Research in Autism Spectrum Disorders 64 (2019) 39–48

Contents lists available at ScienceDirect

Research in Autism Spectrum Disorders

journal homepage: www.elsevier.com/locate/rasd

Psychometric properties, factor structure and cross-cultural validity of the quantitative CHecklist for autism in toddlers (Q-CHAT) in an Italian community setting

Liliana Ruta[a,b,*], Giuseppe M. Arduino[c], Antonella Gagliano[d], Fabio Apicella[b], Elisa Leonardi[a,d], Francesca I. Famà[a], Natasha Chericoni[b], Valeria Costanzo[b], Nazarena Turco[c], Gennaro Tartarisco[a], Carrie Allison[e], Simon Baron Cohen[e], Giovanni Pioggia[a], Flavia Chiarotti[f], Filippo Muratori[b,g]

[a] Institute of Applied Sciences and Intelligent Systems, "Eduardo Caianiello", National Research Council of Italy, Messina, Italy
[b] Department of Developmental Neuroscience, Stella Maris Scientific Institute, Calambrone, Pisa, Italy
[c] Centro Autismo e Sindrome di Asperger ASLCN1, Mondovì, Italy
[d] University of Messina,Italy
[e] Autism Research Centre, Department of Psychiatry, University of Cambridge, Cambridge, United Kingdom
[f] Center for Behavioral Sciences and Mental Health, National Institute of Health, Rome, Italy
[g] University of Pisa, Pisa, Italy

ARTICLE INFO

Keywords:
Autism spectrum disorders
Screening
Cross-cultural validity
Q – CHAT

ABSTRACT

Background: In the past decade, several screening instruments have been developed to detect toddlers at risk for autism, both in clinical and unselected samples.We aimed to assess the validity and cross-cultural stability of the Quantitative CHecklist for Autism in Toddlers (Q – CHAT), a quantitative and normally distributed measure of autistic traits, during routine paediatric developmental surveillance in a large Italian community of toddlers.

Methods: A group of n = 2400 unselected Italian toddlers were screened by paediatricians during routine well – CHild visits between 18 and 24 months,in three Italian regions, representative of the North, Centre and South of Italy. The psychometric properties and factor structure of the Q – CHAT were explored and compared with published literature from Asia and the UK.

Results: The total Q – CHAT scores in our sample confirmed a normal distribution. Mean Q – CHAT scores across regions were consistent with those reported in UK studies and likewise, a gender effect was found, with boys scoring higher than girls.We also found a remarkable consistency with the three-factor structure of the Asian Q – CHAT. Internal consistency was acceptable for both the Q – CHAT total score and the three factors.

Conclusions: Our results confirm the dimensional distribution of autistic traits and related symptoms in unselected populations across different cultures. The Q – CHAT is a valid and reliable dimensional screening instrument to be used within a routine paediatric setting during well-child assessments.The main normative data and factor structure replicated in our sample

* Corresponding author: Institute of Applied Sciences and Intelligent Systems, "Eduardo Caianiello" (ScienceApp) – National Research Council of Italy (CNR), via Torre Bianca, SNC, Istituto Marino, Pad. 4, 98164 Messina, Italy.
E-mail addresses: l.ruta@cnr.it, liliana.ruta@gmail.com (L. Ruta).

https://doi.org/10.1016/j.rasd.2019.04.004
Received 3 December 2018; Received in revised form 9 April 2019; Accepted 11 April 2019
Available online 20 April 2019
1750-9467/

1. Introduction

Autism Spectrum disorders (ASD) are a group of neurodevelopmental conditions characterized by a dyad of symptoms related to social communication deficits along with restricted, repetitive interests and behaviours and atypical sensory interests (Ref. American Psychiatry Association, 2013). The shift from a categorical system proposed in the Diagnostic and Statistical Manual of Mental Disorders, 4th Edition, Text Revision (DSM 4 – TR) towards the integration of a dimensional approach to diagnosis made by the DSM 5 (Mandy, Charman, & Skuse, 2012; Wing, Gould, & Gillberg, 2011) emphasizes the concept of autism as a continuum of symptoms and traits within the autism spectrum and in the general population (Constantino & Charman, 2016; Constantino, 2011). The broad spectrum of a typicalities and symptom severity seen in ASD reflects brain processes and functions that develop differently from very early in life, leading to different developmental trajectories (Elsabbagh & Johnson, 2010; Johnson, Jones, & Gliga, 2015, 2016). Specific tools designed to capture early warning signs of ASD have been developed in the past two decades, with different results in terms of validity, feasibility and replicability, depending on sampling, screening procedures and methodologies (Charman & Gotham, 2013; Towle & Patrick, 2016). Two of the most well known and used screening instruments for autism, the pioneering CHecklist for Autism in Toddlers (CHAT) (Baird et al., 2000; Baron-Cohen, Allen, & Gillberg, 1992) followed by the Modified −CHecklist for Autism in Toddlers (M−CHAT) (Robins, Fein, Barton, & Green, 2001) both demonstrated adequate positive predictive values (PPV = 0.59 for the CHAT and PPV ranging between 0.54 and 0.75 in low and high risk populations respectively for the MCHAT) and sufficient consistency in different cohorts (García-Primo et al., 2014; Robins et al., 2014). However, both the instruments are designed to get a “yes/no” answer to each item according to a more categorical, “all-or-nothing” approach to autism detection. This approach has the potential limitation of picking up the extreme cases, missing the milder cases and consequently contributing to the low sensitivity demonstrated by the CHAT (Baird et al., 2000). To overcome this limitation, a revised instrument has been developed which aims to quantify autistic traits, adopting a dimensional approach, which is in line with the DSM-V. The Quantitative CHecklist for Autism in Toddlers (Q−CHAT). The Q−CHAT is a 25 item parent-report questionnaire. Items are related to developmentally appropriate ASD associated behaviours, including joint attention, pretend play, social communication, repetitive behaviours, sensory interests, and language development.

In the first study (2008), Allison et al. administered the Q−CHAT to n = 779 caregivers of unselected toddlers and to n = 160 caregivers of young children with a diagnosis of ASD in the UK. The Q−CHAT demonstrated a near-normal distribution in the population and good internal consistency (α value of 0.67) with very good intraclass correlation for test-retest reliability after 1 month (r = .82). A main effect of group and gender was found, with the ASD group scoring significantly higher than the unselected group and males scoring significantly higher than females within the unselected group. In a subsequent study, Allison et al. examined the validity of a short form of the Q−CHAT (Q−CHAT-10), identifying the ten best predicting items of an ASD diagnosis (Allison, Auyeung, & Baron-Cohen, 2012). The Q−CHAT-10 demonstrated high internal consistency (Cronbach 's alpha = .88) and was highly correlated with the original 25-item Q−CHAT (r = .79).

