J Autism Dev Disord (2013) 43:1028–1037
DOI 10.1007/s10803-012-1643-4

ORIGINAL PAPER

Development of T-STAT for Early Autism Screening

Chung-Hsin Chiang · Chin-Chin Wu · Yuh-Ming Hou · Ching-Lin Chu · Jiun-Horng Liu · Wei-Tsuen Soong

Published online: 14 September 2012

Abstract This study's purpose was to modify the Screening Tool for Autism in Two-Year-Olds (STAT) into a Taiwanese version called T-STAT. Study 1 included 15 children with Autism and 15 children with Developmental Delay (DD) or language impairment (LI) aged between 24 and 35 months. Study 2 had 77 young children with Autism, PDD-NOS, or DD/LI as a clinical-based validation sample. In Study 1, the signal detection procedure found that a cutoff score of 2 would yield high sensitivity and specificity in T-STAT. In Study 2, using a score of 2 as a cutoff, the agreement between T-STAT risk and ADOS classification was highly acceptable. Results were promising as a Level 2 screening tool for Autism for ages two to three.

Keywords Screening · Young children with Autism · Taiwan

C.-H. Chiang (✉)
Department of Psychology, National Chengchi University, No.64, Sec.2, ZhiNan Rd., Wenshan District, Taipei 11605, Taiwan
e-mail: chchiang@nccu.edu.tw

C.-H. Chiang
Research Center for Mind, Brain, and Learning, National Chengchi University, No.64, Sec.2, ZhiNan Rd., Wenshan District, Taipei 11605, Taiwan

C.-C. Wu
Department of Psychology, Kaohsiung Medical University, Kaohsiung, Taiwan

Y.-M. Hou · J.-H. Liu
Department of Psychiatry, Chia-Yi Christian Hospital, Chia-Yi, Taiwan

C.-L. Chu
Department of Psychology, National Chung Cheng University, Chai-Yi, Taiwan

J.-H. Liu
Department of Psychiatry, Chi-Mei Medical Center, Liouying, Tainan, Taiwan

W.-T. Soong
Department of Psychiatry, St. Joseph's Hospital, Yunlin, Taiwan

W.-T. Soong
Department of Psychiatry, National Taiwan University and University Hospital, Taipei, Taiwan

Introduction

Autistic disorder (Autism) or Autism Spectrum Disorders (ASD) (including Autism, Asperger Syndrome, Pervasive Developmental Disorder Not Otherwise Specified) are considered as neurodevelopmental impairments in social interaction, communication, and repetitive or restricted patterns of interests or behaviors (Bailey et al. 1996; Gotham et al. 2011). In the past, Autism was considered to be evident at birth (Kanner 1943); however, recent prospective research has demonstrated that the deficiency might not be revealed until 6 months old (Ozonoff et al. 2010; for a review, Rogers 2009). Although symptomatology of ASD can be observed at much younger ages (APA 2000), the current age of clinical diagnosis for ASD remains at approximately three years or older (Barbaro and Dissanayake 2009). Therefore, developing a suitable tool based on behavioral domain for early identification is important work in the clinical settings.

Early identification of ASD signs can provide early intervention to facilitate or even maximize the development of affected children. Development of a screening tool is the first step for early identification of ASD. *Screening* or *screening assessment* is defined as "a relatively brief evaluation intended to identify children who are at risk for developing certain

disorders or disabilities, who are eligible for certain programs, or who have a disorder or disability in need of remediation, or who need a more comprehensive assessment" (Sattler 2008).

There are two levels of screening tools for young children with ASD (Filipek et al. 1999). Level 1 ASD screens identify children at risk for ASD in the settings of regular pediatric clinics, or community health services; Level 2 ASD screens specifically focus on differentiating children at risk for ASD from other Developmental Delays, such as language delays or general developmental difficulties in a specialized clinical setting. Existing Level 1 screening tools for ASD or Autism in children below two years old include the Checklist for Autism in Toddlers (CHAT) (Baird et al. 2000; Baron-Cohen et al. 1992, 1996), the Modified Checklist for Autism in Toddlers (M-CHAT) (Robins et al. 2001; Kleinman et al. 2008), Pervasive Developmental Disorders Screening Test-Stage 1 (PDDST-Stage 1) (Siegel 1996), Infant-Toddler Checklist from Communication and Symbolic Behavior Scales: Developmental Profile (ITC CSBS DP) (Wetherby and Prizant 2002), and Early Screening of Autistic Traits Questionnaire (ESAT) (Swinkels et al. 2006). Each tool has its strengths and weaknesses, which were discussed in recent review articles (Barton et al. in press; Martínez-Pedraza and Carter 2009).

