

# The Brief Infant-Toddler Social and Emotional Assessment: Screening for Social-Emotional Problems and Delays in Competence

Margaret J. Briggs-Gowan,<sup>1</sup> Alice S. Carter,<sup>2</sup> Julia R. Irwin,<sup>3</sup> Karen Wachtel,<sup>4</sup>  
and Domenic V. Cicchetti<sup>5</sup>

<sup>1</sup>Department of Epidemiology and Public Health, Yale University, <sup>2</sup>Department of Psychology, University of Massachusetts, Boston, <sup>3</sup>Haskins Laboratory, New Haven, Connecticut, <sup>4</sup>Department of Psychology, University of Massachusetts, Boston, and <sup>5</sup>Child Study Center, Yale University

**Objective** To examine the reliability and validity of the 42-item Brief Infant-Toddler Social and Emotional Assessment (BITSEA), a screener for social-emotional/behavioral problems and delays in competence. **Method** Parents in a representative healthy birth cohort of 1,237 infants aged 12 to 36 months completed the Infant-Toddler Social and Emotional Assessment (ITSEA)/BITSEA, the Child Behavior Checklist (CBCL)/1.5-5, the MacArthur Communication Developmental Inventory vocabulary checklist, and worry questions. In a subsample, independent evaluators rated infant-toddler behavior. **Results** Test-retest reliability was excellent and interrater agreement (mother/father and parent/child-care provider) was good. Supporting validity, BITSEA *problems* correlated with concurrent evaluator problem ratings and CBCL/1.5-5 scores and also predicted CBCL/1.5-5 and ITSEA problem scores one year later. BITSEA measures of competence correlated with concurrent observed competence and predicted later ITSEA competence measures. Supporting discriminant validity, only 23% of high BITSEA *problem* scorers had delayed vocabulary. Moreover, the combined BITSEA problem/competence cutpoints identified 85% of subclinical/clinical CBCL/1.5-5 scores, while maintaining acceptable specificity (75%). **Conclusions** Findings support the BITSEA as a screener for social-emotional/behavioral problems and delays in social-emotional competence.

**Key words** screener; social-emotional problems; behavior problems; competence; infant-toddler.

There is increasing recognition of the importance of early detection and provision of intervention services for infants and toddlers with significant social-emotional and/or behavioral problems (AAP, 2001; U.S. Public Health Service, 2000). Part C of the Individuals with Disabilities Education Act (IDEA) amendments of 1997 (Public Law No. 105-17) mandates intervention services for infants and toddlers with delays in social development and provides discretionary services for children with social-emotional/behavioral problems that may place them at risk for later delay. However, early identification of and service provision to infants and toddlers

with social-emotional/behavioral problems have lagged far behind advances in public policy, as well as 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 (Briggs-Gowan, Carter, Skuban, & Horwitz, 2001; Roberts, Attkisson, & Rosenblatt, 1998). Yet, a recent study of a representative sample indicated that fewer than 8% of 1- and 2-year-olds with social-emotional/behavioral problems received any developmental or mental health services (Horwitz, Gary, Briggs-Gowan, & Carter, 2003). Moreover, the

All correspondence concerning this article should be addressed to Margaret Briggs-Gowan, Epidemiology and Public Health, PO Box 208034, Yale University, New Haven, Connecticut, 06520. E-mail: [margaret.briggs-gowan@yale.edu](mailto:margaret.briggs-gowan@yale.edu).

*Journal of Pediatric Psychology* 29(2) pp. 143–155, 2004

DOI: 10.1093/jpepsy/jsh017

*Journal of Pediatric Psychology* vol. 29 no. 2 © Society of Pediatric Psychology 2004; all rights reserved

services that were provided generally focused on developmental issues, such as speech/language, motor, and cognitive delays. Further, early identification of delays in social-emotional competence may be especially important, as competence may play a key role in the longitudinal course of early emotional/behavioral problems (Cicchetti, 1993; Masten & Coatsworth, 1995).

Routine screening in pediatric settings has been recommended to enhance efforts to identify early social-emotional problems (AAP, 2001; Eisert, Sturmer, & Mabe, 1991; Thompson, 1985). The pediatric setting is particularly well suited to the task of detecting social-emotional and behavioral problems (AAP, 2001; Regier, Goldberg, & Taube, 1978). In addition to frequent and routine contact with young children, pediatricians are especially likely to see children with emotional/behavioral problems, because these children tend to visit their pediatricians more frequently than do other children (Zuckerman, Moore, & Glei, 1996). Moreover, pediatricians often serve as gatekeepers for specialty mental health services in managed care systems (Costello et al., 1988; Forrest et al., 1999).

