ORIGINAL RESEARCH

Validity and Reliability Study of the Chinese (Traditional) Version of the Devereux Early Childhood Assessment for Toddlers (DECA-T)

Sophie Hsin-Yi Liang [ID] [1]
Jen-Yu Chou [2]
Yu-Yu Wu [3]
Chin-Pang Lee [4]
Brent Allan Kelsen [ID] [5]
Yi-Chen Lee [6]

[1]Section of Department of Child and Adolescent Psychiatry, Department of Psychiatry, Chang Gung Memorial Hospital at Taoyuan and Chang Gung University College of Medicine, Taoyuan, Taiwan; [2]Graduate Institute of Humanities in Medicine, Taipei Medical University, Taipei, Taiwan; [3]Section of Department of Child and Adolescent Psychiatry, Department of Psychiatry, Chang-Gung Memorial Hospital at Linkou and Chang Gung University College of Medicine, Taiwan, YuNing Psychiatric Clinic, Taipei, Taiwan; [4]Department of Psychiatry, Chang-Gung Memorial Hospital at Linkou and Chang Gung University College of Medicine, Taoyuan, Taiwan; [5]Language Center, National Taipei University, New Taipei City, Taiwan, Department of Psychology, Auckland University of Technology, Auckland, New Zealand; [6]Department of Psychiatry, Chang-Gung Memorial Hospital at Chiayi and Chang Gung University College of Medicine, Chiayi, Taiwan

Correspondence: Yi-Chen Lee
School of Occupational Therapy, College of Medicine, National Taiwan University, Taipei, Taiwan
Email peggylee599@gmail.com

Background: The Devereux Early Childhood Assessment for Toddlers (DECA-T), which is one of the few standardized, norm-referenced behavioral rating scales related to young children's mental health, resilience, and social-emotional development, was developed for toddlers aged between 18 and 36 months.

Objective: The aim of this study was to test the clinimetric properties of the Chinese (Traditional) version of the DECA-T (C-DECA-T) using a classical test theory analysis and an item response theory analysis.

Methods: Seventy-five community-based toddlers aged from 18 to 36 months and 50 clinic-based participants recruited in hospitals in northern Taiwan participated in this study. Social-emotional competence was assessed by the C-DECA-T and children's behavior problems were rated via the Child Behavior Checklist 1.5 to 5 (CBCL/1.5–5). Homogeneity of the C-DECA-T was assessed by Mokken analysis; sensitivity and specificity were assessed via receiver operating characteristic curve.

Results: The results showed the C-DECA-T demonstrated good test–retest reliability (r=0.8) and high internal consistency (Cronbach's alpha = 0.94). Inter-rater reliability between father and mother was fair (ICC = 0.46). Convergent validity of the CBCL/1.5–5 total behavior problems (r=−0.26) demonstrated acceptable psychometric performance. The overall measure of the sampling adequacy of the C-DECA-T assessed by principal component analysis was 0.93. Mokken scale analysis showed the 36-items of the C-DECA-T formed a weak unidimensional scale (H_s=0.35), supporting its construct validity. The area under curve of the C-DECA-T in prediction of social-emotional disturbance was 0.70. The optimal cutoff of the Total Protective Factor score of the C-DECA-T was a T score of 40.1 (T40.1), with a sensitivity of 95% and a specificity of 68%. Item 2 ("show affection for a familiar adult") and item 33 ("calm herself/himself") provide a good amount of information for the assessment of social-emotional strength and needs of a toddler in clinical practice.

Conclusion: The C-DECA-T showed good psychometric properties. Our findings of high internal consistency of the three subscales and total score of the C-DECA-T suggest symptom manifestation of social-emotional competence and needs in Taiwanese toddlers is not culturally different from American toddlers. The clinimetric properties of the C-DECA-T examined by a classical test theory analysis approach and an item response theory analysis approach suggest that the C-DECA-T is a reliable and valid instrument for measuring social-emotional strength and needs in the population in Taiwan.

Keywords: Devereux Early Childhood Assessment for Toddlers, DECA-T, social-emotional competence, item response theory, Mokken scale analysis, principal component analysis

Introduction

Social-emotional development is defined as a child's developing ability to (a) experience, manage and express the full range of different kinds of emotions; (b) develop satisfying and close relationships with adults and other children; and (c) explore their environment actively and learn.[1,2] Successful engagement with parents, teachers, and peers promotes the development of many skills essential for emotional regulation, problem solving, empathy, cognition, and language.[3–5] Given that learning is closely associated with socially mediated interaction in preschool children, unrecognized or undervalued social and emotional problems not only affect overall development but also impact on their readiness for kindergarten or elementary school, and potentially cause serious lifelong developmental issues, premature mortality, and adult socioeconomic status.[1,6] However, early identification of and service provision to infants and toddlers with social-emotional/behavioral problems have lagged far behind advances in early identification and service provision for delays in cognition, language, and motor development. An estimated 10% to 15% of 1- and 2-year-old children experience significant social-emotional problems.[7] Moreover, research indicates that infant–toddler social-emotional/behavioral problems are not transient and highlight the need for early identification in multidomain and the effectiveness of early intervention.[8] Early identification of children at risk of social-emotional problems is critical to provide evidence-based intervention and optimize social-emotional development as well as educational success.[9–14]

Psychopathology in young children is often conceptualized as falling along the three broad domains of internalizing, externalizing, and dysregulation.[15] In early childhood, internalizing problems include difficulties with anxiety, depression/withdrawal, fears, and shyness/inhibition. Externalizing problems include aggression, overactivity, impulsivity, and inattention. Recently, extreme problems in the regulation of state, affect, and sensory processing have been addressed in the regulatory disorders of the Diagnostic Classification System for 0–3-year-olds.[16]

There are some parent-reported psychometric assessments for infant and toddler social-emotional development and problems,[1] including the Ages and Stages Questionnaires: Social-Emotional-2 (ASQ:SE-2),[17] Baby Pediatric Symptom Checklist (BPSC),[18] Preschool Pediatric Symptom Checklist (PPSC),[19] Brief Infant–Toddler Social and Emotional Assessment (BITSEA),[20] Child Behavior Checklist (CBCL/1.5–5),[21] Devereux Early

Childhood Assessment for Infants and Toddlers (DECA-I/T),[22] Early Childhood Screening Assessment (ECSA),[23] Greenspan Social-Emotional Growth Chart (SEGC),[24] Infant–Toddler Social and Emotional Assessment (ITSEA),[25] Merrill–Palmer-Revised Scales of Development (M-P-R)—Social-Emotional domain,[26] and Social-Emotional Assessment/Evaluation Measure (SEAM™).[27] Although the Child Behavior Checklist (CBCL/1.5–5)[21] and the Infant–Toddler Social and Emotional Assessment (ITSEA)[25] questionnaires have good reliability and validity, they have too many items for use as screening tools. Of the other two brief screeners in existence, the 35-item Eyberg Child Behavior Inventory[28] focuses on conduct problems in children aged 2 to 11 years and has demonstrated acceptable reliability and validity, while the 40-item Toddler Behavior Screening Inventory (TBSI)[29] addresses common behavior problems in 1- to 3-year-olds. However, they do not address both problems and competencies.