For Level 2 screening of young children with Autism under three years of age, at least four tools have been developed: the Childhood Autism Rating Scale (CARS) (Schopler et al. 1988), PDDST-Stage 2 (Siegel 1996; Siegel and Hayer 1999), CSBS DP-behavioral sample (Wetherby et al. 2004), and Screening Tool for Autism in Two-Year-Olds (STAT; Stone et al. 2000, 2004, 2008). Since the former two tools lacked consistently validated findings for Autism in children under three years of age, the latter two instruments will be discussed in detail as they have been repeatedly validated and their findings have been more consistent. The original design of the CSBS DP-behavioral sample was developed as a Level 2 screener for Developmental Delay; however, Wetherby et al. (2004) studied the 29 items of CSBS DP that related the symptoms of ASD called Systematic Observation of Red Flags (SORF) for identifying the developmentally delayed children who were later diagnosed as having ASD. The study of SORF indicated that the following nine red flags can differentiate children with ASD from those children with Developmental Delay and typical development: (1) lack of appropriate gaze; (2) lack of warm, joyful expressions with gaze; (3) lack of sharing enjoyment or interest; (4) lack of response to name; (5) lack of coordination of gaze, facial expression, gesture, and sound; (6) lack of showing; (7) unusual prosody; (8) repetitive movement or posturing of body, arms, hands, or fingers; and (9) repetitive movements with objects. In addition, four red flags can differentiate children with ASD from children with typical development but not Developmental Delay: (1) lack of response to contextual cues; (2) lack of pointing; (3) lack of vocalization with consonants; and (4) lack of conventional playing with a variety of toys. These findings suggest that failing these 13 "red flag" items requires referral to comprehensive evaluation for ASD.

The STAT is a 20-min play-based interactive tool designed for use with high-risk children from 24 to 35 months of age. It consists of 12 items, including four social communicative domains: play, requesting, directing attention, and motor imitation. The 12 items were selected from Stone and colleagues' former studies focusing on the differentiation of young children with Autism and Developmental Delay (Stone et al. 1997a, b). Stone et al. (2000) used seven children with Autism and 33 children with Developmental Delay as a developmental sample; the study defined failing on two domains as a "high-risk for Autism", and yielded sensitivity and specificity as 1.00 and 0.91, respectively. Using the above criterion (failing on 2 domains) in a validation sample, the sensitivity and specificity are 0.83 and 0.86. In another study, Stone et al. (2004) used 13 chronological-age and mental-age matched children with Autism and Developmental Delay as a new developmental sample. By using a score of 2 as a cutoff score for "Autism risk" in signal detection theory, they again obtained high sensitivity (0.92) and specificity (0.85) in the validation group. The study also demonstrated psychometric properties including high inter-rater reliability and concurrent validity in a fairly representative sample referred for developmental evaluation. Recently, Stone and colleagues (Stone, et al. 2008) continued to explore the utility of STAT in young children under 24 months and found acceptable discriminative value while using a score of 2.75 as a cutoff in children 14 months and older.

In Taiwan, only the Clancy Behavior Scale was translated and used in clinical settings since the 1980s (Hsieh et al. 1983). However, the scale was a parent-report questionnaire used in Level 2 screening, and the items need to be rewritten to fit the current knowledge of children in early development of ASD. Until now, most professionals in Taiwan use CHAT or M-CHAT as Level 1 screening in their clinical settings. There has been no instrument developed for Level 2 screening in Taiwan. Between the two well-established Level 2 tools—CSBS DP and STAT—the latter was chosen for developing a new tool to be used in Taiwan because it is a more user-friendly tool for administering and scoring. Therefore, the present study used STAT as the base to establish a new version of STAT called the Taiwan version-STAT (T-STAT) in Chinese. The specific aims were: (1) to develop a scoring algorithm for the T-STAT using the signal detection theory in the development sample; and (2) to examine the concurrent validity of T-STAT in the validation sample.

Methods: Study 1—Development of T-STAT Scoring Algorithm

Participants

Participants in Study 1 were 30 children, 15 with a clinical diagnosis of Autism and 15 with Developmental Delay and/or language impairment (DD/LI) as a development sample. All children were recruited for participation between 2004 and 2005 from a local hospital in the southwest area of Taiwan. The hospital hosts an Interdisciplinary Assessment Center for Children with Suspected Developmental Delay (IACCSDD) to provide services for the children suspected of having Developmental Delay and their parents. The children were recruited by two child psychiatrists in the research team. These children's parents provided the informed consent. Appropriate IRB approvals were obtained prior to conducting the study.

Eligibility requirements for participation included: (1) chronological age from 24 to 35 months; (2) absence of an identified genetic or metabolic disorder; and (3) absence of a severe sensory or motor impairment.

Children in the sample were individually matched by chronological age (CA), mental age (MA), verbal mental age (VMA), nonverbal mental age (NVMA), gender, and the parents' social economic status (SES). The Mullen Scales of Early Learning (MSEL) (Mullen 1995) was used for measuring mental function/development. Demographic characteristics for the developmental sample are presented in Table 1. There was no significant group difference found for CA, MA, VMA, NVMA, and gender.