However, shorter office visits under managed care may reduce the likelihood that parents will raise concerns about child behavior (Blumenthal et al., 1999; Horwitz, Leaf, Leventhal, Forsyth, & Speechley, 1992) or that pediatricians will have the opportunity to elicit such concerns or observe problem behaviors. Indeed, although most parents believe that it is appropriate to discuss behavioral/emotional issues with their pediatricians, many do not actually do so when problems exist (Horwitz et al., 1992; Horwitz, Leaf, & Leventhal, 1998; Dulcan et al., 1990). In addition, infant-toddler behavior during an office visit may not be representative of behavior in other settings. For example, the unfamiliar setting or negative prior medical experiences may affect child behavior, resulting in heightened stranger anxiety, negative affect, dampened affect, and/or restricted affective range.

Routine screenings at well-child visits may address these barriers to identification, providing a time-efficient and cost-effective method for bringing possible social-emotional and behavioral issues to the attention of the pediatrician and opening a dialogue between pediatricians and parents (Carter, 2002). Most parents can complete checklists independently, requiring minimal staff time (Glascoe, 2000). Further, computers may be used to administer checklists to parents and provide scores directly to pediatric staff (Carter, 2002). When employed with school-age children, pediatric screening

has been shown to be feasible (Baird et al., 2000; Jellinek et al., 1999) and effective in improving rates of referral for mental health services (Murphy et al., 1996). However, a lack of age-appropriate measures has hindered screening efforts with infants and toddlers (Stancin & Palermo, 1997).

Screeners 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, 1988) and should identify an acceptable percentage (a minimum of 70%) of children who have problems, yet have a false-positive rate of no greater than 30% (Cicchetti, Volkmar, Klin, & Showalter, 1995). Finally, screeners should provide developmentally appropriate (Glascoe, 2000) and clinically useful information (Carter, 2002).

In the infant-toddler period, comprehensive measures appropriate for social-emotional and behavioral problem screening are limited (Glascoe, 2000). Although the Child Behavior Checklist (CBCL)/1.5-5 (Achenbach & Rescorla, 2000) and the Infant-Toddler Social and Emotional Assessment (ITSEA) (Carter & Briggs-Gowan, 2000) have good reliability and validity, they are too long to employ in screening. Two existing brief screeners (Eyberg, 1980; Mouton-Simien, McCain, & Kelley, 1997) do not address both problems and competencies. One of these, the 35-item Eyberg Child Behavior Inventory (Eyberg, 1980), focuses on conduct problems in children aged 2 to 11 years and has demonstrated acceptable reliability and validity. The 40-item Toddler Behavior Screening Inventory (TBSI) (Mouton-Simien et al., 1997) addresses common behavior problems in 1- to 3-year-olds. The TBSI recently demonstrated acceptable sensitivity in detecting children whose pediatrician had referred them for psychological services (McCain, Kelley, & Fishbein, 1999). However, given that pediatricians often underidentify children's mental health problems (e.g., Horwitz et al., 1992), it would be beneficial to evaluate the TBSI's validity in detecting problems in children who were not already referred by their pediatrician. Further, neither measure addresses social-emotional competencies or behaviors that are typical of autism spectrum disorders (e.g., poor social relatedness, social-withdrawal, repetitive behaviors).

Comparatively, more screeners for autism spectrum disorders have been developed. Existing measures, such as the Checklist for Autism in Toddlers (CHAT) (Baird et al., 2000) and the Modified-CHAT (Robins, Fein,

Barton, & Green, 2001), are promising with respect to early identification of children with autism but do not provide adequate coverage of the range of social-emotional problems evident in very young children (e.g., aggression, sleep disruptions).

One promising screener, the Ages & Stages Questionnaire–Social-Emotional version (ASQ-SE) (Squires, Bricker, & Twombly, 2002a), addresses social-emotional and behavioral problems from birth to 5 years of age. The ASQ-SE includes social competencies, including behaviors that when absent may indicate the presence of autism spectrum disorders (e.g., behaviors necessary to maintain social interactions). Acceptable test-retest reliability and sensitivity in detecting children with developmental delay and/or social-emotional diagnoses has been reported (Squires, Bricker, & Twombly, 2002b). However, more information is needed concerning its sensitivity to specific types of social-emotional disorders and its validity across the entire 5-year age span.