The Q−CHAT has also been used to assess autistic traits at 18–24 months in relation to prenatal sex steroid hormones exposition (Auyeung, Taylor, Hackett, & Baron-Cohen, 2010, 2012) and in a cohort of n = 141 very preterm children to explore social communication skills and ASD traits (Wong, Huertas-Ceballos, Cowan, & Modi, 2014). In the latter study, preterm children at 24 months of age demonstrated significantly higher Q−CHAT scores compared to published population scores (Allison et al., 2008). More recently, the psychometric properties and factor structure of the QCHAT was tested in a population-based sample of Singaporean toddlers at 18 (n = 368) and 24 (n = 396) months (Magiati et al., 2015). This is the first study investigating the cross-cultural stability of the Q−CHAT, since all the previous studies have been conducted in the UK. The Q−CHAT was found to have normally distributed scores and good stability between 18 and 24 months. A three-factor structure was found (social/communication, speech/language and non-social/behavioural traits). Furthermore, Q−CHAT total scores at 18 and 24 months were significantly positively correlated with raw scores from the child behaviour checklist’s (CBCL) Internalizing Problems, Withdrawn and Pervasive Developmental Disorders (PDD) subscales at 24 months.

Currently, there is no standardized screening approach for early identification of ASD in Italy. Various screening instruments (mainly the CHAT and the M−CHAT) have been used in different Italian regions but very inconsistently, in circumscribed clinical settings and mostly if the paediatrician already had a concern about the child. Hence, the main goal of the current study was to conduct a systematic population based screening for ASD using the Q−CHAT within routine paediatric developmental surveillance in a large community of toddlers representative of different regions in the North, Centre and South of Italy. Furthermore, we aimed to assess the psychometric properties and the factor structure of the Q−CHAT and to explore its cross-cultural validity comparing score distribution in Italy with respective score distributions reported in published literature from Asia and the UK.

2. Methods

2.1. Participants

A group of 2400 unselected Italian toddlers aged between 18 and 24 months (M:F = 1267:1133, mean age (standard deviation) = 23.2 (5.2) months) were screened by paediatricians during well-child check-up visits.

2.2. Ethical considerations

The study was approved by the local Ethic Committees in each region and all the participants signed a written informed consent form to be enrolled in the study.

2.3. Procedure

The study was conducted as part of a large population based screening programme supported by the Ministry of Health and by the Tuscany Region (GR-2010 – 2319668) and carried out in three regions (Piedmont, Tuscany and Sicily), representative of the North, Centre and South of Italy.

Paediatricians were contacted via National Health Service lists and paediatric associations. Those who agreed to participate in the project (N = 158, 43% of the paediatricians contacted) were asked to administer the Q – CHAT systematically during the 18–24 months well – CHild visits. The Q – CHAT questionnaires were sent to the paediatricians' practices by mail and retrieved on average every two months either by a member of the research team or via courier. A total of 6080 questionnaires were sent and 2478 (40.7%) were returned. 78 (3.1%) QCHAT questionnaires had seven or more Q – CHAT items missing and, in accordance with Allison et al. (2008), were excluded from analyses.

2.4. Validation of the italian Q-CHAT

The Q – CHAT is a 25-item caregiver-report screening instrument for measuring autistic traits in toddlers. Items are rated on a 5-point Likert scale (0–4) with higher ratings indicating more autistic traits. Thirteen items are reverse scored. Items are summed to give a Q – CHAT total score, ranging from 0 to 100. The scoring procedure used in the study was exactly the same as that adopted in the original Q – CHAT study by Allison et al. (2008). Functional and conceptual equivalence of words between English and Italian, was ensured by back-translation as following: first, the Q – CHAT was translated into Italian by one of the authors (LR), a native Italian speaker. After that, a native English speaker who was naive to the Q – CHAT translated the questionnaire back into English. This translated version of the questionnaire was compared with the original version of the Q – CHAT by one of the authors of the published the Q – CHAT (CA) and points of divergence were discussed until agreement was reached. Furthermore, as a supplement procedure, a probe technique was used as following: a group of 20 parents of typically developing children aged 18–36 months, randomly selected in a public preschool, were administered the Q – CHAT and then asked to explain why they answered as they did to individual items. Hence, the Italian Q – CHAT was judged reliably equivalent to the original English version and presented in the same forced- choice format.

The Q – CHAT questionnaires were presented as parent-friendly booklets (B1 format) and simple cartoons depicting the relative question accompanied each Q – CHAT item. The child's personal history and the main demographic and family information were included in the booklet.

2.5. Statistical analyses

All statistical analyses were conducted using the Statistical Package for Social Sciences (SPSS), Version 21.0 for Mac and the STATA Release 8.1. As per Allison et al. (2008), if less than 7 items were missing or ambiguously answered a conservative score of '0′ was attributed. Of the 2400 included questionnaires, 192 (8%) Q – CHAT had less than 7 items missing. The items with the most missing data were items 7, 10 and 21 (1% for all three items).

Descriptive statistics was conducted on the main demographic and personal history information accounting for group, gender and region. Categorical variables were analysed using the Chi-squared test while quantitative variables were analysed using either the Student *t*-test or the Analysis of Variance (ANOVA). The Shapiro-Wilk test and the Standardized Normal Probability plot were used to assess normality in the Q – CHAT scores distribution. A multiple linear regression model was used to assess the contribution of age, gender, parental age and education, gestational age, birth weight and perinatal problems to the Q – CHAT total scores. Q – CHAT item distribution and item-total correlations were also examined using Spearman's rho non-parametric analyses. The Kaiser-Meyer-Olkin (KMO) statistic was computed on the 25 items of the Q – CHAT in order to verify their suitability for the Exploratory Factor Analysis (EFA). Based on the results of KMO statistic, the EFA was performed on the Q – CHAT scores, as done by Magiati et al. (2015), to investigate the factor structure of the Italian Q – CHAT, and compare it with the factor structure of the Asian Q – CHAT. The Iterated Principal Axis Factoring was used to estimate factor eigenvalues. Kaiser criteria (eigenvalue greater than 1) and scree plot (indicating the point of inflexion of the eigenvalues) were used to determine the number of factors to be retained. The EFA was then rerun forcing the number of factors. Oblique rotation was applied to determine the contribution of single items to the extracted factors. Pearson correlation coefficient was used to calculate the correlation between principal factors extracted from EFA, and Cronbach's alpha was computed to estimate the internal consistency of items contributing to the overall Q – CHAT scores and to the factors extracted by EFA. Finally, the Italian Q – CHAT mean scores were compared with the Q – CHAT mean scores reported in published literature from both Asian and UK samples using a one-way ANOVA.