Table 1 Demographic characteristics of developmental sample

	Autism (n = 15)	DD (n = 15)	t	*p*
Chronological age (months)				
M (SD)	29.14 (3.87)	27.67 (3.46)	1.10	0.28
Range	24.0–35.0	25.0–35.0		
NVMA (months)				
M (SD)	20.77 (4.07)	20.27 (1.80)	0.44	0.67
Range	12.0–28.0	17.5–23.5		
VMA (months)				
M (SD)	13.23 (6.23)	15.13 (2.48)	−1.10	0.28
Range	5.0–31.5	8.5–18.0		
Mental age (months)				
M (SD)	17.00 (4.56)	17.70 (1.97)	−0.59	0.56
Range	8.50–29.75	13.00–20.25		
SES				
M (SD)	58.80 (17.91)	49.47 (18.61)	−0.55	0.59
Range	35–91	35–98		
Sex				
Male:female	10:05	10:05		

Measures and Procedures

STAT is an interactive measure administered individually for about 20 min within a playful context (Stone et al. 2000, 2004). It consists of 12 items that assess behaviors in four social-communicative domains: Play, Requesting, Directing Attention, and Motor Imitation. For developing the T-STAT, one item was changed on each of two domains of STAT: Directing Attention and Motor Imitation. In the Directing Attention domain, all four items are designed to measure initiating joint attention, but there was no item for responding to joint attention. In the pilot study, the authors found that the children suspected of having a Developmental Delay or being at risk for Autism, as well as the typically developed toddlers showed very little interest or did not respond to the item of "Puppet". In reviewing former studies in joint attention for young children with Autism (Chiang et al. 2008; Mundy and Burrette 2005), we replaced the "Puppet" item with a new item called "Posters". We put two posters (56 × 65 cm) onto two walls (near 200 cm high) in the testing room. The examiner called the child's name and used a short-arm point to one of the posters three times. If the child could follow the direction of the pointing and looked at the poster within 3 s, it would be scored as a success. Two trials were arranged, and if one trial was a "success," it would be scored as "pass." This new item measured responding to joint attention; therefore, we used the new title "Joint Attention" instead of "Directing Attention" for the domain. Regarding Motor Imitation in STAT, we used "Fist" instead of "Rattle" for two reasons: First, in the pilot study, we found that most typically developed toddlers or young children with Developmental Delay shake the rattle immediately even before the examiner assigned them the task. The phenomenon was related to affordance learning (Whiten 2006). Second, the literature (Rogers et al. 2003) showed that young children with Autism are more impaired on body imitation than object imitation. Therefore, we created the new item "Fist": The examiner made a fist first and then opened/closed the fist three times in front of the child. Three trials were arranged; it would be scored as a "pass" if the child could imitate a fist opening and closing in at least one out of three trials (please contact the corresponding author to obtain the T-STAT manual in Chinese). Table 2 and Fig. 1 showed the items and described how to administer the test.

All item scores and weightings are the same as described in STAT; thus, scores for the domains of Play and Requesting (i.e., two items) can be 0, 0.5, and 1. Scores for the domains of Joint Attention and Imitation (i.e., four

Table 2 T-STAT items and brief description of administration

Domain	Item	Description
Play	Turn-taking	Examiner rolls a ball or toy car to the child and interact in back-and-forth play
	Doll play	Examiner presents the child with a doll or stuffed animal, along with furniture and eating utensils, and observes child's functional play
Requesting	Snack	Examiner presents the child with a clear, tightly sealed jar filled with favorable food treats
	Bubbles	Examiner blows soap bubbles and then hands the tightly sealed far to the child
Joint attention	Balloon	Examiner inflates a balloon and then lets it go to observe child's initiating joint attention behavior while flying in the room as it deflates
	Posters[a]	Examiner uses short-arm point to the posters hung on the upper walls of the room
	Bag of toys	Examiner presents an opaque bag containing interesting toys to the child and encourages looking inside
	Noisemaker	Examiner activates a noisemaker out of view of the child
Imitation	Car	Examiner rolls a small car back and forth on the table and then encourages the child to do the same
	Drum hands	Examiner drums child's hands on the table and then encourages the child to do the same
	Hop elephant[b]	Examiner hops a small toy elephant on the table and then encourages the child to do the same
	Fist[a]	Examiner raise hand horizontally and opens/close three times and encourages the child to do the same

[a] Two items which are different from STAT

[b] Original STAT used dog instead of elephant

Fig. 1 Materials used in T-STAT. **a** Play; **b** requesting; **c** joint attention (posters, actual size is 56.3 × 65.7 cm); **d** imitation

items) can be 0, 0.25, 0.5, 0.75, and 1. The total T-STAT score ranges from 0 to 4. The higher scores represent greater impairment.

