Infant Behavior & Development 32 (2009) 291–297

Contents lists available at ScienceDirect

Infant Behavior and Development

The reliability and validity of the Turkish version of the brief infant–toddler social emotional assessment (BITSEA)

Koray Karabekiroglu[a,*], Ayse Rodopman-Arman[b], Pinar Ay[c], Mustafa Ozkesen[d], Seher Akbas[a], Gokce Nur Tasdemir[a], Omer Boke[e], Yildiz Peksen[f]

[a] Child & Adolescent Psychiatry Department, Ondokuz Mayis University, School of Medicine, Samsun, Turkey
[b] Child & Adolescent Psychiatry Department, Marmara University, School of Medicine, Istanbul, Turkey
[c] Public Health Department, Marmara University, School of Medicine, Istanbul, Turkey
[d] Ministry of Health, Samsun, Turkey
[e] Psychiatry Department, Ondokuz Mayis University, School of Medicine, Samsun, Turkey
[f] Public Health Department, Ondokuz Mayis University, School of Medicine, Samsun, Turkey

ARTICLE INFO

Article history:
Received 14 July 2008
Received in revised form 4 March 2009
Accepted 30 March 2009

Keywords:
BITSEA
Toddlers
Validity
Reliability
Turkey

ABSTRACT

In this study the reliability and validity of the Turkish version of the brief infant–toddler social emotional assessment (BITSEA) were investigated in a community sample. The sample consisted of 462 children (mean age: 24.60 ± 7.93 [12–42] months) who had applied to Turkish health centers for immunization. Both parents completed the BITSEA; mothers completed the child behavior checklist 2/3 (CBCL). Internal consistencies of the BITSEA–problem (P) and competence (C) scales were good to excellent (Cronbach's $\alpha = 0.82$ and 0.72, respectively). Interrater reliability between parents and test–retest reliability were good. BITSEA/P scores were significantly correlated with CBCL internalizing, externalizing and total problem scores ($p < 0.001$). Maternal BITSEA/P cutpoint scores revealed that 30.6% of male toddlers and 28.6% of females were in the subclinical range and 13.1% of males and 17.6% of females were in clinical range. Results reveal that the Turkish version of BITSEA is a reliable, valid and simply applicable instrument for screening social, emotional and behavioral problems among toddlers. Clinical validation of the BITSEA/C and BITSEA/P is warranted.

Cellular migration, synapse formation and structural organization in the central nervous system, as well as psychosocial development, including the formation of the caregiver–infant attachment system, reach the most significant velocity during early childhood. Therefore, this period has been considered a "critical period" for the development of several psychiatric pathologies later in life (Anderson & Cohen, 2002). There is increasing recognition of the importance of early detection and provision of intervention services for infants and toddlers who have significant social-emotional and/or behavioral problems (American Academy of Pediatrics [AAP], 2001). Early assessment and intervention may reduce parental anxiety while making parents more self-confident and successful in problem solving skills during child-rearing practices (Keren, Feldman, & Tyano, 2003). In addition, even simple assessment procedures may have positive contributions to the infant–parent relationships (Lieberman, Van Horn, Grandison, & Pekarsky, 1997; Meisels & Fenichel, 1996). The increasing concerns about the psychiatric problem behaviors that appear in the early childhood require research studies (Shaw, Owens, Giovanelli, & Winslow, 2001) as well as reliable and valid screening instruments for this age group (Briggs-Gowan, Carter, Irwin, Wachtel, & Cicchetti, 2004).

* Corresponding author at: Çocuk Psikiyatrisi Poliklinigi, Çocuk Hastanesi, Ondokuz Mayis Universitesi, 55139 Kurupelit, Samsun, Turkey. Tel.: +90 362 312 19 19x3696–3753; fax: +90 216 363 54 05.
E-mail addresses: drkorayk@hotmail.com (K. Karabekiroglu), aarman@marmara.edu.tr (A. Rodopman-Arman).