In summary, there remains a need for a measure that is sensitive to social-emotional/behavioral problems, autism spectrum disorders, and delays in social-emotional competence in early childhood. A new screener for identifying social-emotional/behavioral problems and delays in social-emotional competence in 12- to 36-month-olds, the Brief Infant-Toddler Social and Emotional Assessment (BITSEA) (Briggs-Gowan & Carter, 2002), was designed to address this need. The current research examined the reliability and validity of the BITSEA. It was hypothesized that the BITSEA would (1) have acceptable test-retest and interrater reliability and (2) demonstrate acceptable construct-related, predictive, and discriminant validity, relative to a more lengthy measure of social-emotional/behavioral problems and independent ratings of child problems and competencies. Analyses also evaluated the BITSEA's performance relative to the longer ITSEA from which it is derived.

## Methods

### Participants

Subjects were recruited from a birth cohort that was age- and sex-stratified and randomly selected from birth records at the State of Connecticut Department of Public Health for children who were born at Yale–New Haven Hospital and lived in the 1990 Census New Haven–Meriden Standard Metropolitan Statistical Area ( $N = 8,404$ ) (see Briggs-Gowan et al., 2001). A total of 1,788 children were randomly selected from 7,433 children

who were eligible for sampling after exclusion criteria were applied (i.e., likely to have significant developmental delays, based on birth record information; deceased; adopted; or the sibling of a child who was already sampled). Eligible families did not differ significantly from original subjects ( $N = 8,404$ ) in maternal education or child race, but tended to have slightly greater birth weights and gestational ages and slightly older mothers. Additional subjects ( $n = 183$ ) were excluded if no parent could participate in English, the family had moved out of state, or the mother lost custody of the child. These excluded families did not differ from remaining families ( $N = 1,605$ ) in maternal education, child race, 1-minute APGAR scores, gestational age, or birth weight. Compared with noneligible families, eligible families had slightly higher maternal age ( $M = 29.2$ ,  $SD = 6.2$  vs.  $M = 27.5$ ,  $SD = 5.4$ ,  $t = 4.07$ ,  $p < .01$ ) and 5-minute APGAR scores ( $M = 8.96$ ,  $SD = .39$  vs.  $M = 8.89$ ,  $SD = .47$ ,  $t = 2.04$ ,  $p < .05$ ).

A total of 1,280 of 1,605 eligible parents participated (80% response rate). Participants tended to be slightly more educated and older and to have slightly heavier newborns than nonparticipants,  $p < .01$ . Although there was a lower response among families of minority groups (70%) than among white families (84%), ( $p < .01$ ),  $t$ -tests and chi-square analyses indicated that minority-group respondents did not differ from minority-group nonrespondents on any birth status or sociodemographic variable ( $p > .05$ ). Analyses included 1,237 parents and excluded 4% of the sample whose children were beyond the BITSEA's age range ( $>36$  months) at the time of participation.

Most respondents (96.2%) were biological mothers. Mean child age was 23.8 months ( $SD = 6.8$ ), with 47.9% 1-year-olds and 52.1% 2-year-olds. Forty-nine percent of the children were boys. Most children (81.5%) came from two-parent homes (71.7% married). The sample was sociodemographically heterogeneous in terms of respondent education (8.0% less than high school, 18.3% high school diploma/GED, 32.3% education beyond high school, and 41.4% bachelor's degree or higher), and ethnicity (66.3% non-Hispanic white, 16.2% African American/black, 5.3% Hispanic, 8.3% multiracial minority, 2.3% Asian, and 1.6% other). Further, 18.1% of families had incomes below the federal poverty level, and an additional 15.8% lived in borderline poverty (income  $<185\%$  of poverty level). Median annual before-tax income was \$50,600. About two thirds (64.7%) of mothers were employed full- or part-time. Participants were sociodemographically

comparable to individuals living in New Haven County. The following statistics have been reported for this region: 10% poor; \$48,834 median yearly income; 78% non-Hispanic white; 11% African American/black; 88% high school education or more; and 73% marriage rate for families with children (FedStats, 2002; NCES, 2001).

Families retained in the one-year follow-up ( $n = 1,135$  of 1,237, 91.8%) did not differ from nonretained families in child sex, minority status, or marital status, but were more likely to have higher respondent education (75.1% having at least a high school education vs. 58.8%,  $p < .01$ ) and to be nonpoor (67.6% vs. 50.5%,  $p < .01$ ).