Few instruments focus on the social-emotional competence specific to the toddler period. The Devereux Early Childhood Assessment for Toddlers (DECA-T), which is one of the few standardized, norm-referenced behavior rating scales related young children's mental health, resilience, and social-emotional development, was developed with the aim of assessing toddlers aged between 18 and 36 months.[22] It is an assessment tool based on parent or childcare staff ratings after a 4-week observation period. It helps users to identify young children who may be experiencing social and emotional challenges, leading to interventions to reduce associated behavioral problems. The DECA-T contains 36 items measuring three protective factors, including Attachment/Relationships (A/R), Initiative (IN), and Self-Regulation (SR). The DECA-T has a Total Protective Factors scale (TPF), which is a composite across the scales for each assessment. Therefore, the scale provides an overall indication of the strength of young children's protective factors.

There are two different approaches for evaluation of psychometric properties of rating scales: classical test theory and item response theory.[30–33] The limitations of classical test theory include the summed score problem and sample-dependent statistics, which may result in different psychometric properties when based on different samples.[34,35] Item response theory analysis provides item-level statistics that are not influenced by differences between samples. Item response theory assumes that scale items can be ordered along the levels of a latent trait, with item "difficulty" demonstrating where items are difficult (rare) or less difficult (common). Mokken scale analysis is a non-parametric item

response theory derived from Guttman scaling.[36,37] On a Guttman scale, a single response can be used to predict responses to all items on the scale. Being a non-parametric analytical method, Mokken scale analysis is robust according to the underlying distribution of the data, thus avoiding the methodological limitations of previous studies.[36] To the best of our knowledge, no study has used Mokken scale analysis to examine the properties of the DECA-T. Therefore, the aim of this study was to evaluate the psychometric and clinimetric properties of the Chinese version of the DECA-T (C-DECA-T) in a population of toddlers aged 18 to 36 months, including those enlisted in public well-baby clinics (control group) and those from at-risk groups who were recruited from child psychiatric clinics (subject group). Homogeneity of the scale, item sensitivity, item severity, reliability, internal consistency, sensitivity, and specificity in discriminating an at-risk group for social-emotional problems and a healthy group were tested.

Methods
Participants and Setting
The study was carried out in Chang Gung Memorial Hospital at Linkou and YuNing Psychiatric Clinics. A co-author and colleagues (child psychiatrists and child psychologists) conducted translation (English to Chinese) and back-translation of the DECA-T, which was approved by the original publisher (Kaplan Press). The community-based sample initially consisted of 81 children aged 18 to 36 months who visited the hospital for conditions other than psychiatric problems. The clinical-based sample consisted of 57 children aged 18 to 36 months who visited the hospital seeking help for psychiatric problems. We excluded six subjects in the community-based sample and seven in the clinical-based sample due to one of the parents failing to complete the questionnaires.

The C-DECA-T was distributed to parents of the subjects and the healthy control group. Sample exclusion criteria included children with severe chronic physical illness. Mental health staff were trained to be familiar with the instruments and the procedures before the study. Both parents of each case were given a brief description of the study and written informed consent was obtained. Mothers and fathers rated the C-DECA-T separately to detect inter-rater reliability. Other information collected by self-rating instruments included demographic data of parents and children, birth weight, birth complication, physical condition, questions about parents' mental health (the Chinese Health

Questionnaire: CHQ-12), and the Child Behavior Checklist (CBCL/1.5–5). For test–retest reliability, recruited parents rated the C-DECA-T again in the following 14 days.

Ethics
The study data did not indicate any identifying personal information. Institutional Review Board approval was obtained from Chang Gung Memorial Hospital.

Instruments
Chinese (Traditional) Version of the Devereux Early Childhood Assessment for Toddlers (C-DECA-T)
The DECA-T is for toddlers ages 18 to 36 months and contains 36 items (on a five-point Likert scale from 0=never to 4=very frequently) measuring the social-emotional competence within the past 4 weeks, and includes three protective factors: Attachment/Relationship (A/R), Initiative (IN), and Self-Regulation (SR). The Attachment/Relationship (A/R) scale (18 items) assesses the relationship between the infant or toddler and significant adults such as family or teachers. The Initiative (IN) scale (11 items) assesses the toddler's ability to use independent thought and action to achieve his/her needs. The "Self-Regulation" (SR) scale (7 items) assesses the ability to gain control of and manage emotions and sustain focus and attention. The DECA-T has a Total Protective Factors (TPF) scale, which is a composite across the scales for each assessment. The TPF scale provides an overall indication of the strength of the young child's protective factors. We converted the raw data of three subscales and TPF into T scores following the original standardization table for the calculation of T scores and national percentiles for children of all ages (18–36 months). The internal reliability of the original DECA-T, rated by parents, was 0.94 in TPF, 0.87 in A/R, 0.92 in IN, and 0.79 in SR.[6]

Chinese Version of Child Behavior Checklist for Age 1.5–5, Parent Form (CBCL/1.5–5)
Children's behavior and emotional problems were assessed by the CBCL/1.5–5 parent form.[21] The CBCL/1.5–5 includes a set of 99 items with three response options (0=not true; 1=somewhat or sometimes true; 2=very true or often true), plus one open-ended item to add problems not listed on the form. Raw scores were analyzed for DSM-oriented scales as well as for original syndrome scales.[38] The DSM-oriented scales were designated as affective disorder, anxiety disorder, somatic disorder, attention-deficit/hyperactivity disorder, oppositional defiant disorder, and conduct disorder. The sum of the 100 items was counted as the total problems score, and

