DOI:
10.1016/j.rasd.2022.101984

Document Version
Peer reviewed version

Research in Autism Spectrum Disorders (in press)
https://doi.org/10.1016/j.rasd.2022.101984

Cross-cultural differences in reporting autistic symptoms in toddlers: A study synthesizing M-CHAT(-R) data from ten countries

Dejan Stevanovic[a], Diana L. Robins[b], Floriana Costanzo[c], Elisa Fucà[c], Giovanni Valeri[c], Stefano Vicari[c], Handan Ozek Erkuran[d], Ferhat Yaylaci[e], Lilia Albores-Gallo[f], Gabriel Gatica-Bahamonde[g], Maia Gabunia[h], Medea Zirakashvili[i], Tony Charman[j], Sayyed Ali Samadi[k], Teck-Hock TOH[l], Windham Gayle[m], Laura Brennan[n], Tatjana Zorcec[o], Alejandra Auza[p], Maretha de Jonge[q], Noordeen Shoqirat[r], Rajna Knez[s]

[a]Clinic for Neurology and Psychiatry for Children and Youth, Belgrade, Serbia, email: stevanovic.dejan79@gmail.com

[b]A.J. Drexel Autism Institute, Drexel University, Philadelphia, Pennsylvania, United States of America, email: DRobins@drexel.edu

[c]Child and Adolescent Neuropsychiatry Unit, Bambino Gesù Children's Hospital, Rome, Italy; email: floriana.costanzo@opbg.net; elisa.fuca.sci@gmail.com; giovanni.valeri@opbg.net ; stefano.vicari@opbg.net

[d]Dr Behcet Uz Children's Research and Training Hospital Child and Adolescent Psychiatry Unit, Izmir, Turkey, email: handanozek@yahoo.com

[e]Bursa Dortcelik Children's Hospital, Child psychiatry unit, Bursa, Turkey, email: drferhatyaylaci@hotmail.com

[f]Research Division, Hospital Psiquiátrico Infantil "Dr. Juan N. Navarro," Secretaría de Salud, Mexico City, Mexico, email: algali.1209@gmail.com

[g]Hospital de Villarrica, Department of Mental Health, Universidad de La Frontera, Temuco, Chile, email: gabriel.gatica@ufrontera.cl, g.gatica01@ufromail.cl

[h]Mental Health Center, Tbilisi, Georgia, email: maiagabunia.mhc@gmail.com

[i]Ilia State University, Tbilisi, Georgia, Email: m_zirakashvili@yahoo.com

[j]King's College London, Institute of Psychiatry, Psychology & Neuroscience, London, UK, Email: tony.charman@kcl.ac.uk

[k]Institute of Nursing and Health Research, University of Ulster, Newtownabbey BT37 0QB, Coleraine, Northern Ireland, United Kingdome, Email: s.samadi@ulster.ac.uk

[l]Clinical Research Centre & Department of Pediatrics, Sibu Hospital, Ministry of Health Malaysia, Sarawak, Malaysia, Email: thtoh@moh.gov.my

[m]Environmental Health Investigations Branch, California Department of Public Health, Richmond, CA, USA, Email: Gayle.Windham@cdph.ca.gov

[n]University of Connecticut, United States of America, Email: laura.a.brennan@gmail.com

[o]Department for Psychophysiology, University Children's Hospital, Skopje, Macedonia, Email: tzorcec@gmail.com

[p]Language and Cognition Laboratory, Hospital General Dr. Manuel Gea González, Mexico City, Mexico, email: alejandra.auza@yahoo.com

[q]Faculty of Social Science, Education and Child Studies, Clinical neuroscience and developmental disorders, Leiden University, Leiden, The Netherlands, email: m.v.de.jonge@fsw.leidenuniv.nl

[r]Faculty of Nursing, Mu'tah University, Mu'tah, Jordan, email: noordeen@mutah.edu.jo

[s]Department of Women´s and Children´s health, Skaraborgs Hospital, Skövde, Sweden; University of Gothenburg, Sahlgrenska Academy, Institute of Neuroscience and Physiology, Gothenburg, Sweden, email: rajna.knez@gu.se

Corresponding author: Stevanovic Dejan Clinic for Neurology and Psychiatry for Children and Youth, Dr Subotica 6a, 11000 Belgrade, Serbia; Tel: +381692412006, Email: stevanovic.dejan79@gmail.com

Abstract

Background: This study aimed to evaluate the endorsement rates of the M-CHAT(-R) items by parents/caregivers of toddlers with autism spectrum disorder (ASD) from ten countries: Albania, Chile, Georgia, Macedonia, Malaysia, Mexico, Serbia, Turkey, United Kingdom (UK), and the United States of America (USA).

Method: Data were aggregated for toddlers aged 14 – 36 months who participated in previous studies or completed clinical screening. An item with < 30% of endorsements was considered to have a low endorsement rate, an item falling within the range of 30-60% to have a moderate endorsement, and an item with > 60% to have a high endorsement.

Results: All items had a low endorsement rate in at least one country and moderate to high in others. Of 20 items, 14 had a moderate to high endorsement rate in seven to nine countries. Of particular relevance in this regard are items with moderate to high endorsement rates in all countries excluding Malaysia, such as points to get help, points to show, brings things to show, follows a point, follows your gaze, and understands what is said. On the other hand, items makes eye contact, responds to name, hearing concerns, and reciprocal smile were consistently differently interpreted across the countries.

Conclusions: This study showed differences in responding to M-CHAT(-R) items by parents/caregivers across ten countries, which may indicate cross-country variations in the recognition and evaluation of early autistic symptoms in toddlers. Items related to joint attention, imitation, social engagement, and language comprehension may be less variable and potentially interpreted as universal atypical behaviors in toddlers with ASD.

Key words: toddlers; autism; screening; culture; equivalence.

Introduction

Autism spectrum disorder (ASD) is present across all cultures worldwide (American Psychiatric Association – APA, 2013; World Health Organization – WHO, 2018). Evidence has emerged over the past decade that expression and/or reporting of ASD symptoms with available instruments may vary across regional and linguistically diverse samples (e.g., de Leeuw, Happé, & Hoekstra, 2020; Harrison, Long, Tommet, & Jones, 2017; Norbury & Sparks, 2013). For example, Matson et al. (2011) showed potential differences in expression and reporting symptoms of ASD using the Autism Spectrum Disorders-Diagnostic for Children (ASD-DC; Matson & Gonzalez, 2007) among parents of children aged 2 – 16 years from Israel, South Korea, the United Kingdom (UK), and the United States of America (USA). This study found significant differences in reporting core ASD symptom domains (i.e., nonverbal communication and socialization, verbal communication, and insistence upon sameness/restricted interests), with 10 out of 40 items having the highest similar level of endorsement across all countries (Matson et al., 2011). In another study, Matson et al. (2017) used the Baby infant screen for children with autism traits (BISCUIT; Matson, Boisjoli, & Wilkins, 2007) with the parents of toddlers aged 17 to 32 months diagnosed with ASD from Greece, Italy, Japan, Poland, and the USA. This study showed significant differences in overall ASD symptom severity, and practically all 53 tested items related to the symptoms were endorsed differently across the countries. However, the endorsement patterns were more consistent across those items related to socialization and communication, whereas more variation was observed among those linked to restricted and repetitive behaviors. Carruthers et al. (2018) used the Autism Spectrum Quotient (AQ; Auyeung, Baron-Cohen, Wheelwright, & Allison, 2008) with parents of children with ASD and typically developing children aged 4 – 9 years across samples from India, Japan and the UK and showed that 28 out of 50 items have acceptable discrimination properties across all three countries, but only five items measuring social situations and communication to have almost identical discrimination power across the three countries.