3. Results

3.1. Demographic and clinical characteristics of the sample

Table 1 reports the demographic and clinical characteristics of the final sample (N = 2400) in the three regions examined.

Although most p-values demonstrated a significant difference due to the large sample, the effect sizes are all very small and indeed the differences between the regions are negligible. Considering the whole sample, the mean age of the screened children was 23.2 months (range 11–81 months, SD = 5.2 months) and the male to female distribution was 1273:1132 (53% males vs 47%

Table 1
Demografic and clinical characteristics of the pediatric-screened sample (N = 2400).

	PIEMONTE	TOSCANA	SICILIA	Statistics	Effect size
N	850	104	1446	P-values	
Age (mean, SD)	22.8 (4.3)	20.1 (3.7)	23.7 (5.8)	$p < .001$	$\eta^2 = 0.020$
Gender [a] M:F (N, %)	449:401 (53-47%)	39:65 (38-62%)	775:666 (54-46%)	$p = .006$	$V = 0.066$
Q-CHAT total	24.9 (7.6)	25.6 (8.8)	26.3 (7.8)	$p < .001$	$\eta^2 = 0.008$
Personal history					
Term pregnancy (N, %)	755 (95%)	71 (85.5%)	1197 (87.6%)	$p < .001$	$V = 0.121$
Pregnancy complications (N, %)	15 (2%)	5 (7%)	119 (9.5%)	$p < .001$	$V = 0.146$
Birth weight gr. (Mean, SD)	3239.3 (502.1)	3163.4 (663.5)	3204.6 (493.2)	$p = .2$	$\eta^2 = 0.002$
APGAR score (Mean, SD)	8.8 (0.9)	9.1 (0.9)	9.4 (0.8)	$p < .001$	$\eta^2 = 0.107$
Perinatal problems (N, %)	23 (3%)	4 (5%)	170 (13%)	$p < .001$	$V = 0.166$
Nursery school (N, %)	224 (26.6%)	32 (32%)	357 (25%)	$p = .2$	$V = 0.035$
Special education teacher at school (N, %)	3 (0.4%)	2 (2.5%)	5 (0.4%)	$p = .02$	$V = 0.059$
SES					
Education mother (N, %)				$p < .001$	$V = 0.298$
Pre-primary, primary	191 (22.9%)	44 (43.1%)	424 (29.9%)		
Secondary	359 (43%)	22 (21.6%)	110 (7.8%)		
Bachelor, Master Degree, PhD	285 (34.1%)	36 (35.3%)	886 (62.4%)		
Occupation mother (N, %)				$p < .001$	$V = 0.230$
Not working	206 (24.9%)	42 (41.2%)	828 (58.7%)		
Manual, technical	56 (6.8%)	7 (6.9%)	37 (2.6%)		
Clerical, sales	265 (32%)	31 (30.4%)	274 (19.4%)		
Administrative, professional, management	300 (36.3%)	22 (21.6%)	272 (19.3%)		
Ethnicity mother				$p = .1$	$V = 0.042$
Caucasian	822 (97.4%)	100 (98%)	1410 (98.6%)		
Asiatic	/	/	/		
African	/	/	/		
Other	22 (2.6%)	2 (2%)	20 (1.4%)		
Education father				$p < .001$	$V = 0.207$
Pre-primary, primary	327 (40.3%)	51 (53.7%)	478 (36.5%)		
Secondary	215 (26.5%)	25 (26.3%)	107 (8.2%)		
Bachelor, Master Degree, PhD	269 (33.2%)	19 (20.0%)	725 (55.3%)		
Occupation father				$p < .001$	$V = 0.195$
Not working	19 (2.4%)	4 (4.3%)	232 (17.3%)		
Manual, technical	346 (43.1%)	35 (37.2%)	359 (26.8%)		
Clerical, sales	148 (18.4%)	26 (27.7%)	367 (27.4%)		
Administrative, professional, management	290 (36.1%)	29 (30.8%)	381 (28.5%)		
Ethnicity father				$p = .05$	$V = 0.051$
Caucasian	812 (98.1)	93 (96.9)	1363 (99.1%)		
Asiatic	/	/	/		
African	/	/	/		
Other	16 (1.9%)	3 (3.1%)	13 (0.9%)		

N = number of subjects; SD = standard deviation; η^2 = eta squared represents the percentage of the outcome variance explained by the grouping factor (in our case, the center).
V = Cramer's V; the reference values depend on d*, that is the minimum dimension of the contingency table, i.e. d* = minimum between r-1 and c-1 (where r and c are the number of rows and columns of the table). For d* = 1, 0.10 = small, 0.30 = medium, and 0.50 = large effect size. For d* = 2, 0.07 = small, 0.21 = medium, and 0.35 = large effect size. All effect sizes are small, except for mother's and father's education and occupation, where there is a medium effect size.
[a] N = 2395.

females). 222 (9.9%) children were born preterm (< 38 weeks of gestation). Perinatal problems were reported in 197 (9.1%) of the sample. Mean birth weight was 3215 gr (standard deviation = 504 g). 613 (25.8%) of the children attended a nursery school and 7 (1.1%) received the support of a special education teacher at school. Parents' mean age at the time of the study was 33.3 years (standard deviation = 5.5 years) for mothers and 36.4 years (standard deviation = 6.1 years) for fathers. As for parents' education distribution we found that 28% of mothers and 38.6% of fathers had a primary level of education (less than or equal to 8 years education), 20.8% of mothers and 15.7% of fathers had a secondary level of education and 51.2% of mothers and 45.7% of fathers had a high level of education (either bachelor, master degree or PhD). According to the Italian 2015 census reported by the Italian National Institute of Statistics (ISTAT, http://www.istat.it/it/files/2016/12/ItaliaCifre2016.pdf), 38.7% of females and 42.3% of males in the Italian population have a primary level of education, 41.5% of females and 42.3% of males received a secondary level of education and 19.8% of females and 15.3% of males reached a high level of education. Comparing the distribution of education with that observed in the general population, in our sample we found a significantly lower percentage of participants with a secondary level of education (Cohen's w = 0.4 and Cohen's w = 0.5 for mothers and fathers respectively), while participants with a high level of education were over-represented (Cohen's w = 0.79 and 0.84 for mothers and fathers, respectively). While reporting and considering these differences in parents' education, we did not exclude anyone on the basis of education from further analysis. With respect to the child's gender distribution and birth weight, our sample was representative of the Italian Population (Cohen's w = 0.02 and 0.11,

