one is more related to "Figures and patterns", while the third is concerned more literally by Attention to details. Attention switching is restricted to items assessing some attachment to "Repetitive behaviors". Interestingly, among those items loading on a different factor than expected, we identified three items (# 34, # 41 and # 46) the face content of which suggests some "error" in the measurement model as applied in French-speaking students in Belgium (e.g., we cannot subscribe to [item # 41] "I like to collect information about categories of things" as inversely related to an Imagination factor [item negatively keyed]). Might we want trying to test the interest of this adapted measurement model of the French AQ, it would then be necessary to assess its replicability in other independent large samples.

4. Discussion

The Autism Spectrum Quotient, French version, cannot be considered as structurally valid following our assessment with CFA in a large Belgian student sample. Our study – which is the first to translate and factorially assess the original scale in French – adds to a series, predominantly in English, that have cautioned against the use of the AQ as a multidimensional unambiguous tool for quantifying specific autistic traits in a non-clinical population.

Our figures for internal consistency are for the majority worse than those of Baron-Cohen et al. (2001) and Austin (2005), for example, but are not at all exceptional given that internal consistency was clearly a problem in several previous studies before, with some dimensions showing alpha's as bad as 0.34 and 0.42 for Imagination and Attention switching, respectively, in Hurst et al. (2007). One factor impacting consistency may relate to the characteristics of the sample studied, be it a clinically selected, a student or a more general population: clinical groups with definite autistic traits may exhibit a more coherent responding pattern because the clinical picture is more typical. However, this hypothesis does not seem to apply readily here because many research studies in the past did explore student samples. One interesting way to assess how a population variable impact the assessment tool is through invariance analyses; only one such study, by Murray et al. (2014), was recently carried over, but on an abridged version of the AQ: it indeed concluded to the presence of a bias in the estimation of latent autistic traits across populations. On the other hand, cultural more than "pure" translational aspects (as much effort was devoted for a rigorous and reliable translation of the AQ in our study) may also be possible sources of inconsistent responding with the French version. However, because low inconsistency is by no way anecdotal in the literature, our finding of poor consistency suggests too much heterogeneity in several dimensions – being defined by as much as ten items – and especially so for Imagination, Communication and Attention switching. This is further confirmed by the results of the EFA, which showed that no less than ten factors were needed to summarize the data, while more than one out of ten items did not contribute significantly to the solution: the original dimensions seem thus too heterogeneous leading to sets of items merging into new factors while twelve percent of items prove unnecessary and/or misleading. Three "new" factors concern social preference/avoidance, social skills and empathy, confirming the multifaceted relational (social and communicative) dimension of the AQ. The cognitive original dimensions (Attention switching and Attention to detail) also split into five items subsets. Considering the results of the EFA, the apparent relation of the AQ content to its underlying diagnostic rationale (based on the so-called "autistic triad" from the DSM-4, American Psychiatric Association [APA], 1994), assumed by the authors (Baron-Cohen et al., 2001) is too loose.

It therefore comes as no surprise that the five-factor measurement model as originally proposed by Baron-Cohen et al. (2001) did not fit our data at all, as the CFA showed. Not only were global fit indices bad, but local areas of bad fit were also evident mainly at the level of estimated parameter estimates: a series of standardized loadings (fourteen out of fifty) were either statistically or quantitatively not significant. Moreover, several estimated factor covariances were negative (exclusively implying the Attention to detail factor), which is theoretically inadequate, as the measurement model was originally conceived. However, our finding was empirically not unexpected. Indeed, in Wheelwright, Auyeung, Allison, and Baron-Cohen (2010), control parents scored higher than parents of children with an autism spectrum disorder on the "Attention to detail" dimension; also in Kloosterman et al. (2011), a higher Attention to detail was related to a better Communication and Imagination; lastly, Russel-Smith et al. (2011) proposed a three-factor solution of the AQ where increased Details/patterns is related to a better Communication. The Attention to detail dimension may thus be an ambiguous variable to interpret and as such may distort the assessment of autistic traits, at least in non-clinical groups: the ability to catch details can actually also constitute a cognitive asset in the social context. The examination of other sources of ill-fit showed cross-loadings and correlated errors being part of the discrepancy between the model and the data. Note that outside redundancy in item formulation (thus producing an exceedingly high covariance with no added content value), correlated errors may be inevitable in psychological assessment tools and therefore considered as trivial; measurement models with imposed uncorrelated error variances may therefore be unrealistic (Bentler & Chou, 1987).