## Measures

### Sociodemographic Variables

Parents answered questions about sociodemographic factors including child sex, age, ethnicity, birth order, maternal age, parental education, marital status, before-tax household income, and numbers of adults and children in the home.

### Infant-Toddler Social and Emotional Assessment

The survey contained the 169-item ITSEA (Carter & Briggs-Gowan, 2000; Carter, Briggs-Gowan, Jones & Little, 2003), which includes measures of internalizing, externalizing, dysregulation, and competence and three indices (social relatedness, atypical behaviors, and maladaptive) that consist of low base rate, clinically significant behaviors. Items are rated on a 3-point scale (0 = *not true/rarely*, 1 = *somewhat true/sometimes*, 2 = *very true/often*). Reliability and validity have been evaluated in three studies. First, in an ethnically and educationally heterogeneous pediatric sample, the ITSEA demonstrated acceptable internal consistency, test-retest reliability, and validity relative to other parent-report checklists (Briggs-Gowan & Carter, 1998). Second, the ITSEA showed validity relative to independent observational ratings of child behavior in a sample of toddlers (Carter, Little, Briggs-Gowan & Kogan, 1999). Most recently, results from the sample used in this report (Carter et al., 2003) indicated acceptable internal consistency (Cronbach's  $\alpha = 0.80$ – $0.90$ ) and test-retest reliability (intraclass correlation =  $0.82$  to  $0.90$ ). Criterion-related validity was documented relative to parent reports on the CBCL/1.5-5 (e.g., same-domain correlations of  $0.57$  and  $0.73$ ), independent ratings of child behavior ( $r = 0.20$  to  $0.31$ ), and, for *competence*, standardized developmental tests ( $r = 0.39$  to  $0.58$ ). Statistical cutpoints for domains have been set at the 90th percentile (Carter et al., in press).

### Brief Infant-Toddler Social and Emotional Assessment

The 42-item BITSEA (Briggs-Gowan & Carter, 2002) is designed as a screener for parents and child-care providers to identify children “at risk” for or currently experiencing social-emotional/behavioral problems and/or delays in social-emotional competence, including autism spectrum disorders.

BITSEA items were drawn from the pool of ITSEA questions. Clinical and empirical considerations informed item selection. An item was selected if (1) the majority of a panel of 12 infant mental health experts rated it as clinically important to include in a screener and/or (2) the item had the highest loading on an ITSEA scale. Twenty-eight BITSEA items were rated as clinically important by the majority of expert clinicians. Of the 14 remaining items, 12 had the highest ITSEA loading, 1 was selected as the most broadly representative of the ITSEA prosocial-peer parameter (*Plays well with other children, not including brother/sister*), and 1 was included due to its clinical significance (*Hurts him/herself on purpose. For example, bangs his or her head*). To minimize the BITSEA's length, two composite items (*Is afraid of certain places, animals, or things* and *Seems very unhappy, sad, depressed, or withdrawn*) originated from more than one ITSEA item. In analyses, composite items were represented by the maximum response on the original items.

The following numbers of BITSEA items were drawn from each ITSEA area: internalizing, eight; externalizing, six; dysregulation, eight; competence, seven; social relatedness, three; maladaptive, three; and atypical, four. Three clinically significant items that are part of the ITSEA, but not on any domain or index, were included. ITSEA inhibition is considered more a dimension of temperament than of psychopathology (Carter et al., 2003). Thus, the BITSEA does not include inhibition. Sample items are:

- Is restless and can't sit still.
- Hits, bites, or kicks you.
- Does not make eye contact.
- Has less fun than other children.
- Refuses to eat.
- Wakes up at night and needs help to fall asleep again.
- Cries or throws tantrums until exhausted.
- Is affectionate with loved ones.
- Follows rules.
- Looks for you (or other parent) when upset.
- Hugs or feeds dolls or stuffed animals.

The BITSEA requires a fourth- to sixth-grade reading level and can be completed in approximately 5 to 7 minutes. Scores are calculated as sums and can be computed by hand (requiring about 5 minutes) or with a pilot computer scoring program (requiring approximately 3 minutes) that also provides a score profile. This is the first published report on the BITSEA.