0163-6383/$ – see front matter.
doi:10.1016/j.infbeh.2009.03.003

An estimated 20% of 2- to 3-year-old children in Turkey experience social-emotional problems at the subclinical level (Erol, Şimşek, Öner, & Münir, 2005) and 10–15% of 1- and 2-year-old children in the United States experience significant social-emotional problems (Roberts, Attkisson, & Rosenblatt, 1998). However, a study in the United States with a representative sample indicated that fewer than 8% of 1- and 2-year-old with social-emotional/behavioral problems received any developmental or mental health services (Horwitz, Gary, Briggs-Gowan, & Carter, 2003). Even though direct observation of the toddlers' behaviors and parent–child interactions remain as a gold standard in diagnosis and assessment, these procedures are time-consuming and expensive for outpatient clinics and/or home visits. In addition, infant–toddler behaviors during an office visit may not be representative of the child's overall behavioral repertoire in other settings. On the other hand, most parents can complete checklists independently, requiring minimal staff time (Glascoe, 2000). Routine pediatric screening using brief screeners have repeatedly been shown to be feasible and they significantly improve the identification of at-risk infants and toddlers whose difficulties might warrant additional follow-up or intervention (Baird et al., 2000; Jellinek et al., 1999; Murphy et al., 1996). Moreover, during the infant–toddler period, comprehensive measures appropriate for social-emotional and behavioral problem screening are limited (Glascoe, 2000).

Among the current scales used in these age groups, the 42-item brief infant–toddler social and emotional assessment (BITSEA) (Briggs-Gowan et al., 2004) was designed as a screening tool for parents and child-care providers to identify children "at risk" for developing as well as currently experiencing social-emotional/behavioral problems. Delays in social-emotional competence, as well as early signs and symptoms of autism spectrum disorders are also covered within the instrument. BITSEA items were drawn from the larger pool of Infant–Toddler Social Emotional Assessment (ITSEA) questions (Carter, Briggs-Gowan, Jones, & Little, 2003). In the reliability and validity study of original version of BITSEA (Briggs-Gowan et al., 2004), test–retest reliability was excellent and interrater agreement (mother/father and parent/child-care provider) was good. Findings support the BITSEA as a consistent and sufficient screener for social-emotional/behavioral problems and delays in social-emotional competence. As the BITSEA items were embedded and answered within the longer ITSEA, parents may answer BITSEA questions differently when they are not asked in the context of the longer ITSEA (Briggs-Gowan et al., 2004). Further, the BITSEA may behave slightly differently when used in pediatric or early intervention samples, which may have greater proportions of children with delays or health problems. Therefore, additional research with the BITSEA is warranted (Briggs-Gowan et al., 2004).

Similar to many countries, there are only a limited number of valid and reliable screening tools for assessing social and emotional problems and delays in competence for toddlers in Turkey. As the BITSEA is a promising and practically applicable screening tool, we aimed to investigate the reliability and validity of a Turkish version of BITSEA in a community sample.

1. Methods

1.1. Participants and setting

The study was carried out in Samsun, a province on the central-north coast of Turkey with approximately 1.2 million inhabitants. One- to four-year-old children constitute 6.1% of the total population in this city. This community sample consisted of 462 children who had applied to the primary health care centers located in the city for vaccination purposes. The sample sizes for all outpatient clinics were calculated according to the population rate—both urban (63.3%) and rural (36.7%) areas of all districts. Instruments were given to the parents of 660 toddlers applying to the vaccination center consecutively. Parents of 462 toddlers (70.0%) returned the completed instruments to the health center. The mean age of the toddlers was 24.60 ± 7.93 (12–42) months. In this study, the upper limit of early childhood is considered as 42 months, which is consistent with literature based on infant and toddler psychiatry (Keren et al., 2003; Thomas & Guskin, 2001). Slightly over half (54.5%; $n = 252$) of the toddlers were male. On average, mothers were 29.11 years of age (SD = 5.03, range = 19–44) and fathers were 33.12 years of age (SD = 6.09; range = 20–72). Only 25.1% of the mothers were working. While 20.7% of the mothers and 25.9% of fathers had a university degree, 25.9% of mothers and 34.1% of fathers graduated from high school. Three mothers and one father had not completed primary school.

1.2. Procedure

Instruments were distributed to all of the primary health care clinics ($n = 47$). Health staff in all centers were trained in the study procedures. Both mothers and fathers were given a brief description of the study and informed consent was obtained. For the test–retest reliability, the first two parents in each center who had brought the completed measures back were given the measures a second time to be completed in the following 15–30 days. Parental education level, the BITSEA and the CBCL scores were not significantly different between the retest group and the rest of the sample.

1.3. Measures

1.3.1. BITSEA–Turkish

The brief infant–toddler social and emotional assessment (BITSEA) consists of 42 items. The BITSEA Problem scale is comprised of 31 items (BITSEA/P) and the BITSEA Competence scale is comprised of 11 items (BITSEA/C) (Briggs-Gowan & Carter, 2002). The response format for each item has three choices: "not true/rarely" (0), "sometimes true/sometimes" (1) and

"very true/often" (2). Higher total scores on BITSEA/P indicate a higher level of behavioral and emotional problems and lower total scores on BITSEA/C indicate a lower level of competence. The BITSEA was translated to Turkish by the first author of this article. Back translation into English was carried out by a professor of medical English in the Medical School. Both English versions were compared and minimal corrections were made for the Turkish version. The final version of BITSEA–Turkish was further adapted based on the results of a pilot study with a community sample of parents of toddlers ($n = 45$).