Many studies, using either EFA or CFA, have put alternative measurement models forward; very few have been replicated but we thought it worthwhile to test several ones, emanating from the best methodological studies. Unfortunately, none fit our data, by large, according to usual criteria. Relatively speaking, the model of Kloosterman et al. (2011) showed the "least bad" CFI, TLI and RMSEA indices, but they were still far from adequate values. Clearly, the models with the complete fifty-item questionnaire, whatever the number of factors (models 2 through 4), scored the worst with our data. The suggestion would therefore be to favor shorter versions of the AQ, or at least, unless the internal consistency of revised subscales of the questionnaire is not improved, to only use the total AQ score. Of course, because our study is the first to test for the construct validity of the translated AQ in French, we caution against premature conclusion until further empirical evidence on large non-clinical groups is gathered.

Disclosure of interest

The authors declare that they have no competing interest.

Acknowledgement

Acknowledgements to Paul De Boeck and Catherine Dehon for their very helpful and friendly contributions, to Gwénolé Loas for his support as Head of the Psychiatric Dept.

Supported by the Belgian National Fund for Scientific Research (grants 1.5.123.04, 1.5.175.06 and 3.4.553.01.F).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.erap.2017.09.001

References

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: American Psychiatric Association.
American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington, VA: American Psychiatric Association.
Austin, E. J. (2005). Personality correlates of the broader autism phenotype as assessed by the Autism Spectrum Quotient (AQ). *Personality and Individual Differences*, *38*, 451–460. http://dx.doi.org/10.1016/j.paid.2004.04.022
Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): Evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders*, *31*, 5–17.
Bentler, P. M. (1990). Comparative fit indices in structural models. *Psychological Bulletin*, *107*, 238–246. http://dx.doi.org/10.1037/0033-2909.107.2.238
Bentler, P. M., & Chou, C.-P. (1987). Practical issues in structural modeling. *Sociological Methods & Research*, *16*, 78–117. http://dx.doi.org/10.1177/0049124187016001004
Brown, T. A. (2006). *Confirmatory factor analysis*. New York: Guilford Press.
Browne, M. W., & Cudeck, R. (1993). Alternate ways of assessing model fit. In K. A. Bollen, & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Newbury Park, CA: Sage.
Byrne, B. (2008). Testing for multigroup equivalence of a measuring instrument: A walk through the process. *Psicothema*, *20*, 872–882.
Gregorich, S. E. (2006). Do self-reports allow meaningful comparisons across diverse population groups? Testing measurement invariance using the