#### Child Behavior Checklist for 1.5-5

The CBCL/1.5-5 (Achenbach & Rescorla, 2000), designed for children 18 months through 5 years, consists of three domains (internalizing, externalizing, and total problem). “Subclinical” ( $t$ -score  $\geq 60$ ) and “clinical” ( $t$ -score  $\geq 63$ ) cutpoints have been developed. The CBCL/1.5-5 has demonstrated very good 8-day test-retest reliability ( $r = .68$  to  $.92$ , mean  $r = .84$ ) and cross-informant agreement (mean mother-father  $r = .61$ , mean parent-child care provider  $r = .65$ ). Validity results, from a sample from mental health and special education facilities and matched subjects from a normative sample, indicated significant effects of referral status on CBCL scores. Children with subclinical/clinical internalizing and externalizing scores were five and six times more likely, respectively, to be clinically referred than children with lower scores. The internalizing and externalizing scales correctly classified 74% of referred children.

#### MacArthur Communicative Development Inventory–Short Form

The MCDI-SF (Dale, Reznick, & Thal, 1998; Fenson et al., 2000) is a parent-report vocabulary checklist developed from the longer MCDI, which has shown excellent reliability and good validity relative to standardized assessments (Fenson et al., 1993). Three age-based levels of the MCDI-SF were used (12–17 months, 18–29 months,  $>30$  months). The two younger levels have correlated highly with the full MCDI ( $r = 0.97, 0.98$ ) and have excellent reliability ( $r = 0.97$ – $0.99$ ) (Fenson et al., 2000). The oldest level has correlated moderately ( $r = 0.63$ ) with standardized language assessments (Dale, Reznick, Thal, & Newton, 2000). Scores below the 10th percentile by age and sex reflect delayed productive vocabulary.

#### Parental Worry

On a 5-point scale from 1 = *not at all worried* to 5 = *extremely worried*, parents rated their concerns for their child’s social and emotional development, behavior, and language.

#### Evaluator Ratings

Following home visits, evaluators rated children’s social–emotional problems and competencies in categories designed to generally parallel ITSEA scales. Problem ratings were made on a 4-point scale (0 = *no problem*, 1 = *possible*, 2 = *probable*, 3 = *definite problem*). Evaluators rated the following problem items: inhibition/extreme shyness; anxiety/fears; depression/social withdrawal; aggression/defiance; aggression with peers; overactivity; behavior problems; sleep; eating; negative emotionality; and unusual sensory sensitivity. Competencies were rated on a 7-point scale from 1 = *definite strength* to 7 = *definite problem* and included these items: compliance; attention skills; empathy; persistence/enjoyment of challenging activities; imitation/pretend play; prosocial peer interactions; and awareness of others’ emotions. Scales were calculated as sums, and had very good to excellent internal consistency (Cronbach’s  $\alpha$ : problems = 0.87, competencies = 0.91). As reported by Carter et al. (2003), acceptable interrater reliability was observed between the research team’s ratings and ratings by early intervention providers in a separate study.

#### Design and Procedure

From June to September of 1998, parents were mailed a letter describing the study, followed one week later by a questionnaire and children’s book. Staff contacted parents via telephone and in person to encourage participation. Whenever a parent declined, staff discontinued efforts to obtain participation. Informed consent procedures were followed and subjects were notified that they would be invited to participate in one or more surveys in the future. Parents received \$25 for participating. One year later, a follow-up survey, employing identical procedures to those of Year 1, was conducted with all Year 1 participants.

#### Methodologic Substudy

After participating in the initial survey, 173 parents and children participated in a home visit substudy. Eligibility required use of 15 or more hours per week of child care, in order to allow the evaluation of parent/child-care-provider agreement on ITSEA ratings. Participants were selected randomly from the pool of eligible subjects. The test-retest, mother-father, and parent/provider data did not differ significantly ( $p < .05$ ) from the overall sample in terms of child age, child sex, marital status, ethnic minority status, respondent education, or poverty. The home visit included a videotaped developmental

evaluation of the child, a parent interview about adaptive behavior, and completion of an ITSEA retest questionnaire. Following home visits, evaluators who were blind to children's ITSEA status made independent ratings of child behavior, based on observations during the home visit. Retest analyses included 119 parents who completed the initial and retest questionnaires within 10 to 45 days ( $M = 26.4$ ,  $SD = 8.14$ ). With parental permission, second parents and child care providers were invited to complete ITSEA questionnaires. Data for 68 mother-father pairs who completed the questionnaires within a 10- to 45-day interval were included in analyses ( $M = 27.9$  days,  $SD = 9.6$ ). Finally, 79 parent/child-care-provider pairs who completed the questionnaires within a 15- to 59-day period were included ( $M = 39.6$ ,  $SD = 10.1$ ). The longer time interval for child care providers reflects their greater delay in participation relative to parents.