67 of them were scored specifically to seven narrowband behavioral syndromes, including emotionally reactive (9 items), anxious/depressed (8 items), somatic complaints (11 items), withdrawn (8 items), sleep problems (7 items), attention problems (5 items), and aggressive behavior (19 items). In addition, two broadband behavioral syndromes were derived, with the former four syndromes constituting an internalizing problem scale and the latter two syndromes constituting an externalizing problem scale. The Chinese version of the CBCL/1.5–5 was found to have good to excellent levels of reliability (internal consistency, test–retest reliability, and inter-parent agreement) and an acceptable model fit for the seven-syndrome factors when used in preschool children in Taiwan.[39]

Chinese Health Questionnaire (CHQ-12)

Parents' depression and anxiety were assessed by the CHQ-12, which has 12 items rated on a four-point Likert scale, with a higher score indicating more severe psychiatric symptoms. The CHQ-12 was a short form of CHQ (60 items) which was derived from a Chinese translation of the General Health Questionnaire,[40] with the addition of specially designed, culturally relevant items. The Cronbach's alpha in the community and clinical settings were 0.84 and 0.83, respectively. Persons with CHQ-12 scores of 4 and over are considered to have minor psychiatric disorders.[41,42]

Statistics and Data Analysis

First, we tested C-DECA-T properties using a classical theory test approach, including analyzing the reliability (test–retest reliability, internal consistency, and inter-rater reliability) and validity (concurrent, discriminative, and sensitivity/specificity). We conducted sensitivity and specificity analysis and receiver operating characteristic (ROC) analysis to find the cut-off value among the at-risk group and the healthy control group. Concurrent validity was hypothesized to be expressed in significantly negative Pearson correlations between the C-DECA-T TPF and the main CBCL/1.5–5 broad-band syndromes (internalizing, externalizing, and total behavior problems), as well as parent's depression and anxiety (CHQ-12). Conversely, small-to-medium negative correlations were expected to be found between the C-DECA-T TPF and the CBCL/1.5–5 and CHQ-12. All data management and statistical analyses were performed using SPSS version 18 software (SPSS Inc., Chicago, IL, USA).

Second, we evaluated C-DECA-T clinimetric properties using an item response theory approach. The C-DECA-T was evaluated by principal component analysis (PCA) first, followed by Mokken scale analysis. Mokken scale analysis was conducted using the Mokken package in R version 3.3.0 (R Foundation for Statistical Computing, Vienna, Austria).[43]

Principal Component Analysis (PCA)

PCA was conducted to classify the 36 items of the C-DECA-T using the psych package in R version 3.3.0 (R Foundation for Statistical Computing, Vienna, Austria). The Kaiser–Meyer–Olkin measure of sampling adequacy was computed to determine the adequacy of sample size. We extracted two unrotated components of the C-DECA-T (PC1 and PC2). PC1 referred to the general factor of the C-DECA-T, and PC2 referred to the dual factor of the C-DECA-T.

Mokken Scale Analysis (MSA)

Mokken scale analysis, a nonparametric form of item response therapy derived from Guttman scaling,[36,37] was developed to avoid the limitations of classical test theory including the summed score problem and sample-dependent statistics, which may result in different clinimetric properties when based on different samples. We checked the four assumptions of Mokken scale analysis, unidimensionality, local independence, monotonicity, and non-intersection. First, the assumption of unidimensionality was checked using Loevinger's scalability coefficients H.[43,44] Loevinger's scalability coefficients comprises three indexes: item-pair (H_{ij}), item (H_i), and scale (H_s) scalability coefficients. A scale is considered weak if $0.3 \leq H \leq 0.4$, moderate if $0.4 \leq H \leq 0.5$, and strong if $H \geq 0.5$. Second, we used conditional association to identify local dependence. Third, we checked serious violations of monotonicity for each item with a cutoff crit statistic >40. Fourth, we checked serious violations of non-intersection for each item with a cutoff crit statistic >80. For those items which met the four assumptions, we then checked invariant item ordering (IIO) of those items with method manifest invariant item ordering (MIIO). The Ligtvoet coefficient H^T was used to evaluate IIO. The IIO of a scale is considered weak if $0.3 \leq H^T \leq 0.4$, moderate if $0.4 \leq H^T \leq 0.5$, and strong if $H^T \geq 0.5$.[36,43]

Results
Data Description

Table 1 presents the subjects' demographic data and parents' education level. There was no statistically significant difference between the subject group and control group regarding

Table 1 Demographic Data and C-DECA-T T Scores

	Subjects (n=50)			Controls (n=75)			p-Value (t-Test)
	n		%	n		%	p-Value (Chi-square)
Sex							0.018
Male	32		64	31		41.3	
Female	18		36	44		58.7	
Paternal Education							0.292
Senior high school or below	11		22.4	10		13.3	
University or college	26		53.1	39		52	
Master degrees or above	12		24.5	26		34.7	
Maternal Education							0.131
Senior high school or below	10		20.4	6		8	
University or college	30		61.2	53		70.7	
Master degree or above	9		18.4	16		21.3	
	Mean		SD	Mean		SD	t-Test
Age (months)	27.905	±	4.46	27.025	±	5.353	0.342
Height (cm)	88.61	±	6.43	88.43	±	5.96	0.881
Body weight (kg)	12.53	±	2.25	12.5	±	1.86	0.946
Paternal C-DECA-T (T Score)							
Attachment/Relationships	46.60	±	9.78	52.27	±	9.55	0.008
Initiative	47.12	±	10.40	51.92	±	9.31	0.002
Self-Regulation	46.73	±	9.77	52.18	±	9.61	0.009
Total Protective Factors	46.60	±	9.78	52.27	±	9.55	0.002
Maternal C-DECA-T (T Score)							
Attachment/Relationships	47.08	±	11.40	51.95	±	8.48	0.007
Initiative	45.85	±	10.22	52.76	±	8.89	<0.001
Self-Regulation	46.02	±	9.43	52.65	±	9.54	<0.001
Total Protective Factors	45.81	±	10.54	52.79	±	8.62	<0.001

age, body weight, height, and parents' education level. In the subject group, there was a higher boy versus girl ratio than the control group. Table 1 also shows the total score and subscales of C-DECA-T in both groups. The results conveyed that the TPF and subscales – A/R, IN and SR in the subjects' group, rated either by father or mother – were significantly lower than those in the control group.