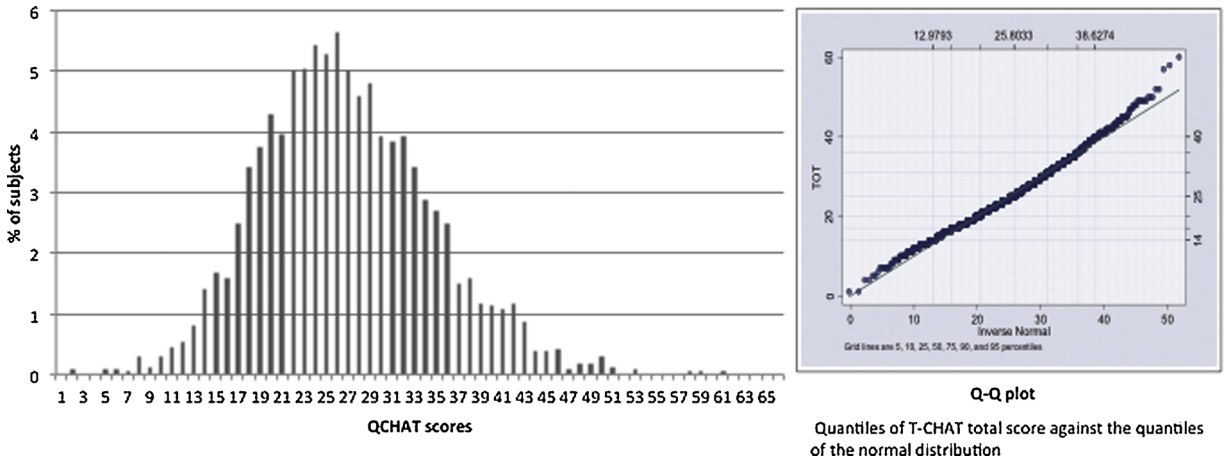

Fig. 1. Distribution and standardized normal probability plot of the Q − CHAT total score in the screened population (n = 2400).

respectively). In a multiple linear regression model, we found that gender (Coeff = -0.95, p = .01), maternal age and education low vs high (Coeff = -0.18, p < .001 and Coeff = 1.04, p = .04), paternal age and education low vs high (Coeff = 0.11, p = .02 and Coeff = 0.93, p = .05) significantly predicted Q − CHAT total scores. Conversely, mother's and father's education intermediate vs high (Coeff = -0.16, p = .76 and 0.51, p = .38), child's age (Coeff = -0.04, p = .24), at-term vs pre-term pregnancy (Beta = -0.10, p = .89), birth weight (Coeff = 0.0003, p = .49) and perinatal problems (Coeff = 0.86, p = .19) were not associated with Q − CHAT total scores.

3.2. Q − CHAT total score distribution, item score distribution and item-total correlations

As shown in Fig. 1 the Q − CHAT total scores are normally distributed in the screened population.

The mean score on the Q − CHAT was 25.8 (SD = 7.8, range: 1–60). The mean score for boys was 26.3 (N = 1263, SD = 7.8, range 4–60) and the mean score for girls was 25.3 (N = 1132, SD = 7.8, range 1–58). This difference was statistically significant (t (2393) = 3.14, p = 0.002, equal variances assumed), although with a very small effect size (Cohen's d = 0.13).

Table 2 shows the distribution of scores for each scored item (i.e. 0–4). All items were significantly correlated with the Q − CHAT

Table 2
Item score distribution and mean item scores.

Q-CHAT Item	Scores %				
	0	1	2	3	4
1. Look when call name	74	22.1	3.2	0.4	0.3
2. Eye contact	66.2	31.8	1.5	0.4	0.1
3. Line objects up[a]	13.5	14.3	29.9	26.5	15.8
4. Understand child's speech	20.5	42.6	24.3	5.8	6.8
5. Protoimperative pointing	80.3	12.3	4.1	1.3	2
6. Protodeclarative pointing	83	9.8	4.2	1.4	1.6
7. Interest maintained by spinning object[a]	38	43.3	12.1	5	1.6
8. Number of words[a]	17.1	14.4	39.1	26.1	3.3
9. Pretend play	66.5	22.3	7.5	1.3	2.4
10. Follow a look	64.6	24.1	7.4	1.3	2.6
11. Sniff/lick unusual objects[a]	12.1	29.8	14.8	20.5	22.8
12. Use of hand as tool[a]	22	9.7	12.6	30.9	24.8
13. Walk on tiptoes[a]	44.2	23.2	24	7.4	1.2
14. Adapt to change in routine	47.2	44.5	5.5	2	0.8
15. Offer comfort	31	34	23.5	7.7	3.8
16. Do same thing over and over again[a]	14.6	12.2	16.4	31.8	25
17. Typicality of first words	65.9	27.7	2.3	0.8	3.3
18. Echolalia[a]	8.5	4.3	8	28.3	50.9
19. Gestures	76.5	19.4	2.5	0.6	1
20. Unusual finger movements[a]	77.6	9	5.3	6.4	1.7
21. Check reaction	38.4	36.5	18.6	4.6	1.9
22. Maintenance of interest[a]	51.8	30	13.2	3.4	1.6
23. Twiddle objects repetitively[a]	66.2	12.4	9.1	9.4	2.9
24. Oversensitive to noise[a]	50.9	26	15.2	4.3	3.6
25. Stare at nothing with no purpose[a]	88.2	7.5	2.6	1.3	0.4

[a] Reverse-scored items.

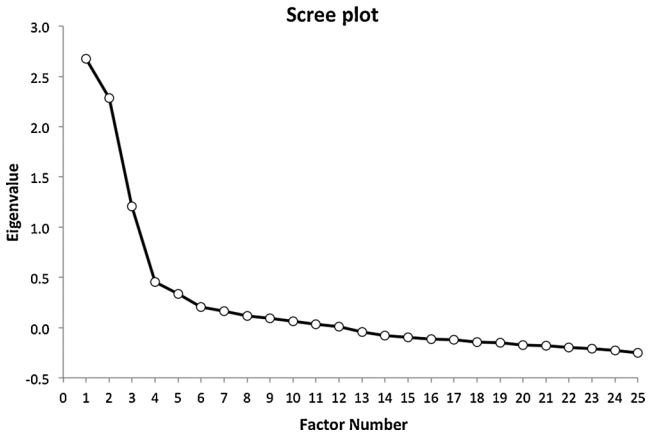

Fig. 2. Scree plot of eigenvalues derived from exploratory factor analysis of the correlation matrix for 25 items of the Q−CHAT.

total score, with large to moderate effect sizes ($0.40 < rho < 0.57$) for items 23, 16, 12, 07, 20, 11 and small effect sizes ($0.20 < rho < 0.40$) for items 03, 13, 17, 25, 02, 22, 04, 08, 24, 10, 01, and 14. Very low effect sizes were found for items 06, 19, 09, 15 ($0.10 < rho < 0.20$) and for items 18, 05, 21 ($-0.05 < rho < 0.10$).