To obtain the diagnosis data blindly, the research team worked in two places, the local hospital and the local university. After, being referred by IACCSDD, the parents

generally attended the first visit to the university lab for psychological assessment carried out by the child psychologists; the parents then visited a child psychiatrist within 1 month for a clinical evaluation based on DSM-IV and the Autism Diagnostic Observation Schedule (ADOS; Lord et al. 1999). Because there might be some inconsistency in clinical judgment between clinicians for the children with ASD under three years (Stone et al. 1999), five children suspected of having ASD and their parents were invited to assist the development of a standardized clinical procedure and the establishment of the inter-rater reliability of the research team. After achieving acceptable interrater reliability (0.90) between the two child psychiatrists, the formal study was started. The assessment at the university lab included using MSEL, T-STAT, and other social communication scales and ADOS sequentially in a session after the parents' informed consent was obtained. The first author had received prior STAT training in 2003 on administering and scoring from Dr. Wendy Stone's group; and in 2004, had received research training on ADOS with Dr. Cathy Lord's group. The T-STAT was administered by the second author, who was a Ph.D. student in the clinical child psychology program and had obtained a high inter-rater reliability with the first author before the start of this study. The ADOS administration was done by the first author, who did not observe the administration of T-STAT and also did not have any information about the child before administering the ADOS. In three cases there was initial diagnostic uncertainty. As a result, a team meeting was held to discuss the cases and final diagnoses were made by the team.

Results: Study 1

Signal detection procedures were used to find the optimal cutoff score for the T-STAT. Results of signal detection associated with different cutoff scores for the developmental sample are shown in Table 3. The data revealed that the optimal cutoff scores for maximizing sensitivity and specificity appeared to be between 1.88 and 2.13. A T-STAT cutoff score of 2 for Autism risk was selected.

In Table 4, using the above cutoff score for Autism risk in the developmental sample yielded a sensitivity of 0.93 and a specificity of 0.87, demonstrating they were good indicators for identifying the Autism and non-Autism classifications. In addition, the positive predictive value (proportion of children who screened at high risk for Autism and those with Autism) was 0.88 and the negative predictive value was 0.93 (proportion of children who screened at low risk for Autism and without Autism). Both predictive values appear to be higher than the standard of suggestions from Glascoe (2005).

Table 3 Sensitivity and specificity for different T-STAT cutoff scores for development sample

Cutoff[a]	Sensitivity	Specificity
−1	1	0
0.13	1	0.03
0.38	1	0.05
0.63	1	0.18
0.88	1	0.26
1.13	0.98	0.39
1.38	0.93	0.48
1.63	0.93	0.58
1.88	**0.93**	**0.74**
2.13	0.83	0.79
2.38	0.72	0.82
2.63	0.52	0.87
2.88	0.46	0.92
3.13	0.41	0.94
3.38	0.33	0.95
3.63	0.24	0.97
3.88	0.11	0.98
5	0	1

[a] A score greater than or equal to the cutoff indicates autism risk

The bold values revealed that the optimal cutoff scores for maximizing sensitivity and specificity

Table 4 Classification between T-STAT and clinical diagnosis

T-STAT risk category	Clinical classification			
	Autism (n = 15)		DD (n = 15)	
High	14	93 %	2	13 %
Low	1	7 %	13	87 %

In the developmental sample, a total of three children were misidentified by the T-STAT. Two children were originally diagnosed as having DD/LI but had scored as high-risk for Autism on the T-STAT due to their shyness and anxiety about playing/interacting with the examiner, which might have caused their failing 3–4 items of the T-STAT. This result might indicate the T-STAT's tendency to overidentify young DD/LI children as being at high-risk for Autism or as having Autism. On the other hand, one child in the Autism group was also misidentified by the T-STAT. The girl had a CA of 33.03 months, a MA of 29.75, and a VMA of 31.5 months, the highest cognitive functioning in the Autism sample. Her high cognitive ability might explain the underestimation of Autism diagnosis by T-STAT.

Methods: Study 2—Validation in a Validation Sample

Participants

Participants included 77 children classified with diagnoses on the ADOS: 32 with Autism, 15 with PDD-NOS, and 30 with DD/LI. These children were recruited between 2005 and 2008 from IACCSDD. Eligibility requirements for participation were the same as in Study 1. Table 5 presents the demographic characteristics in the validation sample. The unselected clinic-based samples showed that there was no significant group difference in CA and SES; however, there were significant group differences in MA, NVMA, VMA, and also gender. Post hoc comparisons demonstrated that the Autism group had lower ages on MA, NVMA, and VMA than the DD/LI and PDD-NOS groups (ps < 0.01), as would be expected in a clinic-based population (Stone et al. 2000, 2004).

Measures and Procedures

All children were tested with an assessment battery after the parents signed informed consent. As mentioned in Study1, for obtaining the diagnostic data blindly, the research team worked in two places, the hospital and university. After referral from IACCSDD, the parents generally first attended the university lab for an assessment battery by the team organized by child psychologists; they then visited one of the two child psychiatrists within 1 month for a clinical evaluation based on DSM-IV-TR. Before evaluation, neither team member knew the child's diagnostic classification. For the assessment in the university, all data were collected in a single session including MSEL, T-STAT, other social communication scales, and the ADOS (Lord et al. 1999). For research purposes, the first author administered ADOS. He did not observe the administration of T-STAT, and he did not know any information about the child before administering the ADOS. The procedures ensured that the child's clinical diagnosis, T-STAT, and ADOS results were obtained independently.

In addition, in comparing the validity of the two changed items between STAT and T-STAT, while administering T-STAT, the original two items in STAT—"Puppet" in the domain of directing attention and "Rattle" in the domain of imitation—were tested.

Results: Study 2

Concurrent validity of the T-STAT was examined to compare the children in the Autism risk category with their ADOS classification (see Table 6). Because T-STAT was used for screening for Autism and not for Autism spectrum

Table 5 Demographic characteristics by ADOS classification

	ADOS classification			*F*	*p*	Group differences
	Autism (n = 32)	PDD-NOS (n = 15)	Non-Autism DD (n = 30)			
Chronological age (months)				0.92	0.41	
M (SD)	30.04 (3.98)	28.4 (4.42)	29.84 (3.83)			
Range	24–39.47	24–36	24–36			
NVMA (months)				11.51	<0.0005	AD, PDD < DD
M (SD)	20.48 (5.07)	21.97 (5.34)	26.22 (4.17)			
Range	12–32	14–34.5	20–42			
VMA (months)				18.86	<0.0005	AD, PDD < DD
M (SD)	12.77 (5.76)	15.77 (6.38)	21.77 (5.6)			
Range	5–25.5	7–31	13–44			
Mental age (months)				17.08	<0.0005	AD, PDD < DD
M (SD)	16.63 (5.21)	18.87 (5.61)	23.99 (4.48)			
Range	8.5–28.25	10.5–32.75	18.25–43			
SES						
M (SD)	61.91 (17.88)	53.2 (20.63)	54.63 (19.32)	1.59	0.21	
Range	35–91	28–91	8–91			
Sex[a]						
Male:female	28:4	11:4	18:12	6.10	<0.05	

[a] χ^2

disorders, children categorized as PDD-NOS on the ADOS were removed from the initial analysis. Hence, 62 children were included in the validation study. Using this sample, Cohen's Kappa for agreement between the T-STAT risk category and the ADOS classification was 0.90. Only three children were misidentified by the T-STAT, based on the ADOS classification: one from the Autism category and two from the DD/LI category. Thus using a cutoff score of 2, as Autism risk in the validation sample, resulted in a sensitivity of 0.97 and a specificity of 0.93, revealing a high hit-rate for identifying children who received a classification of either Autism or non-Autism. Furthermore, the positive predictive value (proportion of children classified with high risk of Autism and those with Autism) was 0.94, and the negative predictive value (the proportion of children who were classified with a low risk of Autism and or without Autism) was 0.97 in the validation sample. However, when classifying the category of PDD-NOS from the ADOS, five children (33 %) were misidentified as being in the no-risk group.

In addition, Table 6 also shows the relationship between the T-STAT risk category and clinical diagnosis. When the children with PDD-NOS were removed from the validation sample, the agreement between T-STAT and clinical diagnosis was 0.78 while using Cohen's Kappa. The data revealed that seven children were identified incorrectly by the T-STAT, one child was misclassified as no-risk Autism, and six children were misidentified as high-risk Autism. Again, when differentiating PDD-NOS from clinical diagnosis, five children were misidentified as being in the no-risk group (39 %) by T-STAT.

Although children with PDD-NOS were not included in the T-STAT category, it is necessary to know their performance on T-STAT compared with the children with Autism and DD/LI based on ADOS and clinical diagnosis. Table 7 shows significant group differences among the three groups with MA as a covariate on both the ADOS and clinical category basis. Post hoc comparisons ($p < 0.01$) revealed that children diagnosed as having Autism by the ADOS and Clinical Diagnosis had the higher T-STAT mean scores (i.e., higher risk for Autism) than the children with PDD-NOS and DD/LI. Children with PDD-NOS also scored significantly higher on T-STAT (i.e., higher risk for Autism) than children with DD/LI.

Because two items were substituted for the original STAT, we also examined the pass/fail ratio between the Autism group and DD/LI group on both the T-STAT and STAT based on the ADOS classification. In T-STAT, for the item "Posters," measuring responding joint attention, only eight children (25 %) with Autism passed the item compared to 30 children (100 %) with DD/LI who passed. Cohen's Kappa for the pass/fail ratio between the two groups was 0.74, revealing a good indicator for differentiation. However, on item "Fist," measuring body imitation, the data showed only four children (13 %) with Autism passed and only 10 children (33 %) with DD/LI passed. Cohen's Kappa was 0.21, manifesting low differentiation. On STAT, for the item "Puppet," measuring directing attention, only two children (6 %) with Autism passed the item compared to 10 (33 %) with DD/LI children who passed. Cohen's Kappa for the pass/fail ratio

Table 6 Concurrent validity of the T-STAT with ADOS and clinical classification

T-STAT risk category	ADOS classification					
	Autism (n = 32)		PDD-NOS (n = 15)		DD (n = 30)	
High	31	97 %	10	67 %	2	7 %
Low	1	3 %	5	33 %	28	93 %
	Clinical classification					
	Autism (N = 30)		PDD-NOS (N = 13)		DD (N = 34)	
High	29	97 %	8	62 %	6	18 %
Low	1	3 %	5	39 %	28	82 %

Table 7 Descriptive statistics for T-STAT scores by ADOS and clinical classification

	Autism	PDD-NOS	Non-Autism/DD	*F*	*p*	Group differences
ADOS (*N*)	32	15	30			
T-STAT						
Score	2.95 (0.75)	2.37 (0.85)	1.05 (0.60)			
Adjusted means[a]	2.76 (0.13)	2.31 (0.17)	1.28 (0.14)	27.820	<0.0001	Autism > PDD-NOS > Non-Autism
Clinical classification (*N*)	30	13	34			
T-STAT						
Score	2.93 (0.76)	2.32 (1.07)	1.26 (1.12)			
Adjusted means[a]	2.69 (0.14)	2.28 (0.20)	1.50 (0.13)	17.82	<0.0001	Autism, PDD-NOS > DD

[a] Covariate: MA

between the two groups was 0.28, suggesting a less powerful indicator for differentiation. Furthermore, for the item "Rattle," which tested object imitation, 22 children (69 %) with Autism passed compared to 26 children (87 %) with DD/LI who passed. Cohen's Kappa was 0.18, suggesting low differentiation.

Discussion

In Study 1, 30 children with Autism and DD/LI were matched on CA, MA, VMA, NVMA, SES, and gender in order to establish T-STAT's scoring algorithm. The study found a score of 2 as an appropriate cutoff score for Autism. Applying the cutoff score to the validation sample in Study 2, it was found to be acceptable for differentiating Autism but not for differentiating PDD-NOS from DD/LI. The two studies demonstrated that the T-STAT is a promising Level 2 screening tool for Autism in Chinese children at ages 2–3. The results not only replicated the original measure of STAT but also showed that even after two items of STAT were changed, the results were quite the same and acceptable.

Because two items were substituted for the original STAT, we wanted to check the two items in detail in both T-STAT and the original STAT. In Study 2, the data revealed that "Posters" from T-STAT was more powerful for classifying the groups of Autism and DD/LI than the item of "Puppet" from STAT. The results suggested that the item "Poster" is a better choice for measuring of responding joint attention than the item "Puppet." However, the item "Fist" for measuring body imitation was too much of a challenge for both groups of children. In addition, the item "Rattle" for measuring object imitation was apparently too easy for both groups of children to pass the test, indicating that it was an item which should not be used in the final version of T-STAT. Further studies are needed to identify one or two suitable items in the T-STAT for higher validity.

Three more issues need to be discussed: (1) the screening of children with PDD-NOS; (2) mental age; and (3) the diagnostic procedure.

First, regarding identifying children with PDD-NOS, because T-STAT was designed specifically for screening Autism, but not for all Autism spectrum disorders, it was not surprising that T-STAT is a good screening tool for Autism at risk but not for PDD-NOS or the ASD groups in Study 2. Since PDD-NOS is considered to be a less severe subtype of ASD and differs from Autism, children with PDD-NOS scored lower on T-STAT than did children with Autism. Further analysis used a score of 1.75, a lower cutoff score, to improve the accuracy of identifying children with PDD-NOS. The results revealed that based on the ADOS classification, four children (27 %) were misidentified as being in the no-risk group (five children were misidentified as being in the no-risk group using 2 as the cutoff score). However, six DD/LI children were misclassified as being in the high-risk Autism (only two children were misclassified as being in the high risk Autism using 2 as the cutoff score). The results suggested that decreasing the false negative for the PDD-NOS group on the T-STAT would increase the false positive for the DD group. Further studies are needed to develop a suitable scoring algorithm that is sensitive to PDD-NOS without over identifying children with non-Autism spectrum disorders.

Second, because our research participants in Study 2 were recruited from a clinical-based sample for the validation group, children with Autism in Study 2 obtained lower cognitive scores than the same-aged peers with Developmental Delay or language impairment. The phenomenon was similar to the Stone et al. (2004) study. It seems to indicate the impact of Autism symptomatology on early cognitive development. To exclude the confounding factor of mental age, the mental age was used as a covariate, which revealed that the highest mean scores on T-STAT were obtained by the Autism group, followed by the PDD-NOS and the DD/LI groups. The data showed that the degree of Autism severity as defined clinically (i.e., highest in Autism, moderate in PDD-NOS, and lowest in DD/LI) paralleled very closely to the T-STAT scoring algorithm for predicting the three groups.

Third, the diagnostic procedure in the study was not the same as that administered in the literature (e.g., Charwarska et al. 2009; Stone et al. 2004). In order to control for biased information before assessment, we identified two areas needing administration. For T-STAT and ADOS, two child psychologists blindly administered the two tests to get the T-STAT scores and ADOS classifications. Such a research design was consistent with the procedure in previous studies. However, to obtain the clinical diagnoses, two psychiatrists independently diagnosed the recruited children in the clinical setting. Although the data revealed an acceptable diagnostic agreement between ADOS and clinical judgment in the study, the diagnostic procedure should be modified in the future by using a multidisciplinary team including two psychiatrists to obtain the final diagnoses.

There are at least four advantages of using T-STAT as a screening tool: First, it provides a structured condition to elicit a child's various types of social communicative behaviors that can be observed directly instead of depending on parents' reports. Second, through interaction with the child, the examiner can have many opportunities to observe a child's behaviors quantitatively and qualitatively, and to learn the details of a child's performance.

Third, T-STAT takes only 20 min to administer, which would facilitate the promotion of T-STAT to practitioners including psychologists, psychiatrists, pediatricians, and other professionals in clinical settings. Fourth, due to the high agreement between ADOS and T-STAT, T-STAT might be used in place of ADOS as a shorter observational assessment in addition to other assessments required for Autism diagnosis in the children at-risk from the age of two to three years.

However, there are some limitations of using T-STAT. First, because T-STAT was developed as a quick assessment tool for screening only, the examiner has only one brief opportunity to do the evaluation. Second, training is required for the administration and scoring of T-STAT to ensure its reliability. Third, because the key items used in the T-STAT focused on coding negative symptoms which tend to include a combination of several behaviors, inexperienced examiners may miss observing part of the complex behaviors. Hence, some items may be coded as "pass" instead of "fail." Therefore, it may be a bit difficult for some junior clinicians to learn how to use T-STAT without getting adequate training time. Nonetheless, a recent pilot study has described a Web-based training format of STAT for professionals with diverse levels of education and experience. Such a format may be an alternative method to increase the junior clinicians' knowledge and observation skills for early Autism screening (Kobak et al. 2011).

As Stone et al. (2004) recommended, the setting should be extended to the community-based settings, and the utility of T-STAT should be studied in children with Autism below the age of two. However, additional work on psychometric properties of T-STAT is also needed.

Some benefits of using the T-STAT in Taiwan should be emphasized. First, T-STAT is the first and only Level 2 screening tool developed and validated in Taiwan. Second, clinical settings are busier in Taiwan than in western countries. Many clinicians are urged or required to take as little time as possible and to use efficient tools in their clinical work. Therefore, although administering the T-STAT requires training, it is conceivable that promoting T-STAT will be easier in Taiwan, Hong Kong, and China because the tool was developed in Taiwan.

Acknowledgments We are grateful to the parents and children who generously donated their time to this project. We would like to acknowledge the valuable contributions of Yen-Ting Yu, Hsing-Fang Lu, Chin-Ting Wang, Shu-Jung Chiang, Ya-Lin Peng, Chia-Ching Lin, Yi-Shan Huang and Hsieh Ya-Chi for the assessment process. Much thanks to Dr. Luke Y. Tsai at the University of Michigan Medical School and University of Michigan Rackham Graduate School for his invaluable comments and suggestions about the English editing. And much thanks to Prof. Al Staal at Eastern Michigan University and Mrs. Mary Staal for their suggestions about the final English editing for publication. This research was supported by grants from the National Science Council in Taiwan (NSC 93-2413-H-194-017, 94-2413-H-194-002, 95-2413-H-194-001 and 96-2413-H-004-021-MY3).

References

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

Bailey, A., Phillips, W., & Rutter, M. (1996). Autism: Towards an integration of clinical, genetic, neuropsychological, and neurobiological perspectives. *Journal of Child Psychology and Psychiatry, 37*, 89–126.

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

Barbaro, J., & Dissanayake, C. (2009). Autism spectrum disorders in infancy and toddlerhood: A review of the evidence on early signs, early identification tools, and early diagnosis. *Journal of Developmental and Behavioral Pediatrtics, 30*, 447–459.

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

Baron-Cohen, S., Cox, A., Baird, G., Swettenham, J., Nightingale, N., Morgan, K., et al. (1996). Psychological markers in the detection of autism in infancy in a large population. *British Journal of Psychiatry, 168*, 158–163.

Barton, M. L., Dumont-Mathieu, T., & Fein, D. (in press). Screening young children for autism spectrum disorders in primary practice. *Journal of Autism and Developmental Disorders.* doi: 10.1007/s10803-011-1343-5.

Charwarska, K., Klin, A., Paul, R., Macari, S., & Volkmar, F. (2009). A prospective study of toddlers with ASD: Short-term diagnostic and cognitive outcomes. *Journal of Child Psychology and Psyhiatry, 50*, 1235–1245.

Chiang, C. H., Soong, W. T., Lin, T. L., & Rogers, S. J. (2008). Nonverbal communication skill in skill in young children with autism. *Journal of Autism and Developmental Disorder, 38*, 1898–1906.

Filipek, P. A., Accardo, P. J., Baranek, G. T., Cook, E. H., Dawson, G., Gordon, B., et al. (1999). The screening and diagnosis of autistic spectrum disorders. *Journal of Autism and Developmental Disorders, 29*, 439–484.

Glascoe, F. P. (2005). Screening for developmental and behavioral problems. *Mental Retardation and Developmental Disabilities Research Review, 11*, 173–179.

Gotham, K., Bishop, S. L., & Lord, C. (2011). Diagnosis of autism spectrum disorders. In D. G. Amaral, G. Dawson, & D. H. Geschwind (Eds.), *Autism spectrum disorders* (pp. 30–43). NY: Oxford University Press.

Hsieh, C.-F., Soong, W.-T., & Hsu, C.-C. (1983). Autism: The validity of Clancy behavior scale. *Bulletin of Chinese Society of Neurology and Psychiatry, 9*, 17–26. (In Chinese).