Reliability

Table 2 displays excellent internal consistency in Total Protective Factors, subscale of Attachment/Relationship (Cronbach's alpha, 0.94 and 0.92, respectively), good internal consistency in the subscale of Initiative (Cronbach's alpha, 0.84) and acceptable internal consistency in the subscale of Self-Regulation (Cronbach's alpha, 0.74). The test–retest reliability was found to be good in the TPF ($r=0.80$)

and acceptable in the subscales of A/R, IN, SR ($r=0.77$, 0.77, 0.72, respectively). Based on the Cicchetti and Sparrow guideline,[45] inter-rater reliability between father and mother for the TPF (ICC, 0.46), the subscales of A/R (ICC, 0.41) and IN (ICC, 0.52) were fair and the subscale of SR was good (ICC, 0.66).

Principle Component Analysis of the C-DECA-T

The overall measure of the sampling adequacy of the C-DECA-T was 0.93. All Kaiser–RMeyer–Olkin values for individual items of the C-DECA-T were between 0.87 and 0.97. Using principle component analysis, we extracted two unrotated components of the C-DECA-T (PC1 and PC2). As shown in Table 3, all items had positive loadings on the PC1.

Table 2 Reliability Statistics of the C-DECA-T

Indicator	Inter-Rater Reliability		Test–Retest Reliability			Scale Reliability				
	ICC [95% CI]	r [95% CI]	Paternal r [95% CI]	Maternal r [95% CI]	Parents r [95% CI]	Cronbach's α	Revelle's β	McDonald's ω_t	McDonald's ω_h	LCRC
TPF	0.46 [0.29, 0.59]	0.47 [0.36, 0.57]	0.82 [0.74, 0.88]	0.77 [0.67, 0.85]	0.80 [0.74, 0.85]	**0.94**	**0.72**	**0.95**	0.63	**0.94**
A/R	0.41 [0.23, 0.55]	0.37 [0.25, 0.48]	0.81 [0.72, 0.87]	0.73 [0.61, 0.81]	0.77 [0.70, 0.82]	**0.92**	**0.88**	**0.94**	**0.76**	**0.93**
IN	0.52 [0.37, 0.63]	0.48 [0.37, 0.58]	0.83 [0.75, 0.88]	0.73 [0.61, 0.81]	0.77 [0.71, 0.83]	**0.84**	0.69	**0.87**	0.68	**0.86**
SR	0.66 [0.55, 0.74]	0.58 [0.49, 0.67]	0.73 [0.61, 0.81]	0.71 [0.59, 0.80]	0.72 [0.64, 0.79]	**0.74**	0.51	**0.81**	0.59	**0.77**

Note: Bold value signifies a reliability index ≥0.7.
Abbreviations: ICC, intraclass correlations; CI, confidence interval; IC, internal consistency; AR, Attachment/Relationships; IN, Initiative; SR, Self-Regulation; TPF, Total Protective Factors.

Such results suggested that the 36 items of the C-DECA-T measured the same latent trait and supported the composite scale of the C-DECA-T, the TPF scale. Regarding the PC2, 18 items had positive loadings on PC2 (3, 8–10, 12, 16, 19, 23, 26–35) and another 18-items had negative loadings (1, 2, 4–7, 11, 13–15, 17, 18, 20–22, 24, 25, and 36). Compared with the original construct of the C-DECA-T, the items with positive loadings on the PC2 corresponded to the IN scale and the SR scale except two items (items 8 and 21).

Scalability of the C-DECA-T Using Mokken Scale Analysis

Table 3 also shows the scalability, monotonicity, and intersection of the C-DECA-T assessed using Mokken scale analysis. The result shows the C-DECA-T has a weak unidimensional scale (H_s=0.35), which referred the C-DECA-T as a multi-dimensional scale, supporting its construct validity. The C-DECA-T was divided into three subscales (H_s ranged from 0.35 to 0.47). The AR has a moderate reliable Mokken scale (H_s=0.47); the other two subscales, IN (H_s=0.38) and SR (H_s=0.35) are weak reliable Mokken scales (Table 3).

Validity, Sensitivity, and Specificity

The concurrent validity of the C-DECA-T was investigated by analysis of association between the C-DECA-T and the Chinese CBCL/1.5–5. The results showed the C-DECA-T was negatively correlated with both CBCL/1.5–5 broadband syndromes, including internalizing scale, externalizing scale, and total problem scale. The TPF was negatively correlated with all DSM-oriented scales of the CBCL/1.5–5 including depression, anxiety, autism spectrum disorder (ASD), attention-deficit hyperactivity disorder (ADHD), and oppositional

defiant disorder (ODD). Moreover, the TPF scale was negatively associated with five narrowband syndromes, including emotional reactive, anxiety/depressed, withdrawn, attention, and aggressive behavior problems (Table 4).

The discriminative validity of the C-DECA-T was carried out by comparing clinical subjects and control group C-DECA-T scores. The difference between the clinical subjects and control group C-DECA-T scores was significant with a medium effect size (Cohen's d=0.67).

The area under curve (AUC) of the C-DECA-T in prediction of social-emotional competence was 0.69 (95% confidence interval 0.59–0.79) (p < 0.000). The optimal cutoff of the TPF score of the C-DECA-T was a T score of 40.1 (T40.1), corresponding to the 19th percentile. The T40.1 had a sensitivity of 95% and a specificity of 68% (Figure 1).

Discussion

To the best of our knowledge, this is the first study to examine the clinimetric properties of the DECA-T in a Chinese ethnic population. Overall, the results showed that the C-DECA-T has a satisfactory psychometric performance with high internal consistency, inter-rater reliability, test–retest reliability, concurrent and discriminative validity. Principal component analysis showed the C-DECA-T has sampling adequacy. Mokken scale analysis showed the 36-items of the C- DECA-T formed a weak unidimensional scale (H_s=0.35), supporting C-DECA-T is a multi-dimensional scale.