1.3.2. Child behavior checklist/2–3 (CBCL/2–3)

The CBCL/2–3 is designed for evaluating young children to obtain ratings of behavioral/emotional problems from parents; it consists of 99 main items rated 0, 1, and 2. It includes internalizing, externalizing, and total problem scales, plus six syndrome scales. Respondents rate the child on each item that describes the child now or in the previous 2 months. The Turkish translation has good 1-week test–retest reliability ($r = 0.88$) and 1-year stability ($r = 0.73$). In addition, the internal consistency (internalizing Cronbach's alpha = 0.77, externalizing Cronbach's alpha = 0.76, total problem Cronbach's alpha = 0.82), and interrater agreement are high (Erol et al., 2005). The Turkish version of CBCL/2–3 identifies 16.5% of toddlers as subclinical and 10.9% as clinical (Erol et al., 2005). A newer version, the CBCL/1.5–5 (Achenbach & Rescorla, 2000), designed for children 18 months through 5 years, has not yet been validated in Turkey.

1.3.3. Zero-to-three psychiatric assessment sociodemographic form

This form included most sociodemographic, medical and developmental variables important for early childhood psychiatric assessment, such as complaint history, family relations, medical problems, and cultural background (Thomas et al., 1997).

1.4. Data analyses

The analytic plan included assessment of the scale structure, test–retest reliability, interrater reliability, and 2- to 4-week stability of the Turkish version of BITSEA. In addition, to inform the assignment of statistically at-risk cutpoints, BITSEA scales were evaluated for age and sex effects. Validity, which was examined only for the BITSEA–problem subscale, was examined dimensionally and through dichotomous sensitivity–specificity analyses. Criterion related validity was evaluated by comparing the BITSEA with the CBCL/2–3.

Eighteen (3.8%) mothers and 45 (9.7%) fathers did not complete the BITSEA. Thus, totally 444 mothers and 417 fathers completed the forms. Among mothers, 58 had some missing data with 27 missing one item, 15 missing two items, and 16 missing more than two items. Sixty-two fathers had some missing data with 41 missing one item, 13 missing two items, and 8 missing more than two items. For completed questionnaires which had one or two missing items, the missing item score was replaced with the score of the item with which the missing item was most significantly correlated. Internal consistency of maternal and paternal BITSEA–problem and competence scores were calculated with Cronbach α. Categorical variables were analysed by the chi square test. Scores were compared with the parametric (Pearson's correlation) and non-parametric tests (Spearman's correlation) depending on the normality of the data. Interrater reliability and test–retest reliability of BITSEA–problem and competence scores, and correlations between the BITSEA–problem scores and CBCL scores were calculated with Spearman's correlations. The Tukey HSD and/or Mann–Whitney U tests were used in the post hoc analysis. As we calculate gender difference in BITSEA (two subscores) and CBCL scores (three subscores) in five subscores, for Bonferroni correction, a p value of less than 0.01 (0.05/5) was determined as the level of statistical significance. The BITSEA problem cutpoints were identified by using subclinical/clinical CBCL/2–3 cutpoint scores. As recommended for CBCL/2–3 (Erol & Simsek, 1997) for subclinical cutpoint 1 standard deviation (SD), and for clinical cutpoint 1.5 SD were selected. In determining the clinical and subclinical cutpoints of BITSEA/P, similar to what have been done in the original validation study of BITSEA, at least 85% sensitivity was selected (Briggs-Gowan et al., 2004). SPSS 11.0 was used for all statistical calculations.

2. Results

There were no significant differences between maternal and paternal ratings on the BITSEA/P or BITSEA/C scores in any age or gender group (Table 1). In addition, both maternal and paternal BITSEA/P scores did not reveal significant differences between age groups ($p > 0.05$). However, consistent with the developmental nature of the BITSEA competence scale, both maternal and paternal BITSEA/C scores increased across the age groups ($p < 0.001$). Exploration for gender difference showed that only paternal BITSEA/C scores of girls were significantly higher than boys ($p < 0.01$); maternal BITSEA/C scores approached significance, $p = 0.014$). Post hoc analysis revealed that the gender difference was restricted to the 24- to 29-month age group ($p < 0.01$). Table 1 presents the mean scores of all scales for all age and gender groups.