Similar findings were observed in studies using diagnostic instruments. For example, Harrison et al. (2017) used the Autism Diagnostic Observation Schedule (ADOS; Lord et al., 2000) with the parents/caregivers of children with ASD aged 4 – 18 years in the USA, but with African American, Hispanic, Asian, or White American ethnic background. They found item-level bias

according to race and/or ethnicity for three out of ten evaluated items, namely unusual eye contact, stereotyped/idiosyncratic use of words or phrases, and immediate echolalia, all of which measures restricted and repetitive behaviors. In our recent study with children with ASD aged 2 – 18 years from India, Jamaica, Mexico, Spain, Turkey, and the USA (Stevanovic et al., 2021), items from the Childhood Autism Rating Scale (CARS; Schopler, Reichler, DeVellis, & Daly, 1980) related to social communication and interaction (i.e., relating to people, imitation, emotional response, and verbal and nonverbal communication) displayed lower levels of cross-country non-invariance compared to items related to stereotyped behaviors and sensory sensitivity (i.e., body and object use, adaptation to change, or taste, smell, and touch response). These results could indicate that items within the available instruments measuring ASD behaviors related to socialization, interaction, and communication, including language, are more universally present, less likely variable, but less likely culturally sensitive, whereas restricted, repetitive behaviors or sensory processing are less universally present and probably more variable among children with ASD when assessed across different cultures (Stevanovic et al., 2021).

The extent to which differences in reporting ASD symptoms are inherent characteristics of the disorder (i.e., the cross-cultural variability), reflected in the variability in responding across cultures, or whether they are the effects of the measurements with these instruments across different cultures (i.e., biases in the cross-cultural validity) is still unclear. Studies testing aspects of cross-cultural validity of ASD instruments for children, especially toddlers, have appeared recently and were not replicated across different societies. In addition, almost all available instruments were developed in Western societies (e.g., Falkmer, Anderson, Falkmer, & Horlin, 2013; Randall et al., 2018), and although many of these tools are being adapted in different languages, the translation and cultural adaptation process itself is not always clearly outlined and often fails to follow the recommended guidelines (DuBay & Watson, 2019). Of particular interest is the extent to which instruments operate the same way and underlying constructs have the same theoretical structure across two or more regional/cultural groups (e.g., Dimitrov, 2010). Another relevant aspect to the measurements of ASD is that there may be differences between behaviors parents endorse and observations of trained clinicians, which may also be related to specific cultural values and practices. Finally, most instruments have not been sufficiently validated across cultures, particularly regarding tasks and materials used for the diagnosis, which may not be culturally relevant (e.g., Abubakar, Sewanyana, & Newton, 2016; Durkin et al., 2015). If clinicians

use materials or measures that lack cultural relevance, the assessment practices employed may misidentify or under-identify ASD (Harris, Barton, & Albert, 2014). Similar to the challenges described for evaluation measures, screening tools to identify children at risk for ASD should be cross-validated in different cultures. Therefore, it is essential to evaluate the equivalence of screening tools across cultures to identify those children needing expert evaluation. In addition, screening tools are often applied both in low- and high-risk settings, such as primary care well-visits and tertiary care centers, respectively, and cross-cultural validation should consider the context in which the screening is administered.

The Modified Checklist for Autism in Toddlers (M-CHAT; Robins, Fein, & Barton, 1999), and its revised form (M-CHAT-R; Robins et al., 2009), is the most frequently used ASD screener in toddlers and it seems to be one of the most promising instruments from a psychometric point of view for ASD (Robins et al., 2014), but also for screening for developmental disorders among toddlers including low-resource regions (Dai et al., 2021; Marlow, Servili, & Tomlinson, 2019; Petrocchi, Levante, & Lecciso, 2020). The M-CHAT-R is also recommended for use in low- and middle-income regions (Marlow et al., 2019) due to a simplified scoring procedure and the two-stage process that pairs the initial items with flow charts for structured Follow-Up (M-CHAT-R/F) questions that reduce the likelihood of false-positive responses. The M-CHAT(-R) is available in several different languages, and there are many psychometric and epidemiological studies with it (for details, see https://mchatscreen.com/); thus, the M-CHAT-R/F is positioned as an ASD screener to be used globally.

In available studies from different world regions, the M-CHAT(-R) showed variable measurement and screening properties across various ethnic, racial, or cultural groups (e.g., Albores-Gallo et al., 2012; Mohamed et al., 2016; Windham et al., 2014). Studies show that although specificity for ASD screening is generally high, sensitivity varies markedly across low-risk community samples in world populations. For example, low to moderate sensitivity was found in Malaysia (17.9 - 63.6%; Toh, Tan, Lau, & Kiyu, 2018) and in France (67%; Baduel et al., 2017), whereas sensitivity was high in Iran for the M-CHAT (90.3%; e.g., Samadi & McConkey, 2015), while for the M-CHAT-R in China (> 85%; Guo et al., 2019) and Turkey (Oner & Munir, 2019). In addition, studies from different regions identified different sets of critical items for ASD identification, with only a few overlapping items across cultures. Four items – *brings things to*

show, *imitation of action*, *follows a point*, and *understands what is said* – were common in at least three of four studies (Albores-Gallo et al., 2012; Darquah, 2021; Kamio et al., 2015; Robins et al., 2001), whereas 14 other items were implicated in one or two studies, suggesting differential performance. In contrast, a recent study indicated that the M-CHAT-R/F is an effective screener with similar screening properties in non-Hispanic Black and non- Hispanic White children in the United States (USA; Dai et al., 2021).

The Autism Spectrum Disorder International Consortium (ASDIC; Stevanovic, 2018) organized this study with the aim to contribute towards a greater understanding of early recognition of autistic symptoms in toddlers across different countries. The study tested for cultural differences in reporting autistic symptoms with the M-CHAT(-R) among toddlers with ASD aggregated from studies and clinical data in ten countries: Albania, Chile, Georgia, Macedonia, Malaysia, Mexico, Serbia, Turkey, the United Kingdom (UK), and the USA. Specifically, the aim was to evaluate the endorsement rates of the M-CHAT(-R) items by parents/caregivers of toddlers with ASD across the countries to identify items that best predict an ASD diagnosis (i.e., universal key indicators) and items with a potential cultural difference (i.e., culturally sensitive) in reporting ASD symptoms.