Similar to the English version, the C-DECA-T is a reliable tool for measuring toddler social-emotional competence. The Cronbach's alpha and test–retest reliability of the C-DECA-T were 0.94 and 0.80, respectively. The

Table 3 Mokken Scale Analysis of the C-DECA-T

Subscale	Item Number	Mean ± SD	Hi	Monotonicity	Non-Intersection[a]
Attachment/Relationships H_s=0.47	6	3.62 ± 0.53	0.42	0	89[b]
	2	3.54 ± 0.59	0.34	0	47
	36	3.51 ± 0.60	0.35	0	57
	18	3.51 ± 0.60	0.33	0	74
	15	3.51 ± 0.61	0.4	0	94[b]
	4	3.50 ± 0.61	0.29	21	83[b]
	14	3.49 ± 0.63	0.34	0	60
	5	3.46 ± 0.63	0.36	0	63
	17	3.45 ± 0.61	0.3	0	82[b]
	20	3.44 ± 0.61	0.39	0	79
	11	3.44 ± 0.63	0.38	0	92
	22	3.44 ± 0.60	0.44	0	83[b]
	25	3.43 ± 0.57	0.43	0	70
	8	3.36 ± 0.65	0.43	0	71
	7	3.33 ± 0.65	0.44	0	84[b]
	1	3.23 ± 0.72	0.38	0	57
	13[c]	3.20 ± 0.73	0.4	0	78
	24	3.17 ± 0.72	0.38	0	53
	21	2.95 ± 0.76	0.23	22	87[b]
Initiative H_s=0.38	27	3.25 ± 0.66	0.39	0	55
	19	2.99 ± 0.77	0.29	18	76
	9	2.88 ± 0.79	0.38	0	63
	31[c]	2.85 ± 0.87	0.34	16	69
	16	2.84 ± 0.86	0.41	0	77
	26	2.81 ± 0.98	0.38	18	79
	32	2.71 ± 0.82	0.39	0	55
	12	2.64 ± 0.91	0.32	0	79
	28	2.61 ± 0.93	0.3	0	79
	10	2.60 ± 0.83	0.36	0	51
	29	2.36 ± 0.88	0.27	27	122[b]
Self-Regulation H_s=0.35	35	3.15 ± 0.77	0.23	0	89[b]
	3[c]	2.80 ± 0.76	0.34	0	77
	30	2.68 ± 0.79	0.32	0	91[b]
	34	2.49 ± 0.81	0.31	25	114[b]
	33	2.14 ± 0.83	0.33	0	55
	23	1.99 ± 0.77	0.28	0	82[b]
Total Protective Factor H_s=0.35					

Notes: [a]Non-intersection was checked using method rest score. [b]The violation of assumptions of monotonicity and non-intersection was assessed using the crit statistic: crit values >80 indicate serious violations. Items 4, 6, 7, 11, 15, 17, 21, 22, 23, 29, 30, 34, and 35 were removed before the backward selection. [c]Items 3, 13, and 31 were removed in the backward selection using method manifest invariant item ordering (MIIO). The coefficient H^T was 0.43 for the remaining 20 items.
Abbreviations: H_s, scale scalability; H_i, item scalability.

inter-rater reliability between mother and father in the C-DECA-T was fair (ICC = 0.46). The validity of the C-DECA-T was also supported by the significant negative correlation with total problems scale of the CBCL/1.5–5. Furthermore, the C-DECA-T was found to be negatively associated with internal and external behavior problems, narrowband syndromes (emotional reactive, anxious/depressed, social withdrawal, attention problems,

aggressive behaviors), and DSM-IV orientated scales (depression, anxiety, ASD, ADHD, ODD).

Mokken Scale Analysis Approach

Since this is the first study evaluating DECA-T clinimetric properties, we cannot compare the results with the literature. The Mokken scale analysis of the C-DECA-T revealed the items 4, 6, 7, 11, 15, 17, 21–23, 29, 30, 34,

Table 4 Pearson's Correlations Between the Subscales of the C-DECA-T and Demographic Date, Child Behaviour Problems, and Parental Emotional Distress

	A/R	IN	SR	TPF
CBCL/1.5–5				
DSM-Oriented Scales				
Depression	−0.174**	−0.214**	−0.192**	−0.221**
Anxiety	−0.069	−0.196**	−0.155*	−0.151*
ASD	−0.217**	−0.274**	−0.138*	−0.255**
ADHD	−0.054	−0.271**	−0.287**	−0.208**
ODD	−0.092	−0.243**	−0.257**	−0.209**
Narrowband syndrome				
Emotional reactive	−0.073	−0.235**	−0.122	−0.162*
Anxious/depressed	−0.104	−0.208**	−0.141*	−0.170**
Somatic complaints	−0.047	−0.122	−0.104	−0.099
Withdrawn	−0.274**	−0.242**	−0.146*	−0.272**
Sleep problems	−0.006	−0.085	−0.112	−0.065
Attention problems	−0.134*	−0.287**	−0.266**	−0.249**
Aggressive behaviour	−0.064	−0.234**	−0.236**	−0.185**
Broadband syndrome				
Internalizing scale	−0.160*	−0.303**	−0.204**	−0.253**
Externalizing scale	−0.083	−0.302**	−0.284**	−0.234**
Total problems scale	−0.134*	−0.307**	−0.259**	−0.255**
CHQ-12				
Father	−0.027	−0.190**	−0.315**	−0.166**
Mother	−0.142*	−0.229**	−0.130*	−0.197**

Notes: **Correlation is significant at the 0.01 level (two-tailed).*Correlation is significant at the 0.05 level (two-tailed).
Abbreviations: A/R, Attachment/Relationships; IN, Initiative; SR, Self-regulation; TPF, Total Protective Factors; ASD, autism spectrum disorder; ADHD, attention-deficit hyperactivity disorder; ODD, oppositional defiant disorder; CBCL/1.5–5, Chinese version of Child Behavior Checklist for Age 1.5–5, parent form; CHQ-12, Chinese Health Questionnaire.

and 35 were removed before the backward selection. Items 3, 13, and 31 were removed in the backward selection using method manifest invariant item ordering (MIIO). There are 20 items of the C-DECA-T having the property of invariant ordering (the coefficient H^T was 0.43 for the remaining 20 items), from the most prevalent (easiest) to the least prevalent (most difficult): 2, 36, 18, 14, 5, 20, 25, 8, 27, 1, 24, 19, 9, 16, 26, 32, 12, 28, 10, and 33. If the most prevalent item (easiest), such as item 2 ("show affection for a familiar adult") is rated negative, it means most items might be rated as negative as well. In contrast, if the least prevalent item (most difficult), such as item 33 ("calm herself/himself") is rated positive, it means most items might be rated as positive as well. The two items could provide the highest amount of information for clinical practice in the assessment of the social-emotional strength and needs in toddlers.