2.1. Reliability

2.1.1. Internal consistency

The internal consistency of the BITSEA–problem (P) and competence (C) were good to excellent. The internal consistency of the BITSEA/P was 0.81 (Cronbach α) for the maternal and 0.83 (Cronbach α) for the paternal scales. The internal consistency of the BITSEA/C was 0.72 (Cronbach α) for both the maternal and paternal scales.

Table 1
Mean scores and standard deviations of all scales for all age and gender groups.

Group	BITSEA–problem scale		BITSEA–competence scale		Child behavior checklist (CBCL)		
	Maternal	Paternal	Maternal	Paternal	Internalizing	Externalizing	Total score
Subjects 12–17 months of age							
M (n = 65)	16.55 ± 7.90	16.56 ± 8.07	13.62 ± 3.91	13.38 ± 3.89	7.85 ± 5.28	16.12 ± 9.01	52.32 ± 28.00
F (n = 40)	17.11 ± 9.76	17.76 ± 9.77	13.53 ± 4.57	13.40 ± 4.65	7.25 ± 5.15	14.00 ± 9.22	47.07 ± 28.79
T (n = 105)	16.77 ± 8.66	17.00 ± 8.70	13.58 ± 4.16	13.39 ± 4.15	7.63 ± 5.18	15.34 ± 9.03	50.35 ± 28.04
p	>0.01	>0.01	>0.01	>0.01	>0.01	>0.01	>0.01
Subjects 18–23 months of age							
M (n = 60)	15.98 ± 8.91	15.44 ± 8.46	15.77 ± 3.51	15.95 ± 3.40	6.80 ± 5.21	14.35 ± 4.68	42.23 ± 24.42
F (n = 58)	16.83 ± 6.95	17.61 ± 8.72	16.92 ± 3.16	17.40 ± 3.07	5.97 ± 3.27	15.60 ± 7.86	42.82 ± 18.68
T (n = 118)	16.39 ± 8.00	6.46 ± 8.61	6.31 ± 3.38	6.63 ± 3.3	6.42 ± 4.41	14.93 ± 7.74	42.50 ± 21.82
p	>0.01	>0.01	>0.01	>0.01	>0.01	>0.01	>0.01
Subjects 24–29 months of age							
M (n = 60)	16.64 ± 7.52	16.40 ± 8.72	15.43 ± 3.68	14.64 ± 3.99	7.08 ± 4.12	16.83 ± 7.74	47.38 ± 18.44
F (n = 43)	15.33 ± 6.63	14.85 ± 6.59	17.00 ± 3.28	17.40 ± 2.71	6.27 ± 4.93	14.08 ± 7.98	40.43 ± 24.64
T (n = 103)	16.06 ± 7.14	15.72 ± 7.85	16.12 ± 3.58	15.82 ± 3.74	6.73 ± 4.48	15.64 ± 7.92	44.35 ± 21.50
p	>0.01	>0.01	0.014	0.002*	>0.01	>0.01	>0.01
Subjects 30–35 months of age							
M (n = 43)	15.10 ± 6.61	16.13 ± 7.34	15.51 ± 3.37	15.58 ± 3.81	6.28 ± 3.94	15.68 ± 8.17	43.83 ± 19.94
F (n = 48)	16.07 ± 8.93	16.13 ± 9.51	16.40 ± 3.82	16.16 ± 3.43	7.11 ± 4.60	15.53 ± 9.17	47.00 ± 25.56
T (n = 91)	15.62 ± 7.90	16.13 ± 8.52	15.99 ± 3.62	15.89 ± 3.60	6.72 ± 4.31	15.60 ± 8.67	45.54 ± 23.07
p	>0.01	>0.01	>0.01	>0.01	>0.01	>0.01	>0.01
Subjects 36–42 months of age							
M (n = 24)	13.74 ± 5.85	14.19 ± 5.51	17.05 ± 3.47	16.10 ± 3.64	6.35 ± 3.72	13.43 ± 6.76	40.87 ± 19.56
F (n = 21)	16.55 ± 9.20	14.06 ± 10.2	15.55 ± 3.48	14.24 ± 3.21	8.45 ± 5.61	17.05 ± 8.92	50.42 ± 25.97
T (n = 45)	15.05 ± 7.63	14.13 ± 7.85	16.33 ± 3.51	15.26 ± 3.53	7.33 ± 4.76	15.12 ± 7.95	45.19 ± 22.90
p	>0.01	>0.01	>0.01	>0.01	>0.01	>0.01	>0.01
All subjects							
M (n = 252)	15.89 ± 7.69	15.84 ± 7.96	15.25 ± 3.73	14.99 ± 3.86	6.86 ± 4.48	15.46 ± 7.91	45.25 ± 21.95
F (n = 210)	16.37 ± 8.13	16.37 ± 8.87	16.04 ± 3.85	16.10 ± 3.73	6.84 ± 4.63	15.25 ± 8.52	44.90 ± 24.26
T (n = 462)	16.11 ± 7.89	16.13 ± 8.38	15.61 ± 3.81	15.49 ± 3.84	6.85 ± 4.54	15.36 ± 8.18	45.09 ± 23.02
p	>0.01	>0.01	0.014	0.002*	>0.01	>0.01	>0.01