Methods

Study sample

Data for this study were aggregated for toddlers who participated in previous studies or from unpublished/clinical data provided by the respective authors. The data were selected in the following steps. First, several searches were done for studies that published any M-CHAT(-R) data in PubMed and Scopus. The authors of the published studies were contacted and invited to share their data. Second, researchers and clinicians already included in the ASDIC, our databases from previous cross-cultural projects (https://www.icmhsg.org/), and from open-calls were separately invited to share data. The main inclusion criteria for the data were: (1) age range at screening 14 – 36 months; (2) a diagnosis of autism, autism spectrum disorder, or pervasive developmental disorder based on the diagnostic evaluation and relevant criteria (APA, 2000, 2013; WHO, 1992), and (3) data available for individual M-CHAT(-R) items completed by parents/caregivers. The age range of 14-36 months was selected considering that the M-CHAT(-R) has been most often tested in samples between 14 and 36 months (e.g., Oner & Munir, 2019; Wieckowski et al., 2021; Yama,

Freeman, Graves, Yuan, & Karen Campbell, 2012; Yuen, Penner, Carter, Szatmari, & Ungar, 2018). Children who were screened with M-CHAT(-R) but did not complete an evaluation, and those who were evaluated and not diagnosed with ASD were excluded.

Overall, M-CHAT(-R) data were available for toddlers with ASD from the following countries (see Table 1): Albania (n = 17 from Brennan, Fein, Como, Rathwell, & Chen, 2016), Chile (n = 15; unpublished/clinical data from Gatica-Bahamonde and from Gatica-Bahamonde et al., 2021), Georgia (n = 189; unpublished/clinical data from Medea Zirakashvili), Macedonia (n = 95; unpublished/clinical data from Tatjana Zorcec), Malaysia (n = 52 from Toh, Tan, Lau, & Kiyu, 2018), Mexico (n = 46 from Albores-Gallo et al., 2012), Serbia (n = 30 from Stevanović, 2021), Turkey (n = 150; unpublished/ clinical data from Handan Ozek Erkuran), UK (n = 28 from Charman et al., 2016), and USA (n = 312 from Chlebowski, Robins, Barton, & Fein, 2013; Robins et al., 2014; Windham et al., 2014). Based on the context in which the M-CHAT(-R) was completed, the samples from Albania, Chile, Malaysia, and USA were considered low-risk, from Mexico and Serbia mixed low-/high risk, and from Georgia, Macedonia, Turkey, and UK high-risk.

Expert clinicians made the final ASD diagnosis based on DSM-IV TR criteria (for Macedonia, Malaysia, Mexico, and the USA), DSM-5 criteria (for Chile, Macedonia, Serbia, Turkey, and the USA), ICD-10 criteria (for Georgia, Serbia, and the UK), or only based on the Autism Diagnostic Observation Schedule (ADOS; Lord et al., 2000) for Albania. Most studies or data provided were missing information related to socio-economic status, race/ethnic background, and/or neuropsychiatric/genetic/medical comorbidities.

The study was approved by the Ethics committee of the Clinic for Neurology and Psychiatry for children and Youths Belgrade, Serbia.

Table 1

Main characteristics of the samples included

Country	N	Boys, n (%)	Age in months, mean (SD); range	Sample	M-CHAT(-R) language
Albania	17	12 (70.6)	26.00 (5.59), 18-36	*Low-risk*	Albanian

					Toddlers positively screened in a community, later diagnosed at a specialized center	
Chile	15	10 (66.7)	21.20	(2.62), 18-24	*Low-risk* Toddlers positively screened during regular visits at pediatricians, later diagnosed by a specialized team	Spanish
Malaysia	52	46 (88.5)	22.60	(5.26), 16-35	*Low-risk* Toddlers positively screened in a community later diagnosed at child health clinics	Malay, Mandarin, English
USA	312	238 (76.3)	21.63	(3.95), 14-36	*Low-risk* Toddlers positively screened in communities or regular pediatric visits, later diagnosed at specialized clinics	English, Spanish
Mexico	46	37 (80.4)	29.07	(5.20), 17-36	*Mixed low-/high-risk* Toddlers positively screened in a community and toddlers suspected for NDD from a psychiatric department, diagnosed at the department by a specialized team	Spanish
Serbia	30	24 (80.0)	28.40	(5.78), 15-36	*Mixed low-/high-risk* Toddlers positively screened in a community and toddlers suspected for NDD from a psychiatric department, diagnosed at the department by specialists	Serbian
Georgia	189	153 (81.0)	30.53	(5.48), 16-36	*High-risk* Toddlers suspected of NDD referred from neurologists or pediatricians, clinically screened, and diagnosed at a mental health center	Georgian
Macedonia	95	75 (78.9)	27.54	(5.54), 17-36	*High-risk* Toddlers suspected for NDD referred from pediatricians, clinically screened, and diagnosed by a specialized team	Macedonian
Turkey	150	118 (78.7)	26.79	(3.63), 17-30	*High-risk* Toddlers suspected of NDD referred from primary care physicians or pediatricians, clinically screened, and diagnosed at a psychiatric department	Turkish
UK	28	26 (92.9)	30.71	(3.58), 23-36	*High-risk* Toddlers referred to child health services and speech and language therapy services mainly suspected for	English

	NDD, clinically screened, and diagnosed by specialized teams

Note: Neurodevelopmental disabilities – NDD

M-CHAT(-R) data

In the study, the term "M-CHAT(-R)" is used to denote both the M-CHAT (Robins, Fein, & Barton, 1999) and M-CHAT Revised (M-CHAT-R; Robins, Fein, & Barton, 2009), if not otherwise stated. The M-CHAT (Robins et al., 2001) has 23 items, and the M-CHAT-R (Robins et al., 2014), dropped three items; these screening instruments ask parents to report about typical behaviors affected in toddlers with ASD as well as atypical behaviors common in ASD. All items are scored as at risk (fail) or not at risk (pass); for 17 items, "yes" is scored as not at risk; three items (oversensitivity to sound, hearing concerns, and finger movements) are reverse scored such that "no" is scored as not at risk. A positive M-CHAT(-R), defined as failing any three items, indicates risk for developing/having ASD, thus requiring further assessment. Beginning in the early 2000s, the structured Follow-Up was available for both original and the revised form. The Follow-Up is recommended for children who score in the moderate risk range (initial score 3-7) to acquire additional information about potential risk behaviors. Since many of the included studies did not incorporate Follow-Up, the current study reports only initial screening results.