confirmatory factor analysis framework. *Medical Care*, *44*, S78–S94. http://dx.doi.org/10.1097/01.mlr.0000245454.12228.8f
Hayton, J. C., Allen, D. G., & Scarpello, V. (2004). Factor retention decisions in exploratory factor analysis: A tutorial on parallel analysis. *Organizational Research Methods*, *7*, 191–205. http://dx.doi.org/10.177/1094428104263675
Hoekstra, R. A., Bartels, M., Cath, D. C., & Boomsma, D. I. (2008). Factor structure, reliability, and criterion validity of the autism-spectrum quotient (AQ): A study in Dutch population and patient groups. *Journal of Autism and Developmental Disorders*, *38*, 1555–1566. http://dx.doi.org/10.1007/s10803-008-0538-x
Hoekstra, R. A., Vinkhuyzen, A. A. E., Wheelwright, S., Bartels, M., Boomsma, D. I., Baron-Cohen, S., et al. (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. http://dx.doi.org/10.1007/s10803-010-1073-0
Horn, J. L. (1965). A rationale and test for the number of factors in factor analysis. *Psychometrika*, *30*, 179–185. http://dx.doi.org/10.1007/BF02289447
Hu, S. (2007). *Akaike information criterion*. (Powerpoint presentation, retrieved as www4.ncsu.edu/~shu3/Presentation/AIC.pdf, on Feb 5, 2015)
Hurst, R. M., Mitchell, J. T., Kimbrel, N. A., Kwapil, T. K., & Nelson-Gray, R. O. (2007). Examination of the reliability and factor structure of the Autism Spectrum Quotient (AQ) in a non-clinical sample. *Personality and Individual Differences*, *43*, 1938–1949. http://dx.doi.org/10.1016/j.paid.2007.06.012
Jackson, D. L., Gillaspy, J. A., Jr., & Purc-Stephenson, R. (2009). Reporting practices in confirmatory factor analysis: An overview and some recommendations. *Psychological Methods*, *14*, 6–23. http://dx.doi.org/10.1037/a0014694
Kempenaers, C., De Boeck, P., Dehon, C., Braun, S., & Linkowski, P. (2014). Metric invariance of the French version of the contingencies of self-worth scale. *Comprehensive Psychology*, *3*, 22. http://dx.doi.org/10.2466/08.CP.3.22
Ketelaars, C., Horwitz, E., Sutema, S., Bos, J., Wiersma, D., Minderaa, R., et al. (2008). Brief report: Adults with mild autism spectrum disorders (ASD): Scores on the Autism Spectrum Quotient (AQ) and comorbid psychopathology. *Journal of Autism and Developmental Disorders*, *38*, 176–180. http://dx.doi.org/10.1007/s10803-007-0358-4
Kloosterman, P. H., Keefer, K. V., Kelley, E. A., Summerfeldt, L. J., & Parker, J. D. A. (2011). Evaluation of the factor structure of the autism-spectrum quotient. *Personality and Individual Differences*, *50*, 310–314. http://dx.doi.org/10.1016/j.paid.2010.10.015
Kose, S., Bora, E., Erermis, S., Ozbaran, B., Bildik, T., & Aydin, C. (2013). Broader autistic phenotype in parents of children with autism: Autism Spectrum Quotient-Turkish version. *Psychiatry and Clinical neurosciences*, *67*, 20–27. http://dx.doi.org/10.1111/pcn.12005
Kuennsberg, R., Murray, A. L., Booth, T., & McKenzie, K. (2012). Structural validation of the abridged Autism Spectrum Quotient-Short Form in a clinical sample of people with autism spectrum disorders. *Autism*, http://dx.doi.org/10.1177/1362361312467708 (Advance online publication)
Lai, M.-C., Lombardo, M. V., Chakrabardi, B., & Baron-Cohen, S. (2013). Subgrouping the autism "spectrum": Reflections on DSM-5. *PLOS Biology*, *11*, e1001544.
Lau, W. Y-P., Gau, S. S-F., Chiu, Y.-N., Wu, Y.-Y., Chou, W.-J., Liu, S.-K., et al. (2013). Psychometric properties of the Chinese version of the Autism Spectrum Quotient (AQ). *Research in Developmental Disabilities*, *34*, 294–305. http://dx.doi.org/10.1016/j.ridd.2012.08.005
Lau, W. Y-P., Kelly, A. B., & Peterson, C. C. (2013). Further evidence on the factorial structure of the Autism Spectrum Quotient (AQ) for adults with and without a clinical diagnosis of autism. *Journal of Autism and Developmental Disorders*, *43*, 2807–2815. http://dx.doi.org/10.1007/s10803-013-1827-6
Ledesma, R. D., & Valero-Mora, P. (2007). Determining the number of factor to retain in EFA: An easy-to-use computer program for carrying out parallel analysis. *Practical Assessment, Research and Evaluation*, *121* (Available online: http://pareonline.net/getvn.asp?v=12&n=2, retrieved on Feb 6, 2015)
Lepage, J.-F., Lortie, M., Taschereau-Dumouchel, V., & Théoret, H. (2009). Validation of French-Canadian versions of the Empathy Quotient and Autism Spectrum Quotient. *Canadian Journal of Behavioural Science*, *41*, 272–276. http://dx.doi.org/10.1037/a0016248
Lord, C., & Jones, R. M. (2012). Re-thinking the classification of autism spectrum disorders. *Journal of Child Psychology and Psychiatry*, *53*, 490–509. http://dx.doi.org/10.1111/j.1469-7610.2012.02547.x
Mohammadi, M. R., Zarafshan, H., & Ghasempour, S. (2012). Broader autism phenotype in Iranian parents of children with autism spectrum disorders vs. normal children. *Iranian Journal of Psychiatry*, *7*, 157–163.
Murray, A. L., Booth, T., McKenzie, K., Kuenssberg, R., & O'Donnell, M. (2014). Are autistic traits measured equivalently in individuals with and without an autism spectrum disorder? An invariance analysis of the Autism Spectrum Quotient Short Form. *Journal of Autism and Developmental Disorders*, *44*, 55–64. http://dx.doi.org/10.1007/s10803-013-1851-6