3.3. Exploratory factor analysis of the Q−CHAT

An exploratory factor analysis of the Q−CHAT was already carried out by Magiati et al. (2015). We replicated the same procedure (EFA) to explore the factor structure of the QCHAT in an independent culturally different sample and to contrast the factor structures of the Italian Q−CHAT with the Singaporean one. The Kaiser-Meyer-Olkin (KMO) statistic was 0.80, indicating that sampling was adequate for EFA. Use of the Kaiser criteria suggested a solution of three factors for our sample with Eigenvalues > 1, which explained 61% of the total variance (Factor 1, 26%; Factor 2, 22%; and Factor 3, 13%). This result was also confirmed by the scree plot (Fig. 2) that indicated one point of inflexion after the third factor.

The EFA was thus recomputed specifying three factors. As for Magiati et al., a factor loading of 0.3 was chosen as the threshold above which items would be retained in the factor structure. Therefore, items 2, 9, 13, 14 and 21 were excluded, as their factor loading did not reach the threshold value. The pattern matrix (based on oblique rotation) with the factor loadings of the 25 items is shown in Table 3.

The three Factors resulted from the EFA and the respective correlation coefficients are shown in Fig. 3.

3.4. Internal consistency

Internal consistency of the Q−CHAT was adequate overall (Cronbach's alpha = 0.68). When the factor structure of the questionnaire was examined, we found comparable internal consistency for the three factors (Cronbach's alpha factor 1 = 0.73, Cronbach's alpha factor 2 = 0.64 Cronbach's alpha factor 3 = 0.63).

3.5. Comparison of Italian Q-CHAT scores with UK and Singaporean scores

Mean Q−CHAT total scores in our sample and in the UK as well as Asiatic samples are shown in Table 4.

Table 3
Cross-cultural comparisons.

Study	Country	N	Age (mean)	Q-CHAT (mean, SD)	t (df)	p	Effect size (Cohen's *d*)
Toddlers	Italy	2400	23.2	25.8 (7.8)			
Allison 2008	UK	754	21.2	26.7 (7.8)	2.8 (3152)	0.005	0.1
Auyeung 2010	UK	129	19	26.5 (7.1)	1.2 (2527)	0.2	0.1
Wong 2014	UK (preterm)	141	24	33.7 (8.3)	12 (2539)	< .001	1.01
Magiati 2015	Singapore	368	18	35.6 (7.2)	24.8 (2766)	< .001	1.27
Magiati 2015	Singapore	396	24	33.2 (7.7)	18.3 (2794)	< .001	0.95

SD = standard deviation; t = *t*-test statistic; df = degrees of freedom; Cohen's *d:* Small = 0.20; Medium = 0.50; Large = 0.80; Very large = 1.20.

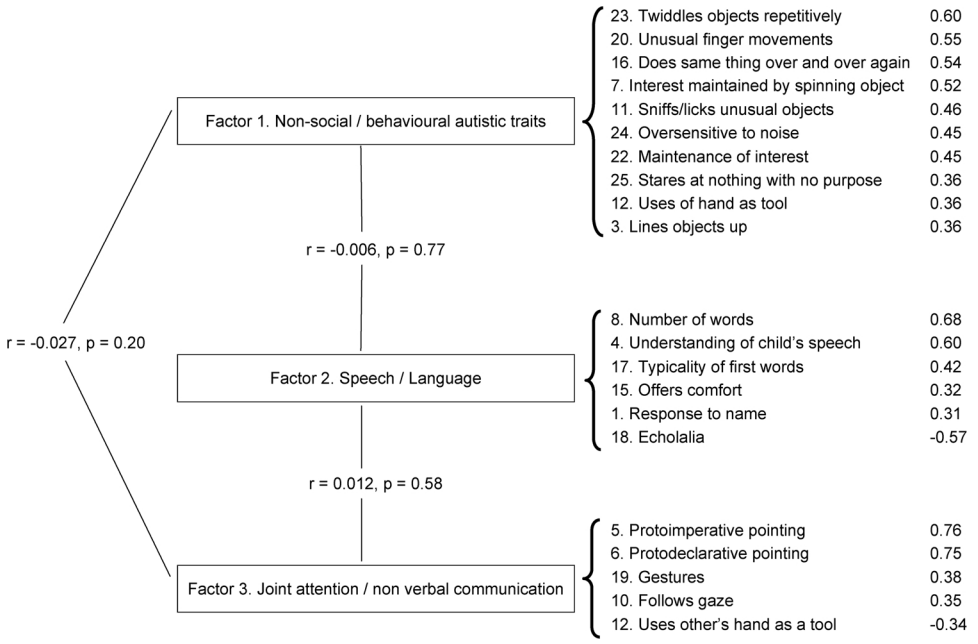

Fig. 3. Three-factor structure of the Q−CHAT.
Factor loadings beside individual items.

In light of the findings reported by Wong et al. (2014) we explored the effect of gestational week on the Q−CHAT. Although the Q−CHAT total score was slightly higher in preterm children (N = 222, mean = 26.4, SD = 8.1) than children born at term (N = 2023, mean = 25.7, SD = 7.8), these differences were not statistically significant (t(2243) = 1.39, p = 0.16, equal variances

Table 4
Pattern matrix (direct oblimin rotation) displaying factor loadings of EFA.