Across the studies/databases from which the respective data were used for the present study, parents/primary caregivers self-completed the M-CHAT in Georgia (in Georgian), Malaysia (in Malay for 60, Mandarin for 21, and English for 13 toddlers), Mexico (in Spanish), Turkey (in Turkish), and the UK (in English), the M-CHAT-R in Albania (in Albanian), Chile (in Spanish), Macedonia (in Macedonian) and Serbia (in Serbian), and both versions in the USA (in English for 737 and in Spanish for 7 toddlers). Cronbach's alpha for the internal consistency of the M-CHAT(-R) with 20 items from the respective studies was .74 for Albania (Brennan et al., 2016), .75 for Chile, .92 for Georgia, .64 for Macedonia. .85 for Malaysia, .78 for Mexico, .87 for Serbia, .89 for Turkey, .81 for the UK, and .78 for the USA.

Data analysis

Considering that the M-CHAT and M-CHAT-R are similar versions that differs in only three items and both had been combined previously (e.g., Khowaja, Hazzard, Robins, 2015), data analyses

included 20 M-CHAT(-R) items. For all 20 items, the percentage of those who failed/scored at risk (i.e., item endorsement) was calculated separately for each country. An item with < 30% endorsements for toddlers with ASD was considered to have a low endorsement rate, an item falling within the range of 30-60% to have a moderate endorsement, and an item with > 60% to have a high endorsement. Adapting the procedure used for other ASD instruments (Carruthers et al., 2018; Gillis, Callahan, & Romanczyk, 2011), any item with a low endorsement rate in at least one country and a moderate to high in the other(s) was considered to represent a potential cultural difference. In contrast, any item with a moderate to high endorsement rate across seven or more countries was considered as a candidate for a universal key indicator. An additional analysis was done with the low-risk vs. high-risk group with items having potential cultural differences (i.e., a low endorsement rate in at least one country and a moderate to high in the other(s) in the group) or being universal key indicators (i.e., a moderate to high endorsement rate across all countries in the group).

Results

The percentages of toddlers with ASD who were positive on each of the 20 M-CHAT(-R) items across ten countries are given in Table 2. All 20 items had a moderate or high endorsement rate for Mexico, 17 for Macedonia and Turkey, 16 for Albania, 15 for Georgia, 14 for Serbia, 12 for Chile, 11 for UK and USA, whereas only one item had a moderate rate for Malaysia, and other items had low. There was no item with a moderate to high endorsement rate across all 10 countries; all items had a low endorsement rate in at least one country and moderate to high in others.

There were 14 items with a moderate to high endorsement rate in seven to nine countries. Of these, four items had moderate to high across nine countries: *points to show* and *understands what is said* had a high endorsement rate for six (and moderate for three), and *follows your gaze* and *gets a parent to watch* had high for five (and moderate for four). Items *points to get help*, *brings things to show* and *follows a point* had a moderate to high endorsement rate in all countries excluding Malaysia. The other items with many moderate or high endorsements were *interest in other children* (for 8 countries), *plays pretend* (for 8), *imitates action* (for 8), *responds to name* (for 7), *unusual finger movements* (for 7), *hearing concerns* (for 7), and *social referencing* (for 8).

Table 3 presents the 20 M-CHAT(-R) items with a potential cultural difference or universal key indicators across the low-risk and high-risk groups. The data from Malaysia were omitted from

this analysis since the low endorsement rate appeared to be an outlier relative to all other countries' data. The ratio between the number of items as universal key indicators and a potential cultural difference was 8/9 in the low-risk sample group and 11/6 in the high-risk. Items *points to get help*, *points to show*, *brings things to show*, *follows a point*, *follows your gaze*, and *understands what is said* had moderate to high endorsement rates across the countries in the groups. Items *makes eye contact, responds to name*, *hearing concerns*, and *reciprocal smile* had low endorsement rates in at least one country and a moderate to high in the other(s) across the groups.

Table 2
Item endorsement (%) for each M-CHAT(-R) item

		Low-risk				Mixed low-/high- risk		High-risk			
Item	Mod-high*	Alb n = 17	Chi n = 15	Mal n = 52	USA n = 312	Mex n = 46	Ser n = 30	Geo n = 189	Mac n = 95	Tur n = 150	UK n = 28
Enjoys movement activities	2	17.6	13.3	0	4.2	39.1	13.8	11.1	27.4	69.3	7.4
Interest in other children	8	58.8	20.0	23.1	32.0	60.9	63.3	51.3	70.2	76.7	53.6
Climbs on things	1	29.4	0	3.8	5.4	54.3	10.0	6.3	20.0	19.3	3.7
Plays pretend	8	70.6	66.7	19.2	49.7	45.7	53.3	51.9	96.8	96.7	21.4
Points to get help	9	41.2	46.7	25.0	63.7	47.8	73.3	53.4	98.9	94.7	42.9
Points to show	9	76.5	73.3	21.2	68.7	47.8	96.7	58.7	96.8	88.0	44.4
Brings things to show	9	88.2	33.3	23.1	45.7	41.3	69.0	56.1	98.9	89.3	42.3
Makes eye contact	5	58.8	20.0	15.4	19.2	39.1	36.7	15.9	86.3	92.0	17.9
Upsets by everyday noises	6	17.6	33.3	25.0	28.6	58.7	10.0	61.9	76.8	43.3	63.0
Reciprocal smile	3	58.8	20.0	11.5	9.4	45.7	13.3	10.1	72.6	28.7	7.7
Imitates actions	8	58.8	13.3	26.9	34.5	63.0	57.1	50.3	96.8	59.3	48.1
Responds to name	7	70.6	40.0	26.9	27.9	34.8	48.3	44.4	94.7	45.3	28.6
Follows a point	9	70.6	60.0	28.8	43.8	45.7	76.7	54.0	93.7	89.3	42.9
Walks	1	5.9	0	7.7	8.7	30.4	3.3	1.1	2.1	4.7	3.6
Follows your gaze	9	94.1	53.3	17.3	54.0	71.7	86.7	39.7	100	99.3	46.4
Unusual finger movements	7	58.8	33.3	17.3	26.1	47.8	21.4	70.9	76.8	43.3	40.7
Gets a parent to watch	9	94.1	20.0	30.8	56.1	63.0	62.1	45.5	100	96.7	32.1
Hearing concerns	7	52.9	60.0	17.3	25.4	45.7	40.0	74.6	88.4	93.3	25.0
Understands what is said	9	70.6	73.3	23.1	48.2	54.3	78.6	70.9	100	94.7	42.3

Social referencing	8	76.5	46.7	26.9	45.0	52.2	53.3	45.2	96.8	64.0	25.0

Note: *Number of countries out of 10 with 30% endorsement or higher. Alb – Albania; Chi – Chile, Geo – Georgia, Mac – Macedonia, Mal – Malaysia, Mex – Mexico, Ser – Serbia, Tur – Turkey, UK – United Kingdom, USA – United States. Light shading indicates moderate endorsement of 30-60% and dark shading high endorsement of > 60%.