Muthén, L. K., & Muthén, B. O. (1998–2004). *Mplus user's guide* (3rd ed.). Los Angeles: Author.
Muthén, L. K., & Muthén, B. O. (1998–2009). *Mplus program and add-on, version 5.21*.
Pisula, E., Kawa, R., Szostakiewicz, L., Luck, I., Kawa, M., & Rynkiewicz, A. (2013). Autistic traits in male and female students and individuals with high functioning autism spectrum disorders measured by the Polish version of the Autism-Spectrum Quotient. *Plos One*, *8*(9), e75236. http://dx.doi.org/10.1371/journal.pone.0075236 (eCollection 2013)
Robel, L., Rousselot-Pailley, B., Fortin, C., Levy-Rueff, M., Golse, B., & Falissard, B. (2014). Subtreshold traits of the broad autistic spectrum are distributed across different subgroups in parents, but not siblings, of probands with autism. *European Child and Adolescent Psychiatry*, *23*, 225–233. http://dx.doi.org/10.1007/s00787-013-0451-5
Rousselot-Pailley, B., Fortin, C., Golse, B., Falissard, B., & Robel, L. (2011). L'autoquestionnaire FAQ: Un outil valable pour le repérage des endophénotypes des parents d'enfants autistes [The FAQ self-report is a valid instrument to characterize endophenotypes of the autistic spectrum in parents of children with autism]. *L'Encéphale*, *37*, 191–198. http://dx.doi.org/10.1016/j.encep.2010.08.014
Russel-Smith, S. N., Mayberry, M. T., & Bayliss, D. M. (2011). Relationships between autistic-like and schizotypy traits: An analysis using the Autism Spectrum Quotient and Oxford-Liverpool Inventory of feelings and experiences. *Personality and Individual Differences*, *51*, 128–132. http://dx.doi.org/10.1016/j.paid.2011.03.027
Ruta, L., Mazzone, D., Mazzone, L., Wheelwright, S., & Baron-Cohen, S. (2011). The Autism-Spectrum Quotient-Italian version: A cross-cultural confirmation of the broader autism phenotype. *Journal of Autism and Developmental Disorders*, *42*, 625–633. http://dx.doi.org/10.1007/s10803-011-1290-1
Ruzich, E., Allison, C., Smith, P., Watson, P., Auyeung, B., Ring, H., et al. (2015). Measuring autistic traits in the general population: A systematic review of the Autism-Spectrum Quotient (AQ) in a nonclinical population sample of 6,900 typical adult males and females. *Molecular Autism*, *6*, 2. http://dx.doi.org/10.1186/2040-2392-6-2 (Advance online publication)
Sonié, S., Kassai, B., Pirat, E., Masson, S., Bain, P., Robinson, J., et al. (2013). The French version of the Autism-Spectrum Quotient in adolescents: A cross-cultural validation study. *Journal of Autism and Developmental Disorders*, *43*, 1178–1183. http://dx.doi.org/10.1007/s10803-012-1663-0
Sonié, S., Kassai, B., Pirat, E., Masson, S., Bain, P., Robinson, J., et al. (2011). Version française des questionnaires de dépistage de l'autisme de haut niveau ou du syndrome d'Asperger chez l'adolescent : Quotient du spectre de l'autisme, quotient d'empathie, et quotient de systématisation. Protocole et traduction des questionnaires [French version of screening questionnaire for high-functioning autism or Asperger syndrome in adolescent: Autism Spectrum Quotient, Empathy Quotient and Systemizing Quotient. Protocol and questionnaire translation]. *Presse Médicale*, http://dx.doi.org/10.106/j.lpm.2010.07.016 (Advance online publication)
SPSS. (2008–2010). *SPSS versions 17.0-19.0*. Chicago Il: SPSS Inc.
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*, 224–228. http://dx.doi.org/10.1016/j.paid.2009.03.004
Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, *38*, 1–10. http://dx.doi.org/10.1007/BF02291170
Volkmar, F. R., State, M., & Klin, A. (2009). Autism and autism spectrum disorders: Diagnostic issues for the coming decade. *Journal of Child Psychology and Psychiatry*, *50*, 108–115. http://dx.doi.org/10.1111/j.1469-7610.2008.02010.x
Wakabayashi, A., Baron-Cohen, S., Wheelwright, S., & Tojo, Y. (2006). The Autism-Spectrum Quotient (AQ) in Japan: A cross-cultural comparison. *Journal of Autism and Developmental Disorders*, *36*, 263–270. http://dx.doi.org/10.1007/s10803-005-0061-2
Watkins, M. (2005). *Determining Parallel Analysis criteria" (2005) + program (Watkins, M.W., 2000). Monte Carlo PCA for parallel analysis [computer software]*. State College, PA: Ed & Psych Associates.
Wheelwright, S., Auyeung, B., Allison, C., & Baron-Cohen, S. (2010). Defining the broader, medium and narrow autism phenotype among parents using the Autism Spectrum Quotient (AQ). *Molecular Autism*, http://dx.doi.org/10.1186/2040-2392-1-10 (Advance online publication)
Williams, Brown, & Onsmans. (2012). Exploratory factor analysis: A five-step guide for novices. *Australasian Journal of Paramedicine*, *8*(3) (Retrieved from http://ro.ecu.edu.au/jephc/vol8/iss3/1)
Woodbury-Smith, M. R., Robinson, J., & Baron-Cohen, S. (2005). Screening adults for Asperger Syndrome using the AQ: A preliminary study of its diagnostic validity in clinical practice. *Journal of Autism and Developmental Disorders*, *35*, 331–335. http://dx.doi.org/10.1007/s1083-005-3300-7