## Test–Retest Reliability and Inter-Rater reliability

The test–retest reliability of the C-DECA-T TPF (r=0.80) was found to be acceptable, though slightly lower than the original DECA-T, in the subscales of A/R, IN and SR (r=0.77, 0.77, 0.72, respectively). One reason for this might be due to the test–retest interval in the original DECA-T being 24 to 72 hrs compared to 14 days in the current study, thus allowing less time for symptoms to change.

The result showed that the correspondence (intraclass correlation coefficients; ICCs) between maternal and paternal rating in the C-DECA-T was fair in the TPF scale and subscales of A/R and IN, and good in the subscale of SR. In a meta-analysis study, Duhig et al suggested higher interparental score correspondence in externalized behaviors than the internalized behavior problems in children aged between 3 and 19 years.[46] Regarding the effect of child age on interparental rating correspondence, the meta-analysis showed lower interparental correspondence for children in early (aged 3 to 5 years old) and middle (aged 6 to 12 years old) childhood than for adolescents.[46] The observed differences between maternal and paternal ratings highlight the importance of gathering reports from both parents when assessing early social-emotional behavior problems. The inter-rater reliability coefficients for the original DECA-T score were 0.70 in TPF, 0.62 in A/R, 0.64 in IN, and 0.72 in SR. The interrater reliability coefficients were lower than the original DECA-T (Table 2); thus, cultural differences in the differences between father's and mother's reporting of toddlers social-emotional behavior/competences might be worthy of future study.

Parental Depression and Toddlers' Social-Emotional Competence in the C-DECA-T

Toddlers rated with lower social-emotional competence in the C-DECA-T would likely include those with problems that may be precursors to psychopathology and whose parents have distorted perceptions of child functioning, as may occur with parental depression.[47] In the present study, maternal depression/anxiety was significantly negatively correlated with toddlers' social-emotional competence on three subscales, while paternal depression/anxiety was also negatively correlated with social-emotional competence and toddler's social initiative and emotion self-regulations. Numerous studies have shown

ROC Curve

Diagonal segments are produced by ties.
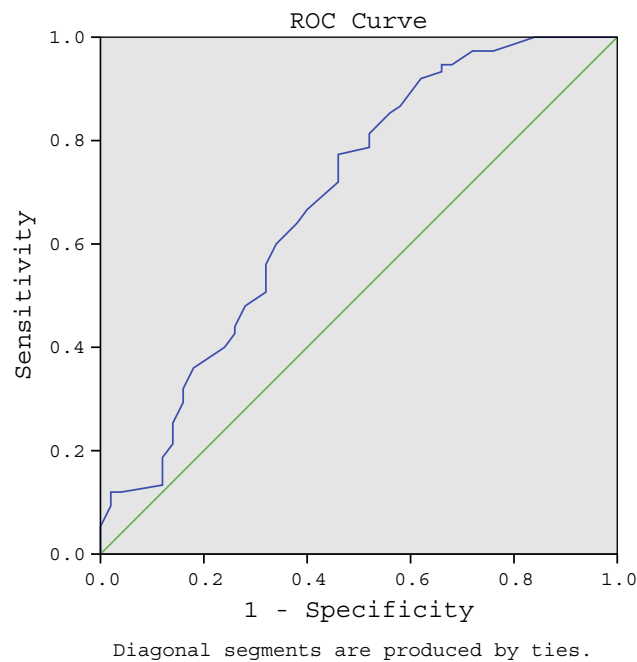
Figure 1 ROC curve of C-DECA-T.

maternal depression was associated with poor mother–infant interaction,[48] reduced infant social responsiveness and engagement,[2,49] increased risk of infant insecurity and preschool attachment,[50] leading to disturbances in the acquisition of competences, such as prosocial peer relations and mastery motivation,[51] and may predict later behavioral problems in school-age boys but not girls.[52] However, the effect of depression in fathers during the early years of a child's life has received little attention. One prospective population study found depression in fathers during the postnatal period was associated with adverse emotional and behavioral outcomes in children aged 3.5 years and an increased risk of conduct problems in boys after controlling for maternal postnatal depression and later paternal depression.[53] The mechanism between parental depression/anxiety and toddlers' social-emotional problems should be investigated in future studies.

Strengths and Limitations

The strength of our study is that it involved at-risk children (e.g., presenting symptoms of pervasive developmental disorders or of disruptive behavior disorders). We applied the Mokken scale analysis to assess the construct validity and dimensionality of the C-DECA-T. Although this investigation provided valuable information on the reliability, validity, and usefulness of a new child social-emotional resilience screening measurement, important study limitations should

be pointed out. One limitation of the present study is that information, such as parenting skill and parent–child interaction, which could affect children's social-emotional development, were not collected in the study. Another limitation of this study was that it was conducted in a highly urbanized setting; therefore, a high number of parents had university-level education, which may limit the generalizability of the findings from this study. Moreover, the sample size for this study was too limited to conduct factor analysis for the C-DECA-T. A larger sample will be required for future study. Despite these limitations, this is the first study examining the clinimetric proprieties of the C-DECA-T and applying clinimetric analyses to a tool measuring social-emotional strength and weakness in toddlers. Therefore, this short screening scale is suitable for use in surveys with two-phase designs, in which a quick screening test is used to select a subset of respondents who are likely to have social-emotional problems requiring more intensive psychiatric diagnostic assessment. Although the inter-rater reliability between father and mother was fair in TPF and acceptable in the subscales, the correspondence and discrepancy between maternal and paternal views toward toddlers' social-emotional problems/competence presents a potential area for future research. Future validation of the C-DECA-T in a larger population with greater variability in socio-economic status would increase the generalizability of our findings and reinforce our conclusions.