Note: M = male; F = female; T = total.
* Significant difference between scores of male and female toddlers.

2.1.2. Interrater reliability

Interrater reliability between parents was good (BITSEA/P, Spearman's rho = 0.68, $p < 0.001$; BITSEA/C, Spearman's rho = 0.71, $p < 0.001$). Interrater reliability, which was examined as percent agreement on the identified cutpoints of BITSEA/P scores revealed significant agreement (subclinical: $p < 0.001$; clinical: $p < 0.001$ [chi-square tests]).

2.1.3. Test–retest reliability

The test–retest reliability was good for mothers (BITSEA/P Spearman's rho = 0.49, $p < 0.001$; BITSEA/C Spearman's rho = 0.41, $p < 0.01$) and fathers (BITSEA/P Spearman's rho: 0.83, $p < 0.001$; BITSEA/C Spearman's rho = 0.60, $p < 0.001$).

2.2. Validity

2.2.1. Criterion validity and correlational analyses of BITSEA with other scales

The maternal and the paternal BITSEA/P scores were significantly correlated with the CBCL internalizing, externalizing, and total scores (Table 2). There were no significant correlations between the BITSEA/C scores and the CBCL scores.

2.2.2. Construct validity, BITSEA subscores, and cutpoints

The BITSEA problem cutpoints were identified by using subclinical/clinical CBCL/2–3 cutpoint scores. Fig. 1 illustrates the ROC curves for maternal BITSEA/P scores based on subclinical and clinical CBCL/2–3 cutpoints.

These BITSEA/P cutpoints revealed that 13.1% of male toddlers' and 17.6% of females' maternal BITSEA/P scores were higher than the clinical cutpoint and 30.6% of males' and 28.6% of females' maternal BITSEA/P scores were higher than the subclinical cutpoint. Similarly, 13.5% of male toddlers' and 17.6% of females' paternal BITSEA/P scores were higher than the clinical cutpoint and 31.0% of males' and 22.9% of females' paternal BITSEA/P scores were higher than subclinical cutpoint. Table 3 presents the identified cutpoints of maternal BITSEA/P scores in gender groups by using subclinical/clinical CBCL/2–3 cutpoint scores.

Table 2
Correlations between subscores of BITSEA and CBCL/2–3.

Subscale	1	2	3	4	5
1. BITSEA/P (maternal)	–	0.69* (427)	0.52* (322)	0.59* (318)	0.69* (316)
2. BITSEA–P (paternal)		–	0.46* (311)	0.58* (308)	0.65* (307)
3. CBCL–internalizing			–	0.47* (329)	[a]
4. CBCL–externalizing				–	[a]
5. CBCL–total					–

* $p < 0.001$ (Spearman's correlations).

[a] CBCL total scores comprise internalizing and externalizing scores, that is why some correlations are not shown in the table. Spearman's rho scores are shown in the table. Numbers in the parenthesis indicate number of cases. BITSEA/P: brief infant and toddler social and emotional assessment scale–problem subscale; CBCL: child behavior checklist: 2–3.

Table 3
Identified cutpoints of maternal BITSEA/P scores in gender groups using subclinical/clinical CBCL/2–3 scores.

Gender	Subclinical			Clinical		
	Cutpoint score	Sensitivity (%)	Specificity (%)	Cutpoint score	Sensitivity (%)	Specificity (%)
Male	≥18	87	73	≥24	83	91
Female	≥21	95	80	≥24	90	88

3. Discussion

Overall results support that the Turkish version of the BITSEA is a valid and reliable screening tool for social and emotional problems of toddlers in primary health care settings. As 70% of parents returned the scales to the clinic and only 3.6% of mothers and 1.9% of fathers had missing values in more than two items, we conclude that the Turkish translation is quite understandable and may be easily used in primary health care settings. Internal consistency of BITSEA/P and BITSEA/C may be assumed excellent. Interrater reliability of both scales were good. Interrater reliability, which was examined as percent agreement on the identified cutpoints of BITSEA/P scores revealed significant agreement. Test–retest reliability of paternal BITSEA/P was excellent and maternal BITSEA/P, paternal and maternal BITSEA/C were fair to good. This relatively lower test–retest reliability outcomes could be associated with the instructions in the scale that the items should be scored for last month. As the symptoms could change in 1-month period, the correlations would be relatively low.

Criterion-related validity of BITSEA/P relative to both subscores and total score of CBCL/2–3 were good. In examination of the cut-points, defined by the ROC curve relative to the CBCL for the BITSEA/P for mothers, 30.6% of boys and 28.6% of girls were in the subclinical level. Similarly, in the validation study of the original BITSEA, Briggs-Gowan et al. (2004) reported that approximately 25% of both male and female toddlers had BITSEA/P scores higher than empirically derived cutpoints. Our results support the assumption that the BITSEA is likely to identify about one in three children as at-risk for problems or delays in competence (Briggs-Gowan et al., 2004). The at-risk group would likely include children with problems that may be precursors to psychopathology and children whose parents have distorted perceptions of child functioning, as may occur with parental depression (Briggs-Gowan, Carter, & Schwab-Stone, 1996). Thus, as a first-stage screener, the BITSEA will likely identify a clinically diverse group.

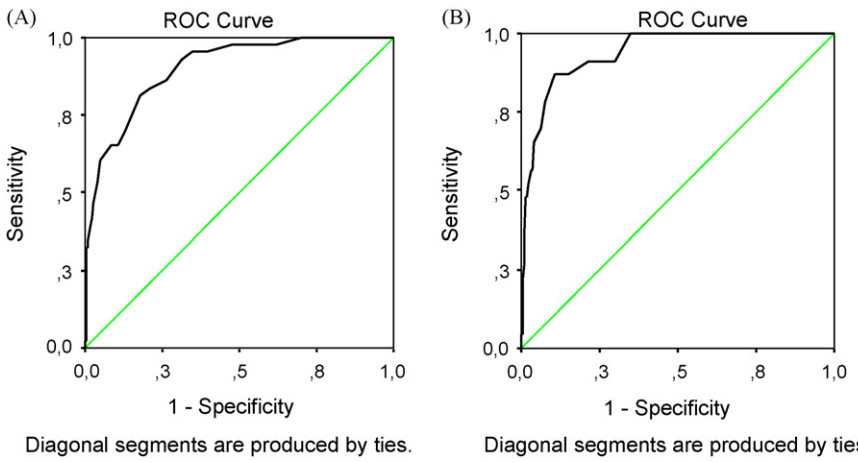

Fig. 1. The ROC curves for maternal BITSEA/problem scores based on subclinical (A) and clinical (B) CBCL/2–3 cutpoints (A: $p < 0.001$, AUC: 0.182; B: $p < 0.001$, AUC: 0.151).

We noticed that, while mean scores of BITSEA/C in Turkish version were found to be relatively similar to the mean scores in the BITSEA/C English version (Briggs-Gowan et al., 2004), BITSEA/P scores were found relatively higher in the Turkish children. As all items are scored from 0 to 2 and would be higher in accordance to the severity of the symptom (in BITSEA/P) and lower in accordance to the problem in competence (in BITSEA/C), this difference may be assumed as a result of language or cultural differences. For instance, "severity" and/or "frequency" may be conceptualized with different words in different languages. Consistent with this finding and in accordance with this assumption, in a national prevalance study in Turkey, total problem scores of CBCL/2–3 were found to be higher in Turkish children (mean: 39.5 ± 21.4) than those of the European and North American children in the same age group (Erol et al., 2005). Nevertheless, these finding may indicate a real difference in the level of problems in Turkish children. However, as in our study, when it is applied in primary health care units, the BITSEA may behave slightly differently when used in pediatric or early intervention samples, which may have greater proportions of children with delays or health problems. Higher CBCL/2–3 total scores (45.09 ± 23.02) in our findings with respect to the previous prevalance study (Erol et al., 2005) support this assumption.

Routine screening in pediatric settings has been recommended to enhance efforts to identify early social-emotional problems (AAP, 2001; Eisert, Sturner, & Mabe, 1991; Thompson, 1985). The pediatric setting is particularly well suited to the task of detecting social-emotional and behavioral problems (AAP, 2001). In addition, like their counterparts in many countries, general practitioners and pediatricians in Turkey are employed in preventive health services such as vaccination clinics. These settings often serve as gatekeepers for specialty mental health services in managed care systems. The BITSEA may be a valuable screening tool to be used in routine assessment in these centers. It is important that pediatricians and other service providers follow up on positive BITSEA scores, by engaging parents in a dialogue about children's difficulties (and strengths) and determining how much the reported behaviors interfere with children's developmental progress and families' day-to-day life (i.e., the extent to which these behaviors are associated with impairment) (Briggs-Gowan et al., 2004). The cutpoints identified in our study appear to be appropriate to apply to a Turkish population and these cuts might not be generalized to other populations or other language translations.

Another significant finding in our study is the higher scores of BITSEA/C in girls with respect to boys between 24 and 29 months. This result is in accordance with the significantly higher mean scores of BITSEA/C in girls with respect to boys in the original validation study of BITSEA ($p < 0.01$) (Briggs-Gowan et al., 2004). Similar to our results, the gender difference in BITSEA/C scores is most prominent in the 24–29-month-old group.

Screening instruments intended for use in pediatric settings should be brief and easy to administer, score, and interpret (Jellinek & Murphy, 1988). They also should have adequate reliability and validity (Eisert et al., 1991; Jellinek & Murphy, 1991) and identify an acceptable percentage (a minimum of 70%) of children who have problems, yet have a false-positive rate of not greater than 30% (Cicchetti, Volkmar, Klin, & Showalter, 1995). Finally, those instruments should provide developmentally appropriate (Glascoe, 2000) and clinically useful information (Carter, 2002). Our results indicate that Turkish version of BITSEA is easy to administer, score, and interpret. It has adequate reliability and validity and it identifies 87–95% of children who have subclinical social, emotional and behavioral problems with a false-positive rate of 20–27%.

3.1. Limitations

The participants in this study were selected randomly and in a stratified design based on population rates. However, as the setting is the primary health care centers, the mean scores of BITSEA should not be assumed valid for a community sample of toddlers. In addition, as 30% of the parents did not return the forms back to the clinic, it was not possible to compare them with the participants, in terms of sociodemographic variables and scale scores. In this study, the validity of BITSEA/C has not been examined. In addition to further validation of the BITSEA/C, as this study did not involve toddlers with severe psychiatric disorders, such as pervasive developmental disorders or disruptive behavior disorders, clinical validation studies are warranted for identification of valid cutpoints.

3.2. Conclusion and clinical implications

This study reveals that Turkish version of BITSEA is a reliable and valid instrument for detecting social, emotional and behavioral problems in toddlers. As it consists of 42 items and is easily scored, it is a promising screening tool for toddlers in primary health care settings. Unlike other instruments that measure social and emotional problems and delays in competence in toddlerhood, other research concerning the BITSEA has indicated that it is a promising screening tool for autistic spectrum disorders as well. Therefore, it is important to establish the BITSEA's clinical validity and to develop cutpoints that reflect clinically significant problems.

References

Achenbach, T., & Rescorla, L. (2000). *Manual for the ASEBA preschool forms and profiles*. Burlington: University of Vermont.

American Academy of Pediatrics [AAP]. (2001). Committee on children with disabilities: Developmental surveillance and screening of infants and young children. *Pediatrics*, *108*, 192–196.

Anderson, G. M., & Cohen, D. J. (2002). Genesis of neocortex. In M. Lewis (Ed.), *Child and Adolescent Psychiatry: A Comprehensive Textbook* (pp. 22–46). Philadelphia, PA: Lippincott Williams & Wilkins.

Baird, G., Charman, T., Baron-Cohen, S., Cox, A., Swettenham, J., Wheelwright, B. A., et al. (2000). A screening instrument for autism at 19 months of age: A 6-year follow-up study. *Journal of Abnormal Child and Adolescent Psychiatry*, *39*(6), 694–702.

Briggs-Gowan, M. J., & Carter, A. S. (2002). *Brief Infant–Toddler Social and Emotional Assessment (BITSEA) manual, version 2.0*. New Haven, CT: Yale University.

Briggs-Gowan, M. J., Carter, A. S., Irwin, J. R., Wachtel, K., & Cicchetti, D. V. (2004). The Brief Infant–Toddler Social and Emotional Assessment: Screening for social-emotional problems and delays in competence. *Journal of Pediatric Psychology*, *29*(2), 143–155.

Briggs-Gowan, M. J., Carter, A. S., & Schwab-Stone, M. (1996). Discrepancies among mother, child, and teacher reports: Examining the contributions of maternal depression and anxiety. *Journal of Abnormal Psychology*, *24*, 749–765.

Carter, A. S. (2002). Assessing social-emotional and behavior problems and competencies in infancy and toddlerhood: Available instruments and directions for Application. In B. Zuckerman, A. Lieberman, & N. Fox (Eds.), *Emotion regulation and developmental health: Infancy and early childhood* (pp. 277–299). New York: Johnson & Johnson Pediatric Institute.

Carter, A. S., Briggs-Gowan, M. J., Jones, S. M., & Little, T. (2003). The Infant–Toddler Social and Emotional Assessment (ITSEA): Factor structure, reliability, and validity. *Journal of Abnormal Child Psychology*, *31*(5), 495–514.

Cicchetti, D. V., Volkmar, F., Klin, A., & Showalter, D. (1995). Diagnosing autism using ICD-10 criteria: A comparison of neural networks and standard multivariate procedures. *Child Neuropsychology*, *1*, 26–37.

Eisert, D. C., Sturner, R. A., & Mabe, P. A. (1991). Questionnaires in behavioral pediatrics: Guidelines for selection and use. *Journal of Developmental and Behavioral Pediatrics*, *12*, 42–50.

Erol, N., & Simsek, Z. (1997). Türkiye Ruh Sağlığı Profili: Çocuk ve gençlerde yeterlik alanları ile sorun davranışlarının dağılımı. In N. Erol, C. Kılıç, M. Ulusoy, M. Keçeci, & Z. Şimşek (Eds.), *Türkiye Ruh Sağlığı Profili: Ön Rapor* (pp. 12–33). Ankara: Mental Health Profile of Turkey Ministry of Health, Aydoğdu Ofset.

Erol, N., Şimşek, Z., Öner, Ö., & Münir, K. (2005). Behavioral and emotional problems among Turkish children at ages 2 to 3 years. *Journal of American Academy of Child and Adolescent Psychiatry*, *44*(1), 80–87.

Glascoe, F. P. (2000). Early detection of developmental and behavioral problems. *Pediatric Review*, *21*, 272–280.

Horwitz, S. M., Gary, L. A., Briggs-Gowan, M. J., & Carter, A. S. (2003). Do needs drive services or do services drive needs? *Pediatrics*, *112*, 1373–1378.

Jellinek, M. S., & Murphy, J. M. (1988). Screening for psychosocial disorders in pediatric practice. *American Journal of Diseases in Children*, *142*, 1153–1157.

Jellinek, M. S., Murphy, J. M., Little, M., Pagano, M. E., Comer, D. M., & Kelleher, K. J. (1999). Use of the pediatric symptom checklist to screen for psychosocial problems in pediatric primary care: A national feasibility study. *Archives of Pediatrics and Adolescent Medicine*, *153*, 254–260.

Keren, M., Feldman, R., & Tyano, S. (2003). A five-year Israeli experience with the DC: 0–3 classification system. *Infant Mental Health Journal*, *24*, 337–348.

Lieberman, A. F., Van Horn, P., Grandison, C., & Pekarsky, J. (1997). Mental health assessment of infants, toddlers and preschoolers in a service program and treatment outcome research program. *Infant Mental Health Journal*, *18*, 158–170.

Meisels, S. J., & Fenichel, E. (1996). *New visions for the developmental assessment of infants and young children*. Washington, DC: ZERO-TO-THREE.

Murphy, J. M., Ichinose, C., Hicks, R. C., Kingdon, D., Crist-Whitzel, J., Jordan, P., et al. (1996). Utility of the pediatric symptom checklist as a psychosocial screen to meet the federal early and periodic screening, diagnosis, and treatment (EPSDT) standards: A pilot study. *Journal of Pediatrics*, *129*, 864–869.

Roberts, R. E., Attkisson, C. C., & Rosenblatt, A. (1998). Prevalence of psychopathology among children and adolescents. *American Journal of Psychiatry*, *155*, 715–725.

Shaw, D. S., Owens, E. B., Giovanelli, J., & Winslow, E. B. (2001). Infant and toddler pathways leading to early externalizing disorders. *Journal of American Academy of Child and Adolescent Psychiatry*, *40*, 36–43.

Thomas, J. M., Benham, A. L., Gean, M., Luby, J., Minde, K., Turner, S., et al. (1997). Practice parameters for the psychiatric assessment of infants and toddlers (0–36 months). *Journal of American Academy of Child and Adolescent Psychiatry*, *36*(10 Suppl.), 21S–36S.

Thomas, J. M., & Guskin, K. A. (2001). Disruptive behavior in young children: What does it mean? *Journal of American Academy of Child and Adolescent Psychiatry*, *40*, 44–51.

Thompson, R. J., Jr. (1985). Delineation of children's behavior problems: A basis for assessment and intervention. *Journal of Developmental and Behavioral Pediatrics*, *6*, 37–50.