Table 3

M-CHAT(-R) items with potential cultural differences or being universal key indicators across three sample groups

Item	Low-risk	High-risk
Points to get help	U	U
Points to show	U	U
Brings things to show	U	U
Follows a point	U	U
Follows your gaze	U	U
Understands what is said	U	U
Interest in other children	C	U
Imitates actions	C	U
Gets a parent to watch	C	U
Plays pretend	U	C
Social referencing	U	C
Unusual finger movements	C	U
Upsets by everyday noises	C	U
Makes eye contact	C	C
Responds to name	C	C
Hearing concerns	C	C
Reciprocal smile	C	C
Climbs on things	UL	UL
Walks	UL	UL
Enjoys movement activities	UL	C

Note: C **-** potential cultural difference (i.e., a low endorsement rate in at least one country and a moderate to high in the other(s) in the group); U – universal key indicator (i.e., a moderate to high endorsement rate across all countries in the group); UL – a universally low endorsement rate across all countries in the group.

Discussion

To clarify the extent to which cross-cultural differences in ASD symptoms are due to measurement methods or true, inherent characteristics of the disorder (i.e., the cross-cultural variability), studies are needed that evaluate whether autistic symptoms are equally reported cross-culturally with existing instruments. Testing the M-CHAT(-R) data from parents/primary caregivers of toddlers in Albania, Chile, Georgia, Macedonia, Malaysia, Mexico, Serbia, Turkey, UK and USA, this study found that within each country and across the countries endorsement rates of at-risk responses varied for all items related to typical behaviors affected in toddlers with ASD as well as atypical behaviors common in ASD. Countries that had the highest number of items with moderate to high endorsement rates (17-20 items) were Mexico, Macedonia, and Turkey; countries with fewer, but more than half of the items endorsed at moderate to high rates (11-16 items) included Albania, Georgia and Serbia, then Chile, UK, and the USA. In contrast, for the data from Malaysia, only one item had a moderate endorsement rate, while all other items were endorsed by fewer than 30% of parents. This finding suggests that the data for Malaysia are an outlier compared to data from other countries. This may not only indicate the variation in the expression and reporting autistic symptoms/behaviors among the parents/caregivers in this country, which could be related to different cultural aspects such as stigma and less willingness to report symptoms, but also on the characteristics of the sample and its assessment, or even a measurement error, even though the internal consistency of the data included was good (.78).

To evaluate cultural sensitivity, individual items were examined to determine whether the endorsement rate was similar across countries. Although all items had a low endorsement rate in at least one country and a moderate to high in the other(s), which could suggest potential cultural difference (Carruthers et al., 2018), 14 items had a moderate to high endorsement rate in seven or more countries; thus, these items represent candidates for universal key indicators. Of particular relevance in this regard are items with moderate to high endorsement rate in all countries excluding Malaysia, *points to get help*, *points to show*, *brings things to show*, *follows a point*, *follows your gaze*, and *understands what is said*. On the other hand, the results of the additional analysis considering whether the sample was low- or high-risk, also confirmed that these six items are universal key indicators. In contrast, items *makes eye contact, responds to name*, *hearing concerns*, and *reciprocal smile* are interpreted to have a potential cultural difference. This analysis also

showed that parents/caregivers across countries reported some ASD symptoms differently depending on the level of risk; *plays pretend* and *social referencing* appeared more culturally different in the high-risk toddlers, whereas *interest in other children*, *imitates actions*, *gets a parent to watch*, *unusual finger movements*, and *upsets by everyday noises* differed among low-risk samples. These findings together suggest that the perception of the M-CHAT(-R) items by parents of toddlers with ASD not only varies markedly across countries, but also that the perception may be culturally sensitive, in line with one previous study showing that items on another toddler screener (Matson et al., 2017) were endorsed differently across Greece, Italy, Japan, Poland, and the USA. However, the level of risk or suspicion for ASD may also affect the reporting in that parents for high-risk toddlers tend to observe the symptoms more universally.

The items that appear to be less culturally sensitive broadly represent joint attention, imitation, social engagement, and language comprehension. Impairment in these domains is more likely endorsed by parents/primary caregivers from different cultures/countries as common atypical behaviors in toddlers with ASD. Although using a different approach from studying the rate of endorsement, Kamio et al. (2015) in the Japanese version of the M-CHAT(-R) also identified item *points to get help* (i.e., imperative pointing) and *understands what is said* to be best discriminating for ASD, while Darquah (2021) identified only item *understands what is said*. The item proto-declarative pointing (i.e., *points to show*) was recognized as one of the most important items within a theoretical framework for ASD screening (Magán-Maganto et al., 2018). Comparing data across several countries for the Quantitative Checklist for Autism in Toddlers (Q-CHAT; Allison et al., 2008), it was observed that items reporting imperative and declarative pointing, as in this study, besides gestures and staring at nothing, were common for toddlers with ASD across different language/cultural groups (Stevanović, 2021).

There are several explanations for our findings. First, differences in responding to the items of the M-CHAT(-R) may reflect different values and norms placed on specific ASD behaviors vs. expected developmental behaviors by parents/ caregivers from different societies or cultures (e.g., Bernier, Mao, & Yen, 2010; Carruthers et al., 2018; Norbury & Sparks, 2013). For example, compared to White parents, Black parents reported significantly fewer ASD concerns and fewer social and restricted and repetitive behavior concerns (Donohue, Childs, Richards, & Robins, 2019). Another example is that the age of parental concern for expression of ASD may vary

especially comparing low- and middle-income countries with high-income countries (Samms-Vaughan, 2014). In addition, local beliefs and knowledge about ASD could also affect the reporting (Norbury & Sparks, 2013), literacy levels, lack of knowledge about milestones in general, and the possibility of parents providing socially acceptable responses (Marlow et al., 2019). Second, differences in responding could represent the construct inequivalence in conceptualizing ASD behaviors with the M-CHAT(-R), linked to genuine differences in expressing ASD symptoms at various levels among children across cultures/nations (e.g., de Leeuw et al., 2020) and how ASD is conceptualized and measured across different societies (e.g., Zaroff & Uhm, 2012). This latter hypothesis may best be explored by testing simultaneously cross-cultural measurement invariance of the M-CHAT(-R). Third, partially linked to the previous point, it could be that the M-CHAT(-R) translations were not culturally adapted for the languages and compared rigorously to ensure conceptual, but also linguistic equivalence, which is one of the most relevant concerns when screeners are to be used in another language/culture (DuBay & Watson, 2019). Finally, impairments in joint attention and social engagement may be more stable signs of ASD in toddlers and more readily reported by parents from different cultures.

The present study has several limitations that should be considered when interpreting the results. Firstly, we used already available data and not data simultaneously collected. In this regard, we could not consider and control for sociodemographic and clinical data for participants and their parents/ caregivers, including race, sex, and children's ASD symptom expression and severity, for example, by an objectively measured/observed instrument such as the ADOS (Lord et al., 2000). It could be possible that children with different cognitive profiles or behavioral difficulties were included in the original studies/clinical data, which also may affect the reporting (e.g., Christopher, Bishop, Carpenter, Warren, & Kanne, 2020), since M-CHAT(-R) scores are affected by the presence of developmental concern (Yuen, Penner, Carter, Szatmari, & Ungar, 2018). A significant limitation related to this aspect is that final ASD diagnosis was based on different criteria across samples, including DSM-IV, DSM-5, ICD-10, and in one case (Albania) based only on the ADOS. In addition, this study is unable to ascertain whether interpretation of ASD symptoms varies across cultures. In addition, because this study's sample aggregated data from several research and clinical sources, it was not possible to examine the endorsements rates of M-CHAT(-R) items in children who were not diagnosed with ASD. Therefore, observed differences across countries could reflect differences in clinical presentation across samples, besides being

actual possible cultural differences. Another limitation is not including the Follow-Up in the analyses since the data were not available for most of the included studies. The Follow-Up may decrease the false positive rates and likely affect the endorsement rates. Recruitment strategies (e.g., screening in high-risk vs. low-risk settings) and inclusion criteria for evaluations (e.g., evaluating randomly selected screen negative cases, children who demonstrated risk other than screening positive on M-CHAT(-R/F), or only evaluating positive screen cases will impact how the tool performed in children who were not diagnosed with ASD. Future studies should examine screener performance, including discriminatory power (Carruthers et al., 2018), in children with other neurodevelopmental delays and typical development, ensuring that samples are comparable across regions. In addition, it was not possible to evaluate the effects of age due to a small number of participants in some countries, an important issue to consider in early ASD screening (Toh et al., 2018; Sánchez-García, Galindo-Villardón, Nieto-Librero, Martín-Rodero, & Robins, 2019), as well as gender effects considering the differences observed in ASD screening with older children (e.g., Samadi, Mohammad, Ghanimi, & McConkey, 2016). Similarly, data collected during different time periods could also be affected by new knowledge available to the public. We also could not evaluate the appropriateness and quality of the clinical assessments since this was not available from all studies even after correspondence with the authors providing the data. Finally, the measurement invariance across the countries was not assessed due to a different number of included subjects.

Conclusion

There were differences in responding to M-CHAT(-R) items related to typical behaviors affected in toddlers with ASD and atypical behaviors common in ASD by parents/primary caregivers across ten countries, which may indicate cross-country variations in the recognition and evaluation of early autistic symptoms in toddlers. Items related to joint attention, imitation, social engagement, and language comprehension may be less variable and potentially interpreted as universal atypical behaviors in toddlers with ASD. The main implications of this study are that ASD screening with the M-CHAT(-R) needs to consider culturally sensitive aspects to improve screening accuracy and to complement the M-CHAT(-R), such as adding additional items or explanations to these 20, but also that finding a universally valid short screener is highly unlikely.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Declaration of interest

Lilia Albores-Gallo receives remuneration for training in the assessment of autism with the CRIDI interview. Dr. Diana Robins is a co-owner of M-CHAT, LLC. M-CHAT, LLC licenses the use of their intellectual property, the Modified Checklist for Autism in Toddlers (M-CHAT) and M-CHAT Revised, with Follow-Up (M-CHAT-R/F), for use in commercial products and collects royalties. Dr. Robins has a 50% share in the LLC. She also is on the advisory board for Quadrant Biosciences, Inc. All other authors have no conflicts of interest.

Author contributions

Dejan Stevanovic: Conceptualization, Methodology, Validation, Formal analysis, Resources, Data Curation, Writing - Original Draft, Project administration; **Diana L. Robins**: Methodology, Resources, Investigation, Data Curation, Writing - Review & Editing, Supervision; **Floriana Costanzo**: Methodology, Data Curation, Visualization, Writing - Review & Editing, Supervision, Project administration; **Elisa Fucà**: Methodology, Data Curation, Visualization, Writing - Review & Editing, Supervision, Project administration; **Giovanni Valeri**: Data Curation, Visualization, Writing - Review & Editing, Supervision, Project administration; **Stefano Vicari**: Methodology, Data Curation, Visualization, Writing - Review & Editing, Supervision, Project administration; **Handan Ozek Erkuran**: Resources, Investigation, Data Curation, Visualization, Writing - Review & Editing; **Ferhat Yaylaci**: Resources, Investigation, Data Curation, Visualization, Writing - Review & Editing; **Lilia Albores-Gallo**: Resources, Investigation, Data Curation, Visualization, Writing - Review & Editing; **Gabriel Gatica-Bahamonde**: Resources, Investigation, Data Curation, Visualization, Writing - Review & Editing; **Maia Gabunia**: Resources, Investigation, Data Curation, Visualization, Writing - Review & Editing, Project administration; **Medea Zirakashvili**: Resources, Investigation, Data Curation, Visualization,

Writing - Review & Editing; **Tony Charman**: Resources, Investigation, Data Curation, Visualization, Writing - Review & Editing; **Sayyed Ali Samadi**: Resources, Investigation, Data Curation, Visualization, Writing - Review & Editing; **Teck-Hock TOH**: Resources, Investigation, Data Curation, Visualization, Writing - Review & Editing; **Windham Gayle**: Resources, Investigation, Data Curation, Visualization, Writing - Review & Editing; **Laura Brennan**: Resources, Investigation, Data Curation, Visualization, Writing - Review & Editing; **Tatjana Zorcec**: Resources, Investigation, Data Curation, Visualization, Writing - Review & Editing; **Alejandra Auza**: Investigation, Data Curation, Visualization, Writing - Review & Editing, Supervision; **Maretha de Jonge**: Investigation, Data Curation, Visualization, Writing - Review & Editing, Supervision; **Noordeen Shoqirat**: Data Curation, Visualization, Writing - Review & Editing; **Rajna Knez**: Investigation, Visualization, Writing - Review & Editing, Supervision, Project administration.

References

Abubakar, A., Sewanyana, D., & Newton, C. R. (2016). A systematic review of research on autism spectrum disorders in sub-Saharan Africa. *Behavioural Neurology*, 2016, 3501910.

Albores-Gallo, L., Roldán-Ceballos, O., Villarreal-Valdes, G., Betanzos-Cruz, B. X., Santos-Sánchez, C., Martínez-Jaime, M. M., ... & Hilton, C. L. (2012). M-CHAT Mexican version validity and reliability and some cultural considerations. *International Scholarly Research Notices*, *2012*, 408694.

Allison, C., Baron-Cohen, S., Wheelwright, S., Charman, T., Richler, J., Pasco, G., & Brayne, C. (2008). The Q-CHAT (Quantitative CHecklist for Autism in Toddlers): A normally distributed quantitative measure of autistic traits at 18–24 months of age: Preliminary report. *Journal of Autism and Developmental Disorders*, *38*(8), 1414-1425.

American Psychiatric Association (2000). *Diagnostic and statistical manual of mental disorders (4th ed., Text Revision)*. Washington, DC: American Psychiatric Publishing.

American Psychiatric Association (2013). *Diagnostic and statistical manual of mental disorders (DSM–5)*. Arlington, VA: American Psychiatric Publishing.

Auyeung, B., Baron-Cohen, S., Wheelwright, S., & Allison, C. (2008). The autism spectrum quotient: Children's version (AQ-Child). *Journal of Autism and Developmental Disorders*, 38(7), 1230-1240.

Baduel, S., Guillon, Q., Afzali, M. H., Foudon, N., Kruck, J., & Rogé, B. (2017). The French version of the Modified-Checklist for Autism in Toddlers (M-CHAT): a validation study on a French sample of 24 month-old children. Journal of autism and developmental disorders, 47(2), 297-304.

Bernier, R., Mao, A., & Yen, J. (2010). Psychopathology, families, and culture: autism. *Child and Adolescent Psychiatric Clinics*, *19*(4), 855-867.

Brennan, L., Fein, D., Como, A., Rathwell, I. C., & Chen, C. M. (2016). Use of the modified checklist for autism, revised with follow up-Albanian to screen for ASD in Albania. *Journal of autism and developmental disorders*, *46*(11), 3392-3407.

Carruthers, S., Kinnaird, E., Rudra, A., Smith, P., Allison, C., Auyeung, B., ... & Hoekstra, R. A. (2018). A cross-cultural study of autistic traits across India, Japan and the UK. *Molecular Autism*, 9(1), 52.

Charman, T., Baird, G., Simonoff, E., Chandler, S., Davison-Jenkins, A., Sharma, A., ... & Pickles, A. (2016). Testing two screening instruments for autism spectrum disorder in UK community child health services. *Developmental Medicine & Child Neurology*, *58*(4), 369-375.

Chlebowski, C., Robins, D. L., Barton, M. L., & Fein, D. (2013). Large-scale use of the modified checklist for autism in low-risk toddlers. *Pediatrics*, *131*(4), e1121-e1127.

Christopher, K., Bishop, S., Carpenter, L. A., Warren, Z., & Kanne, S. (2020). The Implications of Parent-Reported Emotional and Behavioral Problems on the Modified Checklist for Autism in Toddlers. *Journal of Autism and Developmental Disorders*, *51(3),* 884-891.

Dai, Y. G., Porto, K. S., Skapek, M., Barton, M. L., Dumont-Mathieu, T., Fein, D. A., & Robins, D. L. (2021). Comparison of the modified checklist for autism in toddlers, revised with follow-up (M-CHAT-R/F) positive predictive value by race. *Journal of Autism and Developmental Disorders*, *51*(3), 855-867.

Darquah, E. A. (2021). *Cultural Appropriateness of the Modified Checklist for Autism for Screening Autism in Ghana*. Doctoral thesis. https://scholarworks.waldenu.edu/dissertations/9852/

de Leeuw, A., Happé, F., & Hoekstra, R. A. (2020). A conceptual framework for understanding the cultural and contextual factors on autism across the globe. *Autism Research*. https://doi.org/10.1002/aur.2276.

Dimitrov, D.M. (2010). Testing for factorial invariance in the context of construct validation. *Measurement and Evaluation in Counseling and Development*, *43*(1), 121-149.

Donohue, M. R., Childs, A. W., Richards, M., & Robins, D. L. (2019). Race influences parent report of concerns about symptoms of autism spectrum disorder. *Autism*, *23*(1), 100-111.

DuBay, M., & Watson, L. R. (2019). Translation and cultural adaptation of parent-report developmental assessments: Improving rigor in methodology. *Research in Autism Spectrum Disorders*, *62 (1)*, 55-65.

Durkin, M. S., Elsabbagh, M., Barbaro, J., Gladstone, M., Happe, F., Hoekstra, R. A., ... & Tager-Flusberg, H. (2015). Autism screening and diagnosis in low resource settings: challenges and opportunities to enhance research and services worldwide. *Autism Research*, 8(5), 473-476.

Falkmer, T., Anderson, K., Falkmer, M., & Horlin, C. (2013). Diagnostic procedures in autism spectrum disorders: a systematic literature review. *European child & adolescent, 22(6)*, 329-40.

Gatica-Bahamonde, G., Alarcon, P., Fadol, A. M., Allison, C., Baron-Cohen, S., & Roman-Urrestarazu, A. (2021). The Quantitative Checklist for Autism in Toddlers (Q-CHAT): A preliminary psychometric study in Chile. In press

Gillis, J. M., Callahan, E. H., & Romanczyk, R. G. (2011). Assessment of social behavior in children with autism: the development of the Behavioral Assessment of Social Interactions in Young Children. *Research in Autism Spectrum Disorders*, *5*(1), 351-360.

Guo, C., Luo, M., Wang, X., Huang, S., Meng, Z., Shao, J., ... & Jing, J. (2019). Reliability and validity of the Chinese version of modified checklist for autism in toddlers, revised, with follow-up (M-CHAT-R/F). Journal of autism and developmental disorders, 49(1), 185-19

Harris, B., Barton, E. E., & Albert, C. (2014). Evaluating autism diagnostic and screening tools for cultural and linguistic responsiveness. *Journal of Autism and Developmental Disorders*, *44*(6), 1275-1287.

Harrison, A. J., Long, K. A., Tommet, D. C., & Jones, R. N. (2017). Examining the role of race, ethnicity, and gender on social and behavioral ratings within the Autism Diagnostic Observation Schedule. *Journal of Autism and Developmental Disorders*, 47(9), 2770-2782.

Kamio, Y., Haraguchi, H., Stickley, A., Ogino, K., Ishitobi, M., & Takahashi, H. (2015). Brief report: Best discriminators for identifying children with autism spectrum disorder at an 18-month health check-up in Japan. *Journal of Autism and Developmental Disorders*, *45*(12), 4147-4153.

Khowaja, M. K., Hazzard, A. P., & Robins, D. L. (2015). Sociodemographic barriers to early detection of autism: screening and evaluation using the M-CHAT, M-CHAT-R, and follow-up. *Journal of autism and developmental disorders*, 45(6), 1797-1808.

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Leventhal, B. L., DiLavore, P. C., ... & Rutter, M. (2000). The Autism Diagnostic Observation Schedule—Generic: A standard measure of social and communication deficits associated with the spectrum of autism. *Journal of Autism and Developmental Disorders*, *30(3)*, 205-223.

Magán-Maganto, M., Jónsdóttir, S. L., Sánchez-García, A. B., García-Primo, P., Hellendoorn, A., Charman, T., ... & Canal-Bedia, R. (2018). Building a theoretical framework for autism spectrum disorders screening instruments in Europe. *Child and Adolescent Mental Health*, *23*(4), 359-367.