## Results

### Analytic Plan

The scale structure, test-retest reliability, interrater reliability, and one-year stability of the BITSEA were examined. In addition, to inform the assignment of statistically at-risk cutpoints, BITSEA scales were evaluated for age and sex effects. Validity was examined dimensionally and through dichotomous sensitivity-specificity analyses. Criterion-related validity was evaluated by comparing the BITSEA with the CBCL/1.5-5 and independent evaluator ratings. Discriminant validity was assessed by comparing BITSEA cutpoint status with MCDI vocabulary scores. Predictive validity was examined by comparing Year 1 BITSEA scores with Year 2 ITSEA and CBCL/1.5-5 scores. Finally, correlational and sensitivity-specificity analyses were employed to assess the performance of the BITSEA relative to the ITSEA. Most analyses used full birth cohort data. Substudy sample data were used to assess interrater reliability, test-retest reliability, and correlations with evaluator ratings.

### Scale Structure

Internal consistency was acceptable for the BITSEA *problem* scale (BITSEA/P) ( $\alpha = 0.79$ ,  $n = 209$ ) and marginal for BITSEA *competence* (BITSEA/C) ( $\alpha = 0.65$ ,  $n = 1,233$ ), employing criteria recommended by Cicchetti et al. (1995). BITSEA/P item loadings ranged from 0.14 to 0.50 ( $M = 0.30$ ) and BITSEA/C loadings ranged from 0.20 to 0.38 ( $M = 0.30$ ). Low base rates for some items likely contributed to low item loadings. Lower internal

consistency for *competence* is expected because items in that scale address a range of behaviors that may not be expected to co-occur. Similarly, the child care provider BITSEA had acceptable internal consistency for the BITSEA/P ( $\alpha = 0.80$ ,  $n = 100$ ), but marginal internal consistency for the BITSEA/C ( $\alpha = 0.66$ ,  $n = 117$ ).

### Test-Retest and Interrater Reliability

Ten- to 45-day test-retest reliability was excellent ( $n = 119$ , BITSEA/P intraclass correlation coefficient = 0.87, BITSEA/C intraclass correlation coefficient = 0.85) (Cicchetti & Sparrow, 1981). Agreement between parents was quite good ( $n = 68$ , BITSEA/P intraclass correlation coefficient = 0.68, BITSEA/C intraclass correlation coefficient = 0.61). Relative to typical parent-teacher agreement (Achenbach, McConaughy, & Howell, 1987), the parent/child-care-provider correlation ( $n = 79$ ) was higher than expected for *competence* (intraclass correlation coefficient = 0.59) and typical for *problems* (intraclass correlation coefficient = 0.28).

### 1-Year Stability

The 1-year stability of the BITSEA, examined in 1,112 families, was  $r = 0.65$  for problems and  $r = 0.53$  for competence ( $p < .01$ ) and consistent across age and sex groups. Of the 345 children positive on the Year 1 BITSEA problem and/or competence cutpoints, 59.4% continued to be positive on the BITSEA in Year 2,  $\chi^2(1, 1098) = 190.5$ ,  $p < .01$ .

### Child Age and Sex Differences

The distributions of the BITSEA scales were examined to determine whether statistically at-risk cutpoints should be defined based on child age and sex. General linear models testing for effects of sex and age (6-month bands) indicated no age effect for BITSEA/P (Table I). The sex effect approached significance ( $p = .06$ ), with higher *problem* scores in boys than girls. Separate cutpoints were indicated for *problems* because the distributions differed at the extremes, such that a universal cutpoint would identify unequal proportions across age by sex groups (Table I). For example, among 12- to 17-month-olds, a cutpoint of 15 would identify fewer girls than boys (18.5% vs. 24.3%, Fisher exact test = 0.0586). Consistent with the developmental nature of the competence, BITSEA/C scores increased significantly with age group, with 12- to 17-month-olds reported as less competent than each older group ( $p < .05$ ). Competence differed with sex, with lower scores in boys than girls in each age group ( $t$  values = 2.22 to 4.22,

**Table I.** Scale Scores of the Brief Infant-Toