Received: 17 May 2021 | Accepted: 15 October 2021

DOI: 10.1002/aur.2631

RESEARCH ARTICLE

Development and psychometric evaluation of a Thai Diagnostic Autism Scale for the early diagnosis of Autism Spectrum Disorder

Duangkamol Tangviriyapaiboon[1] | **Samai Sirithongthaworn**[2] | **Hataichanok Apikomonkon**[3] | **Chidawan Suyakong**[3] | **Pimwarat Srikummoon**[4,5] | **Suttipong Kawilapat**[4] | **Patrinee Traisathit**[4,5,6]

[1]Rajanagarindra Institute of Child Development, Chiang Mai, Thailand

[2]Department of Mental Health, Ministry of Public Health, Nonthaburi, Thailand

[3]Department of Occupational Therapy, Faculty of Associated Medical Sciences, Chiang Mai University, Chiang Mai, Thailand

[4]Department of Statistics, Faculty of Science, Chiang Mai University, Chiang Mai, Thailand

[5]Department of Statistics, Faculty of Science, Data Science Research Center, Chiang Mai University, Chiang Mai, Thailand

[6]Department of Statistics, Faculty of Science, Research Center in Bioresources for Agriculture, Industry and Medicine, Chiang Mai University, Chiang Mai, Thailand

Correspondence
Duangkamol Tangviriyapaiboon, Rajanagarindra Institute of Child Development, Thailand 196, Moo 10, Don Kaeo Sub-district, Mae Rim District, Chiang Mai, Thailand, 50180.
Email: dtangviriyapaiboon@gmail.com

Funding information
Chiang Mai University; Health Systems Research Institute; Department of Mental Health, Ministry of Public Health

Abstract

The Thai Diagnostic Autism Scale (TDAS) was developed for use as a diagnostic tool for the early diagnosis of Autism Spectrum Disorder (ASD) in Thai children aged 12–48 months old. TDAS consists of 23 items (13 and 17 items in the observational and interview sections, respectively) classified into seven domains (A1–A3 and B1–B4) according to the criteria in the Diagnostic and Statistical Manual of Mental Disorder, fifth edition (DSM-5). Children with a single score in the A1–A3 domains and at least two of the B1–B4 domains were classified with ASD. The item-objective congruence (IOC) index, confirmatory factor analysis, and Kappa coefficient were used to evaluate the content, constructs, and inter-rater validity levels between the evaluators and concurrent validity between TDAS and physicians' diagnoses, respectively. TDAS showed good overall content validity (IOC range 0.71–1.00), suitable construct validity (root-mean-squared errors of approximation of 0.076 and 0.067, comparative fit indexes of 0.902 and 0.858, and Tucker-Lewis indexes of 0.882 and 0.837 for the observation and interview sections, respectively), and excellent diagnostic agreement between TDAS and the evaluators (Kappa = 1.000) as well as between TDAS and the physicians' diagnoses (Kappa = 0.871). The sensitivity and specificity of TDAS were 100% and 82.4%, respectively. In conclusion, TDAS yielded a high level of content validity, concurrent validity, and inter-rater reliability for the early diagnosis of ASD in Thai children. A large-scale study using TDAS is needed to determine an appropriate cut-off point as well as its efficacy.

Lay Summary

The Thai Diagnostic Autism Scale was developed for use as a diagnostic tool for the early diagnosis of Autism Spectrum Disorder (ASD) among Thai children. It contains 23 items in seven domains for the screening via observations and interviews. The psychometric properties of this diagnostic tool provide its reliability and suitability for the early diagnosis of ASD. A large-scale study using it is needed to determine an appropriate cut-off point as well as its efficacy.

KEYWORDS
Autism Spectrum Disorder, diagnostic tool, early diagnosis, Thai Diagnostic Autism Scale

INTRODUCTION

Autism Spectrum Disorder (ASD) is a complex developmental condition characterized by persistent deficits in social communication and social interaction across multiple contexts coupled with restricted and repetitive patterns of behavior, interests, or activities (Huerta et al., 2012). The number of children with ASD in Thailand has been rising similarly to other countries, with a prevalence rate of 7 per 10,000 in 2002 (Plubrukarn et al., 2005) and increasing to 10 per 10,000 in 2005 (Poolsuppasit et al., 2005). Recently, it has been estimated that there are 370,000 Thai people with ASD, with only 15% of them being able to access healthcare (Department of Mental Health, 2014). It is known that early intervention can improve the overall functioning of children with ASD, and thus, its early identification is invaluable (Ben-Itzchak & Zachor, 2007).

Over the past two decades, many standardized screening/symptom measurement tools (e.g., the Modified Checklist for Autism in Toddlers (M-CHAT) (Robins et al., 2001) the M-CHAT revised with follow-up (M-CHAT-R/F) (Robins et al., 2014), the Pervasive Developmental Disorders Screening Questionnaire (PDDSQ) (Pornnoppadol et al., 2002), the Rapid Interactive Screening Test for Autism in Toddlers (RITA-T) (Choueiri & Wagner, 2015), the Infant-Toddler Checklist (ITC) (Wetherby et al., 2008), the Screening Tool for Autism in Two-year-olds (STAT) (Stone et al., 2008), the Childhood Autism Rating Scale (CARS) (Kaat & Lecavalier, 2013; Schopler et al., 1988), and the Gilliam Autism Rating Scale, second edition (GARS-2) (Robinson, 2013) have been developed for the early detection of ASD and subsequent early intervention. Moreover, diagnostic tools such as the Autism Diagnostic Observation Schedule, second edition (ADOS-2) (Carr, 2013; Lord et al., 2012), the Diagnostic Interview for Social and Communication Disorders, ninth edition (DISCO-9) (Leekam, 2013; Leekam et al., 2002; Wing et al., 2002), the Developmental Dimensional and Diagnostic Interview-short version (3Di-sv) (Slappendel et al., 2016), and the Autism Diagnostic Interview-Revised (ADI-R) (Kim et al., 2013; Rutter et al., 2003) have been developed for the confirmation of ASD. However, these tools have limitations. For example, ADOS shows relatively low sensitivity when used to diagnose higher functioning adults with ASD and low specificity in individuals with severe intellectual disability when used without ADI-R (Bastiaansen et al., 2011). Moreover, some of the tools have not been extensively validated in individuals with ASD and intellectual disability in non-western cultures (Rudra et al., 2014). Despite the limitations, the need for ASD screening instruments to ensure timely, comprehensive, and systematic assessment of children with suspected ASD is essential.

In Thailand, the Developmental Surveillance and Promotion Manual (DSPM) (Ministry of Public Health, 2020) was launched to examine early childhood development at 9, 18, 30, and 42 months old. The motivation was to identify developmental problems and confirm specific areas of deficit using the Thai Early Developmental Assessment for Intervention (TEDA4I) (Department of Mental Health, 2015). Children with delayed receptive language, expressive language, and/or personal social skills are referred for ASD diagnosis. The identification of ASD is currently performed by specialist physicians in the fields of child and adolescent psychiatry, pediatrics, and general psychiatry according to the standard protocol, which is currently the diagnostic criteria of the International Classification of Diseases (ICD) or the fifth edition of Diagnostic and Statistical Manual of Mental Disorder (DSM-5) (American Psychiatric Association, 2013; World Health Organization, 1993). Even though there are several screening and diagnostic tools for ASD, its identification and assessment are challenging due to misunderstandings caused by the translation of some instruments such as the ADOS or M-CHAT, as well as interpretation and cultural differences (Varma & Iskandar, 2014). For example, children who are too young and/or living in a rural area might not respond appropriately to a question in ADOS for a child's reactions in a birthday party simulation. Therefore, in our case, some items needed to be adapted to be relevant to the Thai context and society. Another challenge is the comprehensiveness of tools in which evaluation is only based on observation via the parents' viewpoints (such as M-CHAT-R/F and PDDSQ), which might be different and lead to misleading evaluation results. Moreover, the cost of certification training is another barrier against accessing western standardized tools; it is quite high and not covered by the healthcare system in Thailand (Sirithongthaworn, 2018). Meanwhile, the current limited number of practitioners will inevitably have led to missed or late diagnoses of ASD. Therefore, our aim was to develop a new diagnostic tool accounting for both observation and interview items within the Thai context to increase the coverage of the early diagnosis of ASD in Thailand.

METHODS

The procedure for this study consists of two phases. Phase I presents the theoretical conceptualization of developing the Thai Diagnostic Autism Scale (TDAS) and checking its content validity via evaluation by experts. Phase II presents an evaluation of the psychometric properties of TDAS in terms of construct validity, inter-rater reliability of items, inter-rater reliability of ASD diagnosis, and concurrent validity. The details of each phase are provided as follows.

Phase I: TDAS development

The 11 existing tools for ASD (ADOS-2 toddler module and module 1, CARS, DISCO-9, 3Di-sv, M-CHAT, M-CHAT-R/F, PDDSQ for 1–4 years and 4–18 years, RITA-T, ITC, GARS-2, STAT, and ADI-R) were reviewed based on the following criteria: (1) keywords describing the tools (i.e., "early diagnostic", "early screening", or "early detection"), (2) a target population comprising children aged 1–5 years old, and (3) a structure consisting of interview and observational scenarios. The symptoms of ASD identified in DSM-5 were used as guidelines for constructing the TDAS items and applying them to the Thai cultural context. The initial item pool comprised 23 items according to the DSM-5 domains. To deal with the complexity of ASD diagnosis, it has been suggested that tools such as parent interviewing and observational assessment of the functioning of the children should be comprehensively included (Huerta & Lord, 2012; Richler et al., 2007; Sigman & McGovern, 2005). Thus, TDAS was designed to be used by professionals from a range of disciplines and includes both observation and interview sections.

The first expert panel of seven multidisciplinary experts comprising two child and adolescent psychiatrists, a developmental and behavioral pediatrician, a speech therapist, a psychologist, an occupational therapist, and a nurse reviewed the items in the initial item pool and identified their relevance (-1 = disagree, 0 = neutral, 1 = agree) to each domain of DSM-5 criteria. The index of item-objective congruence (IOC) was used to evaluate the content validity: an IOC score lower than 0.50 indicates that the question or evaluation method should be revised or removed (Wangkawan et al., 2020). After achieving unanimous agreement during discussions of its appropriateness for operational methodology and scoring, the preliminary version of TDAS was offered for review to a second expert panel comprising two child and adolescent psychiatrists, a developmental and behavioral pediatrician, a speech therapist, and an occupational therapist. They examined the evaluation method, interview questions, and operational techniques of the preliminary version, and the structure of TDAS was modified until unanimous agreement was reached.

Phase II: Psychometric property evaluation

The psychometric property assessment of TDAS consisted of three trials. The first was to evaluate the construct validity of the observation and interview sections separately. The second trial was conducted for the inter-rater reliability between two evaluators on the same child. The evaluators from a panel of multidisciplinary experts whose qualifications we checked included three nurses, a speech therapist, and two psychologists. Each underwent training and subsequent certification via the TDAS workshop. The course, which was developed and its principles, goals, content, activities, and media assessed for suitability by three experts, covered 13 subjects for a total of 23 h over 3 days. The final trial was conducted to identify the concurrent validity of TDAS compared to the physicians' diagnoses of ASD according to the DSM-5 criteria as the gold standard. The children were diagnosed by using TDAS and by three qualified psychiatrists within 30 days. The results of the diagnoses were double-blinded for both the TDAS evaluators and the psychiatrists.

The procedures for ASD diagnosis using TDAS

The diagnosis of ASD in children using TDAS took around 40 min. The required environment for the diagnosis comprised (1) a well-ventilated private room to prevent outside interference with the child's concentration and/or attention; (2) no stimulating paraphernalia such as pictures or objects with bright colors, or noisy devices; (3) a clean and safe environment to avoid the child harming his/herself; and (4) a comfortable set of chairs to suit the size and height of children of different ages.

The observation section was conducted directly with the child and took approximately 20 min. The 13 observational items were divided into 9 related to various situations and 4 related to behavior and interest in routine activities or narrow settings. The evaluators would give a score for each item after observing the child's behavior during 6 assessment activities: (1) free play, (2) response to name, (3) obtaining toys from the box, (4) bubble blowing, (5) cooking, and (6) peekaboo and touching.

The interview section was conducted with the parents or caregivers and took around 20 min. The 17 items were divided into 8 related to various situations and 9 related to behavior and interest in routine activities or narrow settings. The procedures in this section were (1) introducing and explaining the importance of evaluation using TDAS, giving information, and asking for consent for the assessment; (2) a brief explanation about the assessment procedure, the duration of the assessment, and video recording during assessment; and (3) recommendation and precautions during the assessment, such as turning off their mobile phones and to not speak, advise, or assist the child during the assessment except on request by the evaluators.

Participants

This study was conducted on children with suspected ASD, delayed speech, and delayed development who had been referred to the Rajanagarindra Institute of Child Development (RICD), Chiang Mai, Thailand to confirm the diagnosis in 2016. The required sample size was calculated as 228 children based on an ASD prevalence of

3.2 per 1000 Thai children (Surinkaew et al., 2005) with an acceptable random error of 3% and 99% confidence intervals. The sample size was calculated by using Daniel's formula (Daniel, 1999) as follows:

$$n = \left[Z^2_{\alpha/2} \cdot P(1-P)\right]/e^2 = \left[Z^2_{0.995}(0.032)(1-0.032)\right]/(0.03)^2 = 228.39,$$

where n is the required sample size, $Z_{1-\alpha/2}$ is the statistic corresponding to level of confidence ($Z_{0.995} = 2.576$), P is the expected prevalence of ASD, and e is the acceptable precision.

The inclusion criteria were (1) children with delayed development in receptive language, expressive language, and/or personal social skills (using TEDA4I) or (2) children with suspected ASD based on clinical signs and symptoms or a positive result from the PDDSQ. Children with hearing impairment or intractable epilepsy were excluded from the study. According to previous researchers who suggested 10 times the number of items (Cattell, 1978; Everitt, 1975; Gorsuch, 1983), an adequate sample size was calculated as 170 children based on the maximum number of items in the final version of TDAS. According to Shan (2016), the minimum sample size for inter-rater reliability is 21 subjects based on two-tailed test, significance level of 0.05 (type I error = 0.05), and power of 80% (type II error = 0.20) with minimum acceptable and expected kappa coefficient values of 0.30 and 0.80, respectively. In addition, for kappa coefficient agreement, Donner and Rotondi (2010) suggested that the number of subjects required to achieve a minimum acceptable coefficient value of 0.60 is 116 subjects when assuming that the minimum expected proportion of ASD individuals is 0.10 and an expected coefficient value of 0.80 for 2 raters. Unfortunately, the sample size for trial 3 was lower than the calculated required number of participants due to the limited availability of subjects via the recruitment process. Finally, there were 170, 21, and 62 children aged 12–48 months old and their parents included in trials 1, 2, and 3, respectively.

Statistical analysis

The demographic data of the children are presented as the mean and standard deviation or the median and interquartile range for continuous variables and as frequencies and percentages for categorical variables. Confirmatory factor analysis (CFA) was performed for the construct validity testing in trial 1. Factor loadings are presented for each item. The criterion of acceptable fit was set as a comparative fit index (CFI) of >0.90, a Tucker-Lewis index (TLI) of >0.90, and a root-mean-squared error of approximation (RMSEA) of <0.10 (Kim et al., 2016; Zhao et al., 2017). The inter-rater agreement between paired evaluators in trial 2 and the concurrent validity between the ASD diagnoses by the TDAS evaluators and the psychiatrists were calculated using the Kappa coefficient: <0.40, 0.40–0.59, 0.60–0.74, and 0.75–100 were considered as poor, fair, good, and excellent agreement, respectively (Cicchetti & Sparrow, 1981). Sensitivity and specificity were also calculated to evaluate the ability of TDAS to correctly classify children with and without ASD, respectively (Parikh et al., 2008; Trevethan, 2017). All analyses were performed using Stata 17 (StataCorp, College Station, TX, USA).

RESULTS

The development and content validity of TDAS

The content validity of the items in TDAS ranged from 0.71 to 1.00, which indicates that the content of TDAS is appropriate for use in ASD diagnosis. Thus, all of the 23 items in the initial item pool were accepted for inclusion in the TDAS. Some of the behavioral aspects of the children can be directly observed while others can be obtained from their parents via interviewing, and so the 6, 10, and 7 items were evaluated by observation only, interview only, and both operational methods, respectively. Therefore, the final version of TDAS consisted of 13 items in the observational section and 17 items in the interview section (Table 1).

According to the DSM-5 criteria for ASD, the TDAS items were classified into seven domains: (A1) deficits in social–emotional reciprocity; (A2) deficits in non-verbal communicative behaviors used for social interaction; (A3) deficits in developing, maintaining, and understanding relationships; (B1) stereotyped or repetitive motor movements, use of objects, or speech; (B2) sameness, inflexible adherence to routines, or ritualized pattern; (B3) highly restricted fixed interests that are abnormal in intensity or focus; and (B4) hyper- or hypo-reactivity to sensory input or unusual interest in sensory aspects of the environment. The administration of TDAS generally took 40 min for evaluation and 20–30 min for scoring. ASD was identified based on all single scores in domains A1–A3 and at least two single scores in domains B1–B4.

The construct validity of TDAS

Of the 170 children who participated in trial 1, 135 (79.4%) were male and most of them were 25–36 months old (45.9%), followed by 37–48 months old (40%) and 12–24 months (14.1%). Children who had been diagnosed with ASD comprised 68.8% (Table 2). The construct validity indices of TDAS were suitable in both the observation (RMSEA = 0.076; CFI = 0.902; TLI = 0.882) and interview (RMSEA = 0.067;

TABLE 1 Content validity of TDAS

Item	Content validity				Operational method	
	Agree	Neutral	Disagree	IOC	Observation	Interview
A1: Deficits in social–emotional reciprocity						
1. Getting help	7	0	0	1.00	✓	✓
2. Social initiation	7	0	0	1.00		✓
3. Sharing of interests	7	0	0	1.00	✓	
4. Social response—reciprocal conversation	7	0	0	1.00	✓	✓
5. Sharing of emotions	7	0	0	1.00		✓
A2: Deficits in communicative behaviors used for social interaction						
6. Eye contact	7	0	0	1.00	✓	
7. Non-verbal communication	7	0	0	1.00	✓	
8. Verbal communication	6	0	1	0.71	✓	
9. Understanding emotion	6	0	1	0.71		✓
A3: Deficits in developing, maintaining, and understanding relationships						
10. Developing, maintaining, and understanding relationships	6	0	1	0.71	✓	
11. Difficulties in adjusting behavior to suit the social context	6	0	1	0.71		✓
12. Sharing imaginative play	7	0	0	1.00	✓	
13. Making friends	7	0	0	1.00		✓
14. Absence or lack of interest in others	7	0	0	1.00	✓	
B1: Stereotyped or repetitive motor movements, use of objects, and/or speech						
15. Unusual movements of arms, hands, fingers, and/or the whole body	7	0	0	1.00	✓	✓
16. Repetitive manipulation of objects	7	0	0	1.00	✓	✓
17. Idiosyncratic use of speech	7	0	0	1.00	✓	✓
B2: Sameness, inflexible adherence to routines, and/or ritualized patterns						
18. Difficulty in changing or choosing activities and adherence to routines	6	0	1	0.71		✓
19. Rigid thinking	6	0	1	0.71		✓
B3: Highly restricted fixed interests that are abnormal in intensity and/or focus						
20. Interests that are abnormal in intensity and focused on the same topic, object, or activity; displaying a narrow range of interests; and/or collecting of or clinging to objects	7	0	0	1.00		✓
21. Unusual fears	7	0	0	1.00		✓
B4: Hyper- or hypo-reactivity to sensory input or unusual interest in sensory aspects of the environment						
22. Unusual response to injury and/or violence	6	1	0	0.86	✓	✓
23. Unusual response to sensory stimuli and/or exploration	7	0	0	1.00		✓

Abbreviations: IOC, item objective congruence; TDAS, the Thai Diagnostic Autism Scale.

CFI = 0.858; TLI = 0.837) sections. The item with the highest factor loading in the observation section was item 6 "Eye contact", followed by item 7 "Non-verbal communication" and item 1 "Getting help". The item with the highest factor loading in the interview section was item 13 "Making friends", followed by item 2 "Sharing of interests" and item 9 "Understanding of emotion" (Table 3).

TABLE 2 Demographics of the participants in the TDAS psychometric evaluation

Demographic	Frequency (%)		
	Trial 1 ($N = 170$)	Trial 2 ($N = 21$)	Trial 3 ($N = 62$)
Gender			
Male	135 (79.4%)	18 (85.7%)	47 (75.8%)
Female	35 (20.6%)	3 (14.3%)	15 (24.2%)
Age (months)			
12–24	24 (14.1%)	1 (4.8%)	11 (17.7%)
25–36	78 (45.9%)	9 (42.8%)	19 (30.7%)
37–48	68 (40.0%)	11 (52.4%)	32 (51.6%)
Diagnosis			
ASD	117 (68.8%)	18 (85.7%)	45 (72.6%)
Non-ASD	53 (31.2%)	3 (14.3%)	17 (27.4%)

Abbreviations: ASD, Autism Spectrum Disorder; TDAS, the Thai Diagnostic Autism Scale.

The inter-rater reliability of TDAS

Of the 21 children who participated in trial 2, 18 (85.7%) were male and most of them were 37–48 months old (52.4%), followed by 25–36 months old (42.8%) and 12–24 months old (4.8%). Children who had been diagnosed with ASD comprised 85.7% (Table 2). The inter-rater reliability results show that several items had at least good agreement between the evaluators (Kappa ≥0.6): items 1, 4, 8, 10, 12, 14, 15, 16, and 17 in the observation section and items 2, 3, 4, 9, 11, 15, 16, 18, 19, 20, 21, 22, and 23 in the interview section. In addition, there were some items with poor or fair levels of reliability (Kappa <0.6) in the observation section (items 3, 6, 7, and 22) and the interview section (items 1, 5, 13, and 17) (Table 4). However, the results show that the diagnosis of ASD using TDAS between the evaluators was exceptionally accurate overall (Kappa = 1.000) (Table 5).

The concurrent validity of TDAS

Forty-seven of 62 children (75.6%) who participated in trial 3 were male, and most of them were 37–48 months old (51.6%), followed by 25–36 months old (30.7%), and 12–24 months old (17.7%). The children who were diagnosed with ASD comprised 72.4% (Table 2). The concurrent validity between the TDAS assessments and the physicians' diagnoses based on the DSM-5 criteria was in excellent agreement (Kappa = 0.871). Only one false-negative case was found and there were no false positives. The sensitivity and specificity of TDAS were 100% and 82.4%, respectively (Table 6).

DISCUSSION

We developed a diagnostic tool, TDAS, for the early diagnosis of ASD in Thai children (aged 12–48 months old). It is designed to be used by multidisciplinary teams comprising health professionals and nurses but can also serve as an additional ASD diagnostic instrument for physicians. This tool can be used to comprehensively identify the symptoms related to ASD that are only briefly covered in DSM-5 and presents clearer criteria for their interpretation within the Thai cultural context. This will lead to less reliance on expertise and/or clinical experience for ASD diagnosis and is thus more widely applicable to the Thai healthcare setting. TDAS showed good content validity overall (IOC: 0.71–1.00) and suitable construct validity (RMSEA = 0.076, CFI = 0.902, and TLI = 0.882 for the observation section and RMSEA = 0.067, CFI = 0.858, and TLI = 0.837 for the interview section) Meanwhile, diagnoses using TDAS between evaluators were exceptionally accurate (Kappa = 1.000), while diagnoses of ASD using TDAS and by physicians were in excellent agreement (Kappa = 0.871). The sensitivity and specificity of TDAS were 100% and 82.4%, respectively.

Inter-rater reliability for several items achieved good agreement between the evaluators (9 of 13 items in observation and 13 of 17 items in the interview section). This is consistent with a study on the screening of ASD using DISCO by interviewing parents of children aged 3–11 years old in which a high level of inter-rater reliability (Kappa coefficients at 0.75 or higher) was reported (Wing et al., 2002). Differences in the areas of expertise of the evaluators might influence their evaluation and scoring, and so the evaluators involved in this study came from a variety of disciplines (nurses, psychologists, and speech therapists). For example, speech therapists might have a higher sensitivity toward language problems than nurses. Therefore, the continuation of training and administering TDAS might reduce any discrepancies among diagnoses by health professionals from different backgrounds.

The number of items with poor agreement was higher in the observation section, which might have resulted

TABLE 3 Construct validity of TDAS ($N = 170$)

Item	Factor loading	
	Observation section	**Interview section**
A1: Deficits in social–emotional reciprocity		
1. Getting help	0.683	0.617
2. Social initiation	-	0.487
3. Sharing interests	0.623	0.692
4. Social response—reciprocal conversation	0.667	0.512
5. Sharing of emotions	-	0.560
A2: Deficits in communicative behaviors used for social interaction		
6. Eye contact	0.782	-
7. Non-verbal communication	0.707	-
8. Verbal communication	0.502	-
9. Understanding emotions	-	0.632
A3: Deficits in developing, maintaining, and understanding relationships		
10. Developing, maintaining, and understanding relationships	0.571	-
11. Difficulty in adjusting behavior to suit the social context	-	0.626
12. Sharing imaginative play	0.609	-
13. Making friends		0.705
14. Absence or lack of interest in others	0.641	-
B1: Stereotyped or repetitive motor movements, use of objects, and/or speech		
15. Unusual movements of arms, hands, fingers and/or the whole body	0.497	0.594
16. Repetitive manipulation of objects	0.282	0.437
17. Idiosyncratic use of speech	0.330	0.313
B2: Sameness, inflexible adherence to routines, and/or ritualized patterns		
18. Difficulty in changing or choosing activities and adherence to routines	-	0.416
19. Rigid thinking	-	0.234
B3: Highly restricted fixed interests that are abnormal in intensity or focus		
20. Interests that are abnormal in intensity, focused on the same topics, objects, or activities; displaying a narrow range of interests; and/or collecting or clinging to objects	-	0.205
21. Unusual fears	-	0.257
B4: Hyper- or hypo-reactivity to sensory input or unusual interest in sensory aspects of the environment		
22. Unusual response to injury and/or violence	0.393	0.445
23. Unusual response to sensory stimuli and/or exploration	-	0.495
RMSEA	0.076	0.067
CFI	0.902	0.858
TLI	0.882	0.837

Abbreviations: CFI, comparative fit index; RMSEA, root-mean-squared error of approximation; TDAS, the Thai Diagnostic Autism Scale; TLI, Tucker-Lewis index.

from the difficulty in observing some expressions by the children during the short period of observation or in different environments. For example, in item 3, which had a fair level of agreement, the action of a child to attract his/her parents' attention in a different place with several people might be less than normal compared to his/her

TABLE 4 Inter-rater reliability of items in TDAS ($N = 21$)

Item	Kappa coefficients	
	Observation section	**Interview section**
A1: Deficits in social–emotional reciprocity		
1. Getting help	0.696	0.382
2. Social initiation	-	0.613
3. Sharing of interests	0.577	0.753
4. Social response—reciprocal conversation	0.718	1.000
5. Sharing emotions	-	0.588
A2: Deficits in communicative behaviors used for social interaction		
6. Eye contact	0.577	-
7. Non-verbal communication	0.539	-
8. Verbal communication	0.877	-
9. Understanding emotions	-	0.622
A3: Deficits in developing, maintaining, and understanding relationships		
10. Developing, maintaining, and understanding relationships	0.859	-
11. Difficulties in adjusting behavior to suit the social context	-	0.897
12. Sharing imaginative play	0.889	-
13. Making friends		0.462
14. Absence or lack of interest in others	0.691	-
B1: Stereotyped or repetitive motor movements, use of objects, and/or speech		
15. Unusual movements of arms, hands, fingers, and/or the whole body	0.667	0.712
16. Repetitive manipulation of objects	0.632	0.858
17. Idiosyncratic use of speech	1.000	0.417
B2: Sameness, inflexible adherence to routines, and/or ritualized patterns		
18. Difficulty in changing or choosing activities and adherence to routines	-	0.772
19. Rigid thinking	-	0.879
B3: Highly restricted fixed interests that are abnormal in intensity or focus		
20. Interests that are abnormal in intensity, focused on the same topics, objects, or activities; displaying a narrow range of interests; and/or collecting or clinging to objects	-	0.634
21. Unusual fears	-	0.708
B4: Hyper- or hypo-reactivity to sensory input or unusual interest in sensory aspects of the environment		
22. Unusual response to injury or violence	0.507	0.892
23. Unusual response to sensory stimuli and exploration	-	0.922

Abbreviation: TDAS, the Thai Diagnostic Autism Scale.

daily life. The importance of assessing information from the parents toward the diagnosis has been suggested in several previous studies (Kim & Lord, 2010; Schutte & Hewitson, 2014). The parents are more likely to notice abnormalities in behavior after spending a lot of time with their children, which is reflected by the high level of agreement for several items in the interview section in the present study. Discrepancies between observations and parents' perspectives have been reported as a source of inconsistency in inter-rater reliability analysis between the observation and interview sections. Our results also show that some items with fair agreement in one section achieved stronger agreement in the other (such as items 3, 17, and 22).

Although there were some items with a fair level of agreement, our results show excellent consistency of ASD diagnosis between the evaluators, suggesting that any disagreement between them on the severity of abnormalities

TABLE 5 Inter-rater reliability of ASD diagnosis using TDAS ($N = 21$)

	Second evaluator			
	Non-ASD	ASD	Total	Kappa coefficient
First evaluator				
Non-ASD	3	0	3	1.000
ASD	0	18	18	
Total	3	18	21	

Abbreviations: ASD, Autism Spectrum Disorder; TDAS, the Thai Diagnostic Autism Scale.

TABLE 6 Concurrent validity between TDAS and physicians for ASD diagnosis ($N = 62$)

	Physicians' diagnosis			
	Non-ASD	ASD	Total	Indices
TDAS assessment				Kappa = 0.871
Non-ASD	14	0	14	Sensitivity = 100%
ASD	3	45	48	Specificity = 82.4%
Total	17	45	62	

Abbreviations: ASD, Autism Spectrum Disorder; TDAS, the Thai Diagnostic Autism Scale.

did not affect the overall diagnosis of ASD. However, a consistent perspective on diagnosis might also reduce evaluation bias and increase the accuracy of the severity classification. Guidelines of scoring should be explained with a clear statement to mitigate this problem.

Cultural differences might be influential on the differences between evaluation methods. It has been suggested that issues with translation, interpretation, and/or cultural understanding of behavior might contribute to false-positive results (Varma & Iskandar, 2014). For instance, like in most other Far Eastern countries, Thai children are not expected to make eye contact nor initiate social interaction with adults. Thus, from the parent's perspective, this issue might be a temporary problem that will improve when the child has grown up. Therefore, the parents might miss and/or not report this behavior.

In a previous study to compare the sensitivity and specificity between the M-CHAT and Parent's Observations of Social Interactions (POSI) among children aged 16–48 months old (Salisbury et al., 2018), the authors found that POSI had a higher sensitivity than M-CHAT among children aged 16–30 months (93.6% versus 77.5%; $p < 0.001$). In addition, both tools seemed to have lower sensitivity when used to screen older children. In the present study, we did not evaluate the effect of age on sensitivity and specificity due to the limited number of participants, and so further study factoring in age or other characteristics might be advantageous for evaluating the psychometric properties of TDAS. In addition, since the diagnosis of ASD using TDAS was a single score in at least five domains, the lowest possible score is 5 points, and one of the advisory expert panelists suggested that identification of ASD with a very low score might lead to a false-positive result. Therefore, an appropriate cut-off point for the diagnostic criteria should be examined in future studies.

Our initial literature review included studies from several non-English speaking countries regarding the process they undertook to develop a culturally sensitive diagnostic tool for Autism in their native language, such as the Behavior Development Screening for Toddlers (BeDevel) (Lee et al., 2020). However, the tools reviewed did not match some of the criteria we wanted to include, such as age ranges.

There were some limitations of this study. According to the Developmental Surveillance Project of Thailand, children should be screened for developmental delay and disability during regular well-baby clinic visits at 9, 18, 30, and 42 months old, and indeed, the participants in the present study were referred from the Developmental Surveillance Project. Thus, the age range of the participants was not evenly distributed. Outreach recruitment of participants in healthcare services in addition to depending solely on referrals from the Developmental Surveillance Project might increase the sample size and more evenly distribute the target population by age (1–4 years). Another limitation is that the samples were only recruited from the Chiang Mai center, and so our findings might not be generalizable for the incidence and severity of ASD nationwide. Thus, a follow-up study to examine the efficacy of this tool from a national perspective should be conducted. In addition, some potential confounding variables such as parental socio-economic status were not included in this study, which could have resulted in the differences in severity diagnosis between the evaluators, and so a further study involving confounding by these factors should be considered. In addition, the term "getting help" used in the TDAS tool was included in the examiner's dialogue with the child to respond to back-and-forth interaction when seeking assistance; revision of this misleading phrase might be useful in future studies.

In summary, TDAS for the early diagnosis of ASD in Thai children yielded a high level of content validity, concurrent validity, and inter-rater reliability. This tool comprehensively includes items related to DSM-5 criteria for both observational and interviewing aspects. Further study to examine the efficacy of TDAS nationwide and to establish an appropriate cut-off point should be conducted. A large-scale study using TDAS is needed to determine an appropriate cut-off point, its efficacy, and intervention related to its interpretation.

ACKNOWLEDGMENTS

We are very grateful to all of the participants and their families who were involved in this study. We would also like to thank the multidisciplinary team at

Rajanagarindra Institute of Child Development for their support of this study. This research work was supported by the Health Systems Research Institute, the Department of Mental Health, and partially supported by Chiang Mai University.

CONFLICT OF INTEREST

The authors declare that there are no competing interests to disclose.

ETHICS STATEMENT

Ethical approval for this study was obtained from the RICD ethics committee (number 14/2015 and 1/2017).

ORCID

Duangkamol Tangviriyapaiboon https://orcid.org/0000-0002-0221-8666

REFERENCES

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*. American Psychiatric Publishing.

Bastiaansen, J. A., Meffert, H., Hein, S., Huizinga, P., Ketelaars, C., Pijnenborg, M., Bartels, A., Minderaa, R., Keysers, C., & de Bildt, A. (2011). Diagnosing autism spectrum disorders in adults: The use of autism diagnostic observation schedule (ADOS) module 4. *Journal of Autism and Developmental Disorders*, *41*, 1256–1266. https://doi.org/10.1007/s10803-010-1157-x

Ben-Itzchak, E., & Zachor, D. A. (2007). The effects of intellectual functioning and autism severity on outcome of early behavioral intervention for children with autism. *Research in Developmental Disabilities*, *28*, 287–303. https://doi.org/10.1016/j.ridd.2006.03.002

Carr, T. (2013). Autism diagnostic observation schedule. In F. R. Volkmar (Ed.), *Encyclopedia of autism Spectrum disorders* (pp. 349–356). Springer, New York.

Cattell, R. B. (1978). *The scientific use of factor analysis in behavioral and life sciences*. Plenum.

Choueiri, R., & Wagner, S. (2015). A new interactive screening test for autism spectrum disorders in toddlers. *The Journal of Pediatrics*, *167*, 460–466. https://doi.org/10.1016/j.jpeds.2015.05.029

Cicchetti, D. V., & Sparrow, S. A. (1981). Developing criteria for establishing interrater reliability of specific items: Applications to assessment of adaptive behavior. *American Journal of Mental Deficiency*, *86*, 127–137.

Daniel, W. W. (1999). *Biostatistics: A foundation for analysis in the health sciences* (7th ed.). John Wiley & Sons.

Department of Mental Health, Ministry of Public Health. (2014). *Khaosan Sukkhaphapchit* (Vol. 21). WVO Officer of Printing Mill, War Veterans Organization of Thailand.

Department of Mental Health, Ministry of Public Health (2015). In S. Sirithongthaworn, A. Benjaponpithak, & N. Sriwongpanich (Eds.), *Thai early developmental assessment for intervention: TEDA4I*. The Agricultural Cooperative Federation of Thailand., Ltd.

Donner, A., & Rotondi, M. A. (2010). Sample size requirements for interval estimation of the kappa statistic for interobserver agreement studies with a binary outcome and multiple raters. *The International Journal of Biostatistics*, *6*, 1–11. https://doi.org/10.2202/1557-4679.1275.

Everitt, B. S. (1975). Multivariate analysis: The needfor data, and other problems. *The British Journal of Psychiatry: the Journal of Mental Science*, *126*, 237–240.

Gorsuch, R. L. (1983). *Factor analysis* (2nd ed.). Lawrence Erlbaum Associates.

Huerta, M., Bishop, S. L., Duncan, A., Hus, V., & Lord, C. (2012). Application of DSM-5 criteria for autism spectrum disorder to three samples of children with DSM-IV diagnoses of pervasive developmental disorders. *The American Journal of Psychiatry*, *169*, 1056–1064. https://doi.org/10.1176/appi.ajp.2012.12020276

Huerta, M., & Lord, C. (2012). Diagnostic evaluation of autism spectrum disorders. *Pediatric Clinics of North America*, *59*, 103–xi. https://doi.org/10.1016/j.pcl.2011.10.018

Kaat, A., & Lecavalier, L. (2013). Childhood autism rating scale. In F. R. Volkmar (Ed.), *Encyclopedia of autism Spectrum disorders* (pp. 590–593). Springer New York.

Kim, H., Ku, B., Kim, J., Park, Y., & Park, Y. (2016). Confirmatory and exploratory factor analysis for validating the phlegm pattern questionnaire for healthy subjects. *Evidence-based Complementary and Alternative Medicine*, *2016*, 2696019. https://doi.org/10.1155/2016/2696019

Kim, S., & Lord, C. (2010). Restricted and repetitive behaviors in toddlers and preschoolers with autism spectrum disorders based on the autism diagnostic observation schedule (ADOS). *Autism Research*, *3*, 162–173. https://doi.org/10.1002/aur.142

Kim, S. H., Hus, V., & Lord, C. (2013). Autism diagnostic interview-revised. In F. R. Volkmar (Ed.), *Encyclopedia of autism Spectrum disorders* (pp. 345–349). Springer, New York.

Lee, K. S., Jung, S. J., Yoo, H. J., Shin, Y. W., & Cho, Y. I. (2020). Validity and reliability of the behavior development screening for toddlers-questionnaire/parents(BeDevel-Q/P): A Korean autism screening instrument for infants aged 24-35 months. *Psychiatry Investigation*, *17*, 47–54. https://doi.org/10.30773/pi.2018.0271

Leekam, S. (2013). Diagnostic interview for social and communication disorders. In F. R. Volkmar (Ed.), *Encyclopedia of autism spectrum disorders* (pp. 926–931). Springer, New York.

Leekam, S., Libby, S., Wing, L., Gould, J., & Taylor, C. (2002). The diagnostic interview for social and communication disorders: Algorithms for ICD-10 childhood autism and Wing and Gould autistic spectrum disorder. *Journal of Child Psychology and Psychiatry*, *43*, 327–342. https://doi.org/10.1111/1469-7610.00024

Lord, C., Rutter, M., DiLavore, P. C., Risi, S., Gotham, K., & Bishop, S. (2012). *Autism diagnostic observation schedule, second edition (ADOS-2) manual*. Western Psychological Services.

Ministry of Public Health. (2020). *Developmental assessment for intervention manual (DAIM)*. WVO Officer of Printing Mill, War Veterans Organization of Thailand.

Parikh, R., Mathai, A., Parikh, S., Chandra Sekhar, G., & Thomas, R. (2008). Understanding and using sensitivity, specificity and predictive values. *Indian Journal of Ophthalmology*, *56*, 45–50. https://doi.org/10.4103/0301-4738.37595

Plubrukarn, R., Piyasil, V., Moungnoi, P., Tanprasert, S., & Chutchawalitsaku, I. V. (2005). Trend study of autistic spectrum disorders at queen Sirikit National Institute of child health. *Journal of the Medical Association of Thailand*, *88*, 891–897.

Poolsuppasit, S., Panyayong, B., Liknapichitkul, D., Serisathien, P., & Chutha, W. (2005). Holistic care for Thai autism. *Journal of Mental Health Care in Thailand*, *13*, 1–7.

Pornnoppadol, C., Thongngen, A., Gaevalin, A., & Sangratanayont, D. (2002). Development of the pervasive developmental disorders screening questionnaires. *Journal of the Psychiatric Association of Thailand*, *47*, 75–96.

Richler, J., Bishop, S. L., Kleinke, J. R., & Lord, C. (2007). Restricted and repetitive behaviors in young children with autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *37*, 73–85. https://doi.org/10.1007/s10803-006-0332-6

Robins, D. L., Casagrande, K., Barton, M., Chen, C. M., 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. https://doi.org/10.1542/peds.2013-1813

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. https://doi.org/10.1023/a:1010738829569

Robinson, J. (2013). Gilliam autism rating scale (GARS). In F. R. Volkmar (Ed.), *Encyclopedia of autism spectrum disorders* (pp. 1441–1444). Springer, New York.

Rudra, A., Banerjee, S., Singhal, N., Barua, M., Mukerji, S., & Chakrabarti, B. (2014). Translation and usability of autism screening and diagnostic tools for autism spectrum conditions in India. *Autism Research*, *7*, 598–607. https://doi.org/10.1002/aur.1404

Rutter, M., Le Couteur, A., & Lord, C. (2003). *Autism diagnostic interview-revised*. Western Psychological Services.

Salisbury, L. A., Nyce, J. D., Hannum, C. D., Sheldrick, R. C., & Perrin, E. C. (2018). Sensitivity and specificity of 2 autism screeners among referred children between 16 and 48 months of age. *Journal of Developmental and Behavioral Pediatrics*, *39*, 254–258. https://doi.org/10.1097/dbp.0000000000000537

Schopler, E., Reichler, R., & Renner, B. (1988). *The childhood autism rating scale (CARS)*. Western Psychological Services.

Schutte, C., & Hewitson, L. (2014). Relationship between stereotyped behaviors and restricted interests (SBRIs) measured on the autism diagnostic observation schedule (ADOS) and diagnostic results. *OA Autism*, *2*, 15–20.

Shan, G. (2016). Sample size calculation for agreement between two raters with binary endpoints using exact tests. *Statistical Methods in Medical Research*, *27*, 2132–2141. https://doi.org/10.1177/0962280216676854

Sigman, M., & McGovern, C. W. (2005). Improvement in cognitive and language skills from preschool to adolescence in autism. *Journal of Autism and Developmental Disorders*, *35*, 15–23. https://doi.org/10.1007/s10803-004-1027-5

Sirithongthaworn, S. (2018). The development of developmental surveillance and promotion manual; DSPM. *Journal of the Psychiatric Association of Thailand*, *63*, 3–12.

Slappendel, G., Mandy, W., van der Ende, J., Verhulst, F. C., van der Sijde, A., Duvekot, J., Skuse, D., & Greaves-Lord, K. (2016). Utility of the 3Di short version for the diagnostic assessment of autism spectrum disorder and compatibility with DSM-5. *Journal of Autism and Developmental Disorders*, *46*, 1834–1846. https://doi.org/10.1007/s10803-016-2713-9

Stone, W. L., McMahon, C. R., & Henderson, L. M. (2008). Use of the screening tool for autism in two-year-olds (STAT) for children under 24 months: An exploratory study. *Autism*, *12*, 557–573. https://doi.org/10.1177/1362361308096403

Surinkaew, D., Louthrenoo, O., Charnsil, C., & Witoonchart, C. (2005). Prevalence of pervasive developmental disorders in preschool children in Chiang Mai. *Chiang Mai Medical Journal*, *44*, 29–34.

Trevethan, R. (2017). Sensitivity, specificity, and predictive values: Foundations, pliabilities, and pitfalls in research and practice. *Frontiers in Public Health*, *5*, 307. https://doi.org/10.3389/fpubh.2017.00307

Varma, A., & Iskandar, J. W. (2014). Challenges in diagnosis of autism and the struggle of using western screening tools in different cultures. Psychiatrist's perspective. *Indian Pediatrics*, *51*, 356–357.

Wangkawan, T., Lai, C., Munkhetvit, P., Yung, T., & Chinchai, S. (2020). The development and psychometric properties of the visuospatial working memory assessment (VWMA) for children. *Occupational Therapy International*, *2020*, 8736308. https://doi.org/10.1155/2020/8736308

Wetherby, A. M., Brosnan-Maddox, S., Peace, V., & Newton, L. (2008). Validation of the infant-toddler checklist as a broadband screener for autism spectrum disorders from 9 to 24 months of age. *Autism*, *12*, 487–511. https://doi.org/10.1177/1362361308094501

Wing, L., Leekam, S. R., Libby, S. J., Gould, J., & Larcombe, M. (2002). The diagnostic interview for social and communication disorders: Background, inter-rater reliability and clinical use. *Journal of Child Psychology and Psychiatry*, *43*, 307–325. https://doi.org/10.1111/1469-7610.00023

World Health Organization. (1993). *The ICD-10 classification of mental and behavioural disorders: Diagnostic criteria for research*. World Health Organization.

Zhao, Y., Chan, W., & Lo, B. C. (2017). Comparing five depression measures in depressed Chinese patients using item response theory: An examination of item properties, measurement precision and score comparability. *Health and Quality of Life Outcomes*, *15*, 60. https://doi.org/10.1186/s12955-017-0631-y

How to cite this article: Tangviriyapaiboon, D., Sirithongthaworn, S., Apikomonkon, H., Suyakong, C., Srikummoon, P., Kawilapat, S., & Traisathit, P. (2022). Development and psychometric evaluation of a Thai Diagnostic Autism Scale for the early diagnosis of Autism Spectrum Disorder. *Autism Research*, *15*(2), 317–327. https://doi.org/10.1002/aur.2631