Marlow, M., Servili, C., & Tomlinson, M. (2019). A review of screening tools for the identification of autism spectrum disorders and developmental delay in infants and young children: recommendations for use in low-and middle-income countries. *Autism Research*, *12*(2), 176-199.

Matson, J. L., Boisjoli, J., & Wilkins, J. (2007). *Baby and infant screen for children with aUtIsm traits (BISCUIT)*. Baton Rouge, LA: Disability Consultants, LLC.

Matson, J.L., & Gonzalez, M.L. (2007). *Autism Spectrum Disorders – Diagnosis – Child Version*. Baton Rouge, La. Disability Consultants, LLC.

Matson, J. L., Matheis, M., Burns, C. O., Esposito, G., Venuti, P., Pisula, E., ... & Ishitobi, M. (2017). Examining cross-cultural differences in autism spectrum disorder: A multinational comparison from Greece, Italy, Japan, Poland, and the United States. *European Psychiatry, 42(1)*, 70-76.

Matson, J. L., Worley, J. A., Fodstad, J. C., Chung, K. M., Suh, D., Jhin, H. K., ... & Furniss, F. (2011). A multinational study examining the cross-cultural differences in reported symptoms of autism spectrum disorders: Israel, South Korea, the United Kingdom, and the United States of America. *Research in Autism Spectrum Disorders, 5(4),* 1598-1604.

Mohamed, F. E., Zaky, E. A., Youssef, A., Elhossiny, R., Zahra, S., Khalaf, R., ... & Eldin, W. S. (2016). Screening of Egyptian toddlers for autism spectrum disorder using an Arabic validated version of M-CHAT; report of a community-based study (Stage I). *European Psychiatry*, *34*, 43-48.

Norbury, C. F., & Sparks, A. (2013). Difference or disorder? Cultural issues in understanding neurodevelopmental disorders. *Developmental Psychology*, 49(1), 45-58.

Oner, O., & Munir, K. M. (2019). Modified Checklist for Autism in Toddlers Revised (MCHAT-R/F) in an Urban Metropolitan Sample of Young Children in Turkey. *Journal of Autism and Developmental Disorders*, *50(9):*3312-3319.

Petrocchi, S., Levante, A., & Lecciso, F. (2020). Systematic review of level 1 and level 2 screening tools for autism spectrum disorders in toddlers. *Brain Sciences, 10*(3), 180.

Randall, M., Egberts, K. J., Samtani, A., Scholten, R. J., Hooft, L., Livingstone, N., ... & Williams, K. (2018). Diagnostic tests for autism spectrum disorder (ASD) in preschool children. *Cochrane Database of Systematic Reviews*, 24 (7), CD009044.

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*(1), 37-45.

Robins, D.L., Fein, D., & Barton, M. (1999). *The Modified Checklist for Autism in Toddlers (M-CHAT)*. Self-published. www.mchatscreen.com

Robins, D.L., Fein, D., & Barton, M. (2009). *The Modified Checklist for Autism in Toddlers, Revised, with Follow-up (M-CHAT-R/F)*. Self-published. www.mchatscreen.com

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.

Samadi, S. A., & McConkey, R. (2015). Screening for autism in Iranian preschoolers: Contrasting M-CHAT and a scale developed in Iran. *Journal of autism and developmental disorders*, *45*(9), 2908-2916.

Samadi, S. A., Mohammad, M. P., Ghanimi, F., & McConkey, R. (2016). The challenges of screening pre-school children for autism spectrum disorders in Iran. Disability and rehabilitation, 38(17), 1739-1747.

Samms-Vaughan, M. E. (2014). The status of early identification and early intervention in autism spectrum disorders in lower-and middle-income countries. *International Journal of Speech-Language Pathology*, *16*(1), 30-35.

Sánchez-García, A. B., Galindo-Villardón, P., Nieto-Librero, A. B., Martín-Rodero, H., & Robins, D. L. (2019). Toddler screening for autism spectrum disorder: A meta-analysis of diagnostic accuracy. *Journal of Autism and Developmental Disorders*, *49*(5), 1837-1852.

Schopler, E., Reichler, R. J., DeVellis, R. F., & Daly, K. (1980). Toward objective classification of childhood autism: Childhood autism rating scale (CARS). Journal of Autism and Developmental Disorders, 10(1), 91–103.

Stevanovic, D. (2018). Autism Spectrum Disorder International Consortium (ASDIC): Moving Towards Sustainable and Quality of Life-Improving Services for Autism in Low and Middle Income Regions. *IACAPAP Bulletin*, 51, 90-91.

Stevanović, D. (2021). Quantitative Checklist for Autism in Toddlers (Q-CHAT): A psychometric study with Serbian Toddlers. *Research in Autism Spectrum Disorders*, *83*, 101760.

Stevanovic, D., Costanzo, F., Fuc`a, F., Valeri, G., Vicari, S., Robins, D. L., … Knez, R. (2021). Measurement invariance of the Childhood Autism Rating Scale (CARS) across six countries. *Autism Research*, doi: 10.1002/aur.2586.

Toh, T. H., Tan, V. W. Y., Lau, P. S. T., & Kiyu, A. (2018). Accuracy of Modified Checklist for Autism in Toddlers (M-CHAT) in detecting autism and other developmental disorders in community clinics. *Journal of Autism and Developmental Disorders*, *48*(1), 28-35.

Wieckowski, A. T., Hamner, T., Nanovic, S., Porto, K. S., Coulter, K. L., Eldeeb, S. Y., ... & Robins, D. L. (2021). Early and repeated screening detects autism spectrum disorder. *Journal of Pediatrics*, 234, 227-235.

Windham, G. C., Smith, K. S., Rosen, N., Anderson, M. C., Grether, J. K., Coolman, R. B., & Harris, S. (2014). Autism and developmental screening in a public, primary care setting primarily serving Hispanics: challenges and results. *Journal of Autism and Developmental Disorders*, *44*(7), 1621-1632.

World Health Organization (WHO) (1992). *The ICD-10 classification of mental and behavioural disorders: Clinical descriptions and diagnostic guidelines*. Geneva: World Health Organization.

Yama, B., Freeman, T., Graves, E., Yuan, S., & Karen Campbell, M. (2012). Examination of the properties of the Modified Checklist for Autism in Toddlers (M-CHAT) in a population sample. Journal of autism and developmental disorders, 42(1), 23-34.

Yuen, T., Penner, M., Carter, M. T., Szatmari, P., & Ungar, W. J. (2018). Assessing the accuracy of the Modified Checklist for Autism in Toddlers: a systematic review and meta-analysis. *Developmental Medicine & Child Neurology*, *60*(11), 1093-1100.

Zaroff, C. M., & Uhm, S. Y. (2012). Prevalence of autism spectrum disorders and influence of country of measurement and ethnicity. *Social psychiatry and psychiatric epidemiology*, 47(3), 395-398.