		Factors		
		1	2	3
23.	Twiddles objects repetitively[a]	**0,60**	0,03	002
20.	Unusual finger movements[a]	**0,55**	−0,02	0,07
16.	Does same thing over and over again[a]	**0,54**	0,14	−0,10
7.	Interest maintained by spinning objects[a]	**0,52**	−0,06	0,06
11.	Sniffs/licks unusual objects[a]	**0,46**	0,11	−0,16
24.	Oversensitive to noise[a]	**0,45**	−0,04	0,08
22.	Maintenance of interest[a]	**0,45**	−0,13	0,12
25.	Stares at nothing with no purpose[a]	**0,36**	0,05	0,13
12.	Uses of hand as tool[a]	**0,36**	0,20	−0,34
3.	Lines up objects[a]	**0,36**	−0,24	0,05
8.	Number of words[a]	0,11	**0,68**	−0,30
4.	Understand child's speech	0,04	**0,60**	−0,07
18.	Echolalia[a]	0,11	**−0,57**	0,03
17.	Typicality of first words	011	**0,42**	0,05
15.	Offers comfort	−0,14	**0,32**	0,05
1.	Looks when name called	0,06	**0,31**	0,26
5.	Protoimperative pointing	0,01	−0,21	**0,76**
6.	Protodeclarative pointing	0,02	−0,09	**0,75**
19.	Gestures	−0,02	0,16	**0,38**
10.	Follows a look	−0,06	0,17	**0,35**
13.	*Walks on tiptoes*[a]	*0,28*	*0,03*	*−0,03*
14.	*Adapts to changes in routine*	*0,16*	*0,08*	0,11
21.	*Checks reaction*	*−0,18*	*0,12*	*0,16*
2.	*Eye contact*	*0,14*	*0,28*	*0,21*
9.	*Pretend play*	*−0,04*	*0,25*	*0,27*

Extraction Method: Principal Axis Factoring. Rotation Method: Oblimin with Kaiser Normalization. Rotation converged in six iterations. Italicized entries indicate which items load on which factor.
[a] Reverse-scored items.
[b] Items 13, 14, 21, 2 and 9 were excluded due to loading < 0.3.

assumed, Cohen's d = 0.10). However, when we compared the distribution of children having a Q – CHAT total score above 2SD (> 42) in preterm and at term children, we found that the proportion of preterm children having a Q – CHAT total score above 2SD (5.9%) was significantly higher than those born at term (3%, Fisher's exact-test, p = 0.03).

4. Discussion

Our study reported the preliminary psychometric properties of the Q – CHAT as a quantitative measure for autistic traits to be implemented by paediatricians during routine developmental surveillance at 18–24 months. A large population of toddlers representative of the North, Centre and South of Italy was screened through the Q – CHAT.

When the main demographic and clinical characteristics of the sample were analyzed we found that lower maternal age and higher paternal age significantly predicted higher Q – CHAT scores. Previous studies have reported an association between father's age and risk of ASD (Hultman, Sandin, Levine, Lichtenstein, & Reichenberg, 2011; Kong et al., 2012; Reichenberg et al., 2006). Total Q – CHAT scores in our sample followed a normal distribution supporting the conceptualisation that autistic traits are distributed along a continuum in the population (Constantino, 2011; Bölte, Westerwald, Holtmann, Freitag, & Poustka, 2011; Robinson et al., 2011). Internal consistency, and item to total score correlations were acceptable and consistent with previous findings in both UK (Allison et al., 2008, 2012) and Singaporean studies (Magiati et al., 2015).

Q – CHAT mean scores in our population were comparable to the mean scores reported by Allison et al. (2008) in an age-matched sample, and by Auyeung et al. (2010); Auyeung et al., 2012 in younger children. No effect of age on the Q – CHAT scores was found in our study or in those mentioned above. Furthermore, consistent with Allison et al. (2008), we found a gender difference in the Q – CHAT scores, with boys scoring significantly higher (more autistic traits) than girls. As expected for large unselected populations, we replicated a small magnitude of difference (Cohen's d = 0.13 in Allison et al. (2008) and Cohen's d = 0.2 in our study, respectively). In our sample, gender significantly predicted the Q – CHAT scores, with maleness being positively associated with higher scores, in line with previous findings reporting sex differences in autism traits in unselected populations (Baron-Cohen, Wheelwright, Skinner, Martin, & Clubley, 2001; Constantino & Todd, 2003; Leekam et al., 2007; Messinger et al., 2015; Reinhardt, Wetherby, Schatschneider, & Lord, 2015; Ruzich et al., 2015; Williams et al., 2008) and with the Extreme Male Brain theory of autism (Baron-Cohen, 2002). These results confirm a similar distribution of autistic traits in the toddler population in Italy and in the UK.

In the light of the findings of higher Q – CHAT scores in premature children in the UK (Wong et al., 2014), we explored the effect of gestational week and birth weight on the Q – CHAT. Different to their results, we did not find a statistically significant difference between preterm born children and those born at term (although preterm children in our sample exhibited slightly higher Q – CHAT scores than children born at term). Nevertheless, we found that a significantly larger proportion of premature children in our sample displayed very high Q – CHAT scores (above 2 SD) as compared with children born at term, suggesting that social-communication developmental trajectories in preterm children should be monitored.

When compared with Singaporean toddlers (Magiati et al., 2015), the Italian mean Q – CHAT scores were significantly lower, with a large effect size. This finding is in line with a number of other studies reporting higher autistic traits in the Asiatic population using quantitative assessments such as the CAST (Sun et al., 2014), the SRS (Gau, Liu, Wu, Chiu, & Tsai, 2013) and the AQ (Freeth, Sheppard, Ramachandran, & Milne, 2013). It is arguable that these differences are partly due to different parental interpretations and expectations regarding their child's behaviours and to different cultural belief systems (Chuang & Su, 2009; Freeth, Milne, Sheppard, & Ramachandran, 2014. However, other studies, using the SRS, did not confirm cross-cultural differences in autistic traits (Takei et al., 2014; Wang, Lee, Chen, & Hsu, 2012) and other individual children's characteristics such as age, language and cognitive level or other behavioural problems were found to influence the interpretation of parents' report measures (Hus, Bishop, Gotham, Huerta, & Lord, 2013). Despite the differences in the mean scores, the Italian Q – CHAT demonstrated a remarkable consistency in its three-factor structure with the Asian Q – CHAT, with most of the items falling on to the same factors as reported by Magiati et al. (2015). In our study, Factor 1 "Non-social/behavioural autistic traits" has a close correspondence with Factor 2 of Magiati et al. (2015) with eight out of ten items in common ("Twiddles objects repetitively", "Unusual finger movements", "Does same thing over and over again", "Interest maintained by spinning object", "Sniffs/licks unusual objects", "Oversensitive to noise", "Maintenance of interest", "Stares at nothing with no purpose"). The other two items, which load on to this Factor in our study, are "Lines up objects", still directly related to repetitive behaviour and "Use of other's hand as a tool" that may have been interpreted by parents as an unconventional behavioural strategy and/or an atypical self-management behaviour. Factor 2 in our study is related to "Speech and Language" as for Factor 3 in Magiati et al. (2015) and four out of six items were overlapping ("Number of words", "Echolalia", "Typicality of first words", "Understanding of child's speech"). Factor 3 ("Joint attention / non verbal communication") consists of shared attention behaviours such as gaze following, use of conventional gestures (pointing and other gestures) for reference sharing and non-verbal communication as well as, with reversed score, use of unconventional gestures such as placing another's hand on objects. All the items in this Factor overlap with Factor 1 "Social-communicative autistic traits" in Magiati et al. (2015) and indeed directly relate to young children's social skills development.

The factor structure of the Q – CHAT indicated that two of the three factors (Factor 1 and 3 respectively) are directly related to the core dyadic structure of ASD impairments proposed in DSM-V (American Psychiatry Association, 2013) while Factor 2 appears to be more generally related to language development.

Furthermore, the low inter-correlation of the three factors replicated in our study the findings reported in other studies (Constantino et al., 2004; Mandy & Skuse, 2008; Sun et al., 2014) that the social-communication domain and the non-social domain of stereotyped and repetitive behaviors have independent contribution to the core symptomatology of ASD.

4.1. Limitations and future directions

Although this represents the first study in a large paediatric unselected population across different regions of the North, Centre and South of Italy, the paediatricians involved were relatively few (43% of those contacted). Also, the majority of the children were enrolled in one of the three regions and this sampling bias may have implication on generalization of the results. Furthermore, the parental response rate was 40.7%, less than we expected but still higher compared to other population based screenings (Allison et al., 2008; Reznick, Baranek, Reavis, Watson, & Crais, 2007) and a test-retest analysis was not conducted. Lastly, the socio-demographic characteristics of our sample were different from those reported in the Italian 2015 census in terms of parents' education, with secondary level of education under-represented and higher education over-represented. All these aspects have to be taken into account when considering the generalization of our findings to the general population.

4.2. Implications

Our study confirms the validity of the Q – CHAT as a dimensional screening instrument for autism traits in the general population very early on.

The main normative data and factor structure replicated in our sample support the cross-cultural stability of the Q – CHAT and the independent contribution of the social communication domain and repetitive stereotyped behaviours to the core autistic symptoms.

Furthermore, the study demonstrated that the QCHAT is a feasible and easy to administer instrument to be used within the paediatric setting during well-child assessments. The study has important implications, demonstrating that the QCHAT is a reliable tool to support clinicians' decision-making about further steps for referral and early diagnostic assessment fostering early clinical diagnosis and intervention with a substantial impact on the prognosis and parent-child relationship.

Acknowledgements

This research ("Toddlers Project") has been supported by the Ministry of Health and Tuscany Region (GR-2010-2319668). The funding body had no involvement in the study. We warrant that all the authors have no conflict of interests to declare.

We would like to acknowledge all the paediatricians, parents/caregivers and children involved in the study. We are grateful to Bhismadev Chakrabarti for their valuable feedback and to Professor Giovanni Cioni and Maria Luisa Scattoni for their support and contribution. CA and SBC were supported by the Autism Research Trust during the period of this work. The research was supported by the National Institute for Health Research (NIHR) Collaboration for Leadership in Applied Health Research and Care East of England at Cambridgeshire and Peterborough NHS Foundation Trust. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.

References

Allison, C., Auyeung, B., & Baron-Cohen, S. (2012). Toward brief "Red Flags" for autism screening: the short autism spectrum quotient and the short quantitative checklist for autism in toddlers in 1,000 cases and 3,000 controls [corrected]. *Journal of the American Academy of Child and Adolescent Psychiatry, 51*, 202–212.

Allison, C., Baron-Cohen, S., Wheelwright, S., Charman, T., Richler, J., Pasco, G., et al. (2008). The Q-CHAT (Quantitative CHecklist for Autism in Toddlers): a normally distributed quantitative measure of autistic traits at 18–24 months of age: preliminary report. *Journal of Autism and Developmental Disorders, 38*, 1414–1425.

American Psychiatry Association (2013). *Diagnostic and statistical manual of mental disorders* (5th edn.). Arlington: American Psychiatric Publishing.

Auyeung, B., Ahluwalia, J., Thomson, L., Taylor, K., Hackett, G., O'Donnell, K. J., et al. (2012). Prenatal versus postnatal sex steroid hormone effects on autistic traits in children at 18 to 24 months of age. *Molecular Autism, 3*, 17.

Auyeung, B., Taylor, K., Hackett, G., & Baron-Cohen, S. (2010). Foetal testosterone and autistic traits in 18 to 24-month-old children. *Molecular Autism, 1*, 11.

Baird, G., Charman, T., Baron-Cohen, S., Cox, A., Swettenham, J., Wheelwright, S., et al. (2000). A screening instrument for autism at 18 months of age: A 6-year follow-up study. *Journal of the American Academy of Child and Adolescent Psychiatry, 39*, 694–702.

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

Baron-Cohen, S., Allen, J., & Gillberg, C. (1992). Can autism be detected at 18 months? The needle, the haystack, and the CHAT. *The British Journal of Psychiatry, 161*, 839–843.

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.

Bölte, 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*, 66–72.

Charman, T., & Gotham, K. (2013). Measurement Issues: Screening and diagnostic instruments for autism spectrum disorders - lessons from research and practice. *Child and Adolescent Mental Health, 18*, 52–63.

Chuang, S. S., & Su, Y. (2009). Do we see eye to eye? Chinese mothers' and fathers' parenting beliefs and values for toddlers in Canada and China. *Journal of Family Psychology, 23*, 331–341.

Constantino, J. N. (2011). The quantitative nature of autistic social impairment. *Pediatric Research, 69*, 55R–62R.

Constantino, J. N., & Charman, T. (2016). Diagnosis of autism spectrum disorder: Reconciling the syndrome, its diverse origins, and variation in expression. *The Lancet Neurology, 15*, 279–291.

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

Constantino, J. N., 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*, 719–726.

Elsabbagh, M., & Johnson, M. H. (2016). Autism and the social brain: The first-year puzzle. *Biological Psychiatry, 80*, 94–99.

Elsabbagh, M., & Johnson, M. H. (2010). Getting answers from babies about autism. *Trends in Cognitive Sciences, 14*, 81–87.

Freeth, M., Milne, E., Sheppard, E., & Ramachandran, R. (2014). Autism across cultures: Perspectives from non-Western cultures and implications for research. In F. R. Volkmar, S. J. Rogers, P. Rhea, & K. A. Pelphrey (Eds.). *Handbook of autism and pervasive developmental disorders*(4th edn). Wiley.

Freeth, M., Sheppard, E., Ramachandran, R., & Milne, E. (2013). A cross-cultural comparison of autistic traits in the UK, India and Malaysia. *Journal of Autism and*

Developmental Disorders, 43, 2569–2583.
García-Primo, P., Hellendoorn, A., Charman, T., Roeyers, H., Dereu, M., Roge, B., et al. (2014). Screening for autism spectrum disorders: State of the art in Europe. *European Child & Adolescent Psychiatry, 23*, 1005–1021.
Gau, S., Liu, L., Wu, Y. Y., Chiu, Y. N., & Tsai, W. C. (2013). Psychometric properties of the Chinese version of the social responsiveness scale. *Research in Autism Spectrum Disorders, 7*, 349–360.
Hultman, C. M., Sandin, S., Levine, S. Z., Lichtenstein, P., & Reichenberg, A. (2011). Advancing paternal age and risk of autism: New evidence from a population-based study and a meta-analysis of epidemiological studies. *Molecular Psychiatry, 16*, 1203–1212.
Hus, V., Bishop, S., Gotham, K., Huerta, M., & Lord, C. (2013). Factors influencing scores on the social responsiveness scale. *Journal of Child Psychology and Psychiatry, 54*, 216–224.
Johnson, M. H., Jones, E. J., & Gliga, T. (2015). Brain adaptation and alternative developmental trajectories. *Development and Psychopathology, 27*, 425–442.
Kong, A., Frigge, M. L., Masson, G., Besenbacher, S., Sulem, P., Magnusson, G., et al. (2012). Rate of de novo mutations and the importance of father's age to disease risk. *Nature, 23*, 471–475.
Leekam, S., Tandos, J., McConachie, H., Meins, E., Parkinson, K., Wright, C., et al. (2007). Repetitive behaviours in typically developing 2-year- olds. *Journal of Child Psychology and Psychiatry, 48*, 1131–1138.
Magiati, I., Goh, D. A., Lim, S. J., Gan, D. Z., Leong, J. C., Allison, C., et al. (2015). The psychometric properties of the Quantitative-Checklist for Autism in Toddlers (Q-CHAT) as a measure of autistic traits in a community sample of Singaporean infants and toddlers. *Molecular Autism, 6*, 40.
Mandy, W. P., & Skuse, D. H. (2008). What is the association between the social- communication element of autism and repetitive interests, behaviours and activities? *Journal of Child Psychology and Psychiatry, 49*, 795–808.
Mandy, W. P. L., Charman, T., & Skuse, D. H. (2012). Testing the construct validity of proposed criteria for DSM-5 Autism Spectrum disorder. *Journal of the American Academy of Child and Adolescent Psychiatry, 51*, 41–50.
Messinger, D. S., Young, G. S., Webb, S. J., Ozonoff, S., Bryson, S. E., Carter, A., et al. (2015). Early sex differences are not autism-specific: A Baby Siblings Research Consortium (BSRC) study. *Molecular Autism, 6*, 32.
Reichenberg, A., Gross, R., Weiser, M., Bresnahan, M., Silverman, J., Harlap, S., et al. (2006). Advancing paternal age and autism. *Archive General Psychiatry, 63*, 1026–1032.
Reinhardt, V. P., Wetherby, A. M., Schatschneider, C., & Lord, C. (2015). Examination of sex differences in a large sample of young children with autism spectrum disorder and typical development. *Journal of Autism and Developmental Disorders, 45*, 697–706.
Reznick, J. S., Baranek, G. T., Reavis, S., Watson, L. R., & Crais, E. R. (2007). A parent-report instrument for identifying one-year- olds at risk for an eventual diagnosis of autism: The first year inventory. *Journal of Autism and Developmental Disorders, 37*, 1691–1710.
Robins, D. L., Casagrande, K., Barton, M., Chen, C. M. A., Dumont-Mathieu, T., & Fein, D. (2014). Validation of the modified checklist for autism in toddlers, revised with follow-up (M- CHAT-R/F). *Pediatrics, 133*, 37–45.
Robins, D. L., Fein, D., Barton, M. L., & Green, J. A. (2001). The Modified Checklist for Autism in Toddlers: An initial study investigating the early detection of autism and pervasive developmental disorders. *Journal of Autism and Developmental Disorders, 31*, 131–144.
Robinson, E., Munir, K., Munafò, M. R., Hughes, M., McCormick, M., & Koenen, K. C. (2011). The stability of autistic traits in the general population: Further evidence for a continuum of impairment. *Journal of the American Academy of Child and Adolescent Psychiatry, 50*, 376–384.
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.
Sun, X., Allison, C., Auyeung, B., Matthews, F. E., Zhang, Z., Baron-Cohen, S., et al. (2014). Comparison between a Mandarin Chinese version of the Childhood Autism Spectrum Test and the Clancy Autism Behaviour Scale in mainland China. *Research in Developmental Disabilities, 35*, 1599–1608.
Takei, R., Matsuo, J., Takahashi, H., Uchiyama, T., Kunugi, H., & Kamio, Y. (2014). Verification of the utility of the social responsiveness scale for adults in non-clinical and clinical adult populations in Japan. *BMC Psychiatry, 14*, 302.
Towle, P. O., & Patrick, P. A. (2016). Autism Spectrum disorder screening instruments for very young children: A systematic review. *Autism Research and Treatment, 2016* 4624829.
Wang, J., Lee, L. C., Chen, Y. S., & Hsu, J. W. (2012). Assessing autistic traits in a Taiwan preschool population: Cross-cultural validation of the Social Responsiveness Scale (SRS). *Journal of Autism and Developmental Disorders, 42*, 2450–2459.
Williams, J. G., Allison, C., Scott, F. J., Bolton, P. F., Baron-Cohen, S., Matthews, F. E., et al. (2008). The childhood autism Spectrum test (CAST): Sex differences. *Journal of Autism and Developmental Disorders, 38*, 1731.
Wing, L., Gould, J., & Gillberg, C. (2011). Autism spectrum disorders in the DSM-5: better or worse than the DSM-IV? *Research in Developmental Disabilities, 32*, 768–773.
Wong, H. S., Huertas-Ceballos, A., Cowan, F. M., & Modi, N. (2014). Evaluation of early childhood social-communication difficulties in children born preterm using the Quantitative Checklist for Autism in Toddlers. *Jornal de Pediatria, 164*, 26–33.