Abbreviations

C-DECA-T, the Chinese (Traditional) version of the Devereux Early Childhood Assessment for Toddlers; CBCL/1.5–5, Child Behavior Checklist 1.5 to 5; CHQ-12, Chinese Health Questionnaire; ADHD, attention-deficit hyperactivity disorder; ODD, oppositional defiant disorder; ASD, autism spectrum disorder; TPF, Total Protective Factors; A/R, Attachment/Relationships; IN, initiative; SR, Self-Regulation.

Disclosure

The authors report no conflicts of interest in this work.

References

1. Pontoppidan M, Niss NK, Pejtersen JH, Julian MM, Væver MS. Parent report measures of infant and toddler social-emotional development: a systematic review. *Fam Pract*. 2017;34(2):127–137. doi:10.1093/fampra/cmx003

2. Feinberg ME, Jones DE, Kan ML, Goslin MC. Effects of family foundations on parents and children: 3.5 years after baseline. *J Fam Psychol*. 2010;24(5):532–542. doi:10.1037/a0020837

3. Bulotsky-Shearer RJ, Bell ER, Romero SL, Carter TM. Identifying mechanisms through which preschool problem behavior influences academic outcomes: what is the mediating role of negative peer play interactions? *J Emot Behav Disord*. 2014;22(4):199–213. doi:10.1177/1063426613484806

4. Sektnan M, McClelland MM, Acock A, Morrison FJ. Relations between early family risk, children's behavioral regulation, and academic achievement. *Early Child Res Q*. 2010;25(4):464–479. doi:10.1016/j.ecresq.2010.02.005

5. Lillemyr OF. *Taking Play Seriously: Children and play in Early Childhood Education–An Exciting Challenge*. Charlotte (NC): Information Age Publishing; 2009.

6. Power C, Kuh D, Morton S. From developmental origins of adult disease to life course research on adult disease and aging: insights from birth cohort studies. *Annu Rev Public Health*. 2013;34:7–28. doi:10.1146/annurev-publhealth-031912-114423

7. Briggs-Gowan MJ, Carter AS, Skuban EM, Horwitz SM. Prevalence of social-emotional and behavioral problems in a community sample of 1-and 2-year-old children. *J Am Acad Child Adolesc Psychiatry*. 2001;40(7):811–819. doi:10.1097/00004583-200107000-00016

8. Briggs-Gowan MJ, Carter AS, Bosson-Heenan J, Guyer AE, Horwitz SM. Are infant-toddler social-emotional and behavioral problems transient? *J Am Acad Child Adolesc Psychiatry*. 2006;45 (7):849–858. doi:10.1097/01.chi.0000220849.48650.59

9. Pediatrics AAo. Committee on Children with Disabilities. Developmental surveillance and screening of infants and young children. *Pediatrics*. 2001;108(1):192–196.

10. Fantuzzo J, Stoltzfus J, Lutz MN, et al. An evaluation of the special needs referral process for low-income preschool children with emotional and behavioral problems. *Early Child Res Q*. 1999;14 (4):465–482. doi:10.1016/S0885-2006(99)00030-7

11. Hirai AH, Kogan MD, Kandasamy V, Reuland C, Bethell C. Prevalence and variation of developmental screening and surveillance in early childhood. *JAMA Pediatr*. 2018;172(9):857–866. doi:10.100 1/jamapediatrics.2018.1524

12. Gamel-McCormick M. *Status of Young Children and Their Mental Health Service Needs in Delaware*. Newark (DE): University of Delaware; 2002.

13. Denham SA, Zinsser KM, Brown CA. *The Emotional Basis of Learning and Development in Early Childhood Education*. Routledge; 2013.

14. Graziano PA, Reavis RD, Keane SP, Calkins SD. The role of emotion regulation in children's early academic success. *J Sch Psychol*. 2007;45(1):3–19. doi:10.1016/j.jsp.2006.09.002

15. Zeanah CH. *Handbook of Infant Mental Health*. New York (NY): Guilford Press; 2018.

16. Three Zt. *Diagnostic Classification, 0-3: Diagnostic classification of Mental Health and Developmental Disorders of Infancy and Early Childhood*. Washington (DC): National Center for Clinical Infant Programs; 2005.

17. Squires J, Bricker D, Twombly E. *The ASQ: SE User's Guide: For the Ages & Stages Questionnaires: Social-Emotional*. Baltimore (MD, USA): Paul H Brookes Publishing; 2002.

18. Sheldrick RC, Henson BS, Neger EN, Merchant S, Murphy JM, Perrin EC. The baby pediatric symptom checklist: development and initial validation of a new social/emotional screening instrument for very young children. *Acad Pediatr*. 2013;13(1):72–80. doi:10.1016/j.acap.2012.08.003

19. Sheldrick RC, Henson BS, Merchant S, Neger EN, Murphy JM, Perrin EC. The Preschool Pediatric Symptom Checklist (PPSC): development and initial validation of a new social/emotional screening instrument. *Acad Pediatr*. 2012;12(5):456–467. doi:10.1016/j.acap.2012.06.008

20. Karabekiroglu K, Briggs-Gowan MJ, Carter AS, Rodopman-Arman A, Akbas S. The clinical validity and reliability of the Brief Infant–toddler Social and Emotional Assessment (BITSEA). *Infant Behav Deve*. 2010;33(4):503–509. doi:10.1016/j.infbeh.2010.07.001

21. Achenbach TM, Rescorla LA. *Manual for ASEBA Preschool Forms & Profiles*. Burlington (VT): University of Vermont, Research Center for Children, Youth and Families; 2000.

22. Powell G, Mackrain M, LeBuffe P, Lewisville N. *Devereux Early Childhood Assessment for Infants and Toddlers Technical Manual*. Villanova (PA): Devereux Center for Resilient Children; 2007.

23. Gleason MM, Zeanah CH, Dickstein S. Recognizing young children in need of mental health assessment: development and preliminary validity of the early childhood screening assessment. *Infant Ment Health J*. 2010;31(3):335–357. doi:10.1002/(ISSN)1097-0355

24. Bayley N. *Bayley Scales of Infant and Toddler Development*. PsychCorp, Pearson; 2006.

25. Carter A, Briggs-Gowan M. *Manual of the Infant-Toddler Social-Emotional Assessment*. New Haven (CT): Yale University; 2000.