Kanner, L. (1943). Autistic disturbances of affective contact. *The Nervous Child, 2*, 217–250.

Kleinman, J. M., Robins, D. L., Ventola, P. E., Pandey, J., Boorstein, H., Esser, E., et al. (2008). The modified checklist for autism in toddlers: A follow-up study investigating the early detection autism spectrum disorders. *Journal of Autism and Developmental Disorders, 38*, 827–839.

Kobak, K., Stone, W., Ousley, O. Y., & Swanson, A. (2011). Web-based training in early autism screening: Results from a pilot study. *Telemedicine and e-Health, 17*, 640–644.

Lord, C., Rutter, M., DiLavore, P. C., & Risi, S. (1999). *Autism diagnostic observation schedule (ADOS)*. LA: WPS.

Martínez-Pedraza, F. L., & Carter, A. S. (2009). Autism spectrum disorders in young children. *Child and Adolescent Psychiatric Clinics of North America, 18*, 645–663.

Mullen, E. (1995). *Mullen scales of early learning*. Cranston, RI: T.O.T.A.L. Child, Inc.

Mundy, P., & Burrette, C. (2005). Joint attention and neurodevelopmental models of autism. In F. R. Volkmar, R. Paul, A. Klin, & D. Cohen (Eds.), *Handbook of autism and pervasive developmental disorders* (3rd ed., pp. 650–681). NJ: John Wiley & Sons.

Ozonoff, S., Losif, A. M., Baguio, F., Cook, I. C., Hill, M. M., Hutman, T., et al. (2010). A prospective study of the emergence of early behavioral signs of autism. *Journal of American Academy of Child and Adolescent Psychiatry, 49*, 256–266.

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.

Rogers, S. (2009). What are infant siblings teaching us about autism in infancy? *Autism Research, 2*, 125–157.

Rogers, S. J., Hepburn, S. L., Stackhouse, T., & Wehner, E. (2003). Imitation performance in toddlers with autism and those with other developmental disorders. *Journal of Child Psychology and Psychiatry, 44*, 763–781.

Sattler, J. M. (2008). *Assessment of children: Cognitive foundation* (5th ed.). San Diego: Jerome M. Sattler, Publisher, Inc.

Schopler, E., Reichler, R. J., & Tenner, B. R. (1988). *The childhood autism rating scale*. LA: WPS.

Siegel, B. (1996). *Pervasive developmental disorders screening test (PDDST),* University of California at San Francisco (Unpublished manuscript).

Siegel, B., & Hayer, C. (1999, April). *Detection of autism in the 2nd and 3rd year: The pervasive developmental disorders screening test (PDDST).* Paper presented at the Society of Research in Child Development. Albuquerque, NM.

Stone, W. L., Coonrod, E., & Ousley, O. Y. (2000). Screening tool for autism in two-year-olds (STAT): Development and preliminary data. *Journal of Autism and Developmental Disorders, 30*, 607–612.

Stone, W. L., Coonrod, E. E., Turner, L. M., & Pozdol, S. L. (2004). Psychometric properties of the STAT for early autism screening. *Journal of Autism and Developmental Disorders, 34*, 691–701.

Stone, W. L., Lee, E. B., Ashford, L., Brissie, J., Hepburn, S. L., Coonrod, E. E., et al. (1999). Can autism be diagnosed accurately in children under three years? *Journal of Child Psychology and Psychiatry, 40*, 219–226.

Stone, W., McMahon, C. R., & Henderson, L. M. (2008). Use of the screening tool for autism in two-year-olds (STAT) for children under 24 months. *Autism, 12*, 557–573.

Stone, W. L., Ousley, O. Y., & Littleford, C. (1997a). Motor imitation in young children with autism: What's the object? *Journal of Abnormal Child Psychology, 25*, 475–485.

Stone, W. L., Ousley, O. Y., Yoder, P. J., Hogan, K. L., & Hepburn, S. L. (1997b). Nonverbal communication in two- and there-year-old children with autism. *Journal of Autism and Developmental Disorders, 27*, 677–696.

Swinkels, S. H., Dietz, C., van Daalen, E., Kerkhof, I. H., van Engeland, H., & Buitelaar, J. K. (2006). Screening for autistic spectrum in children aged 14 to 15 months. I: The development of the Early Screening of Autistic Traits Questionnaire (ESAT). *Journal of Autism and Developmental Disorders, 36*, 723–732.

Wetherby, A., & Prizant, B. (2002). *Communication and symbolic behavior scales developmental profile-first normed edition*. Baltimore, MD: Paul H. Brookes.

Wetherby, A., Wood, J., Allen, L., Cleary, J., Dickinson, H., & Lord, C. (2004). Early indicators of autism spectrum disorders in the second year of life. *Journal of Autism and Developmental Disorders, 34*, 473–493.

Whiten, A. (2006). The dissection of imitation and its "cognitive kin" in comparative and developmental psychology. In S. J. Rogers & J. G. Williams (Eds.), *Imitation and the social learning* (pp. 227–250). New York: The Guildford Press.