26. Roid GH, Sampers JL. *Merrill-Palmer-Revised: Scales of Development*. Wood Dale (IL): Stoelting; 2004.

27. Squires JK, Waddell ML, Clifford JR, Funk K, Hoselton RM, Chen C-I. A psychometric study of the infant and toddler intervals of the social emotional assessment measure. *Topics Early Child Spec Educ*. 2013;33(2):78–90. doi:10.1177/0271121412463445

28. Eyberg S, Boggs SR, Reynolds LA. *Eyberg Child Behavior Inventory*. Lutz (FL): PAR; 1998.

29. Mouton-Simien P, McCain AP, Kelley ML. The development of the toddler behavior screening inventory. *J Abnorm Child Psychol*. 1997;25(1):59–64. doi:10.1023/A:1025759408417

30. Fava GA, Carrozzino D, Lindberg L, Tomba E. The clinimetric approach to psychological assessment: a tribute to per bech, MD (1942–2018). *Psychother Psychosom*. 2018;87(6):321–326. doi:10.11 59/000493746

31. Carrozzino D. Clinimetric approach to rating scales for the assessment of apathy in Parkinson's disease: a systematic review. *Prog Neuro-Psychopharmacol Biol Psychiatry*. 2019;109641. doi:10.1016/j.pnpbp.2019.109641

32. Carrozzino D, Morberg BM, Siri C, Pezzoli G, Bech P. Evaluating psychiatric symptoms in Parkinson's Disease by a clinimetric analysis of the Hopkins Symptom Checklist (SCL-90-R). *Prog Neuro-Psychopharmacol Biol Psychiatry*. 2018;81:131–137. doi:10.1016/j.pnpbp.2017.10.024

33. Carrozzino D, Svicher A, Patierno C, Berrocal C, Cosci F. The euthymia scale: a clinimetric analysis. *Psychother Psychosoms*. 2019;88(2):119. doi:10.1159/000496230

34. Chang -C-C, Su J-A, Tsai C-S, Yen C-F, Liu J-H, Lin C-Y. Rasch analysis suggested three unidimensional domains for Affiliate Stigma Scale: additional psychometric evaluation. *J Clin Epidemiol*. 2015;68 (6):674–683. doi:10.1016/j.jclinepi.2015.01.018

35. Amin L, Rosenbaum P, Barr R, et al. Rasch analysis of the PedsQL: an increased understanding of the properties of a rating scale. *J Clin Epidemiol*. 2012;65(10):1117–1123. doi:10.1016/j.jclinepi.2012.04.014

36. Mokken RJ. *A Theory and Procedure of Scale Analysis: With Applications in Political Research*. Vol. 1. Walter de Gruyter; 2011.

37. Guttman L. *The Basis for Scalogram Analysis*. Bobbs-Merrill, College Division; 1974.

38. Achenbach TM, Dumenci L, Rescorla LA. DSM-oriented and empirically based approaches to constructing scales from the same item pools. *J Clin Child Adolesc Psychol*. 2003;32(3):328–340. doi:10.1207/S15374424JCCP3203_02

39. Wu Y-T, Chen WJ, Hsieh W-S, et al. Maternal-reported behavioral and emotional problems in Taiwanese preschool children. *Res Dev Disabil*. 2012;33(3):866–873. doi:10.1016/j.ridd.2011.11.018

40. Goldberg DP, Hillier V. A scaled version of the General Health Questionnaire. *Psychol Med*. 1979;9(1):139–145. doi:10.1017/S0033291700021644

41. Chong M-Y WG. Validation of 30-and 12-item versions of the Chinese Health Questionnaire (CHQ) in patients admitted for general health screening. *Psychol Med*. 1989;19(02):495–505. doi:10.1017/S0033291700012526

42. Cheng T-A, Williams P. The design and development of a screening questionnaire (CHQ) for use in community studies of mental disorders in Taiwan. *Psychol Med*. 1986;16(2):415–422. doi:10.1017/S0033291700009247

43. Van der Ark LA. New developments in Mokken scale analysis in R. *J Stat Softw*. 2012;48(5):1–27.

44. Loevinger J. The technic of homogeneous tests compared with some aspects of "scale analysis" and factor analysis. *Psychol Bull*. 1948;45 (6):507. doi:10.1037/h0055827

45. Cicchetti DV, Sparrow SA. Developing criteria for establishing inter-rater reliability of specific items: applications to assessment of adaptive behavior. *Am J Ment Defic*. 1981;86(2):127–137.

46. Duhig AM, Renk K, Epstein MK, Phares V. Interparental agreement on internalizing, externalizing, and total behavior problems: a meta-analysis. *Clin Psychol-Sci Pr*. 2000;7(4):435–453. doi:10.10 93/clipsy.7.4.435

47. Briggs-Gowan MJ, Carter AS, Schwab-Stone M. Discrepancies among mother, child, and teacher reports: examining the contributions of maternal depression and anxiety. *J Abnorm Child Psychol*. 1996;24(6):749–765. doi:10.1007/BF01664738

48. Campbell SB, Cohn JF, Meyers T. Depression in first-time mothers: mother-infant interaction and depression chronicity. *Dev Psychol*. 1995;31(3):349–357. doi:10.1037/0012-1649.31.3.349

49. Granat A, Gadassi R, Gilboa-Schechtman E, Feldman R. Maternal depression and anxiety, social synchrony, and infant regulation of negative and positive emotions. *Emotion*. 2017;17(1):11–27. doi:10.1037/emo0000204

50. Murray L, Fiori-Cowley A, Hooper R, Cooper P. The impact of postnatal depression and associated adversity on early mother-infant interactions and later infant outcome. *Child Deve*. 1996;67 (5):2512–2526. doi:10.2307/1131637

51. Elicker J, England M, Sroufe LA. Predicting peer competence and peer relationships in childhood from early parent-child relationships. In Parke RD, Ladd GW (editors).*Family-Peer relationships: Modes of Linkage*. Hillsdale (NJ, USA): Lawrence Erlbaum Associates, Publishers; 1992.

52. Shaw DS, Vondra JI. Infant attachment security and maternal predictors of early behavior problems: a longitudinal study of low-income families. *J Abnorm Child Psychol*. 1995;23(3):335–357. doi:10.1007/BF01447561

53. Ramchandani P, Stein A, Evans J, O'Connor TG, Team AS. Paternal depression in the postnatal period and child development: a prospective population study. *Lancet*. 2005;365(9478):2201–2205. doi:10.1016/S0140-6736(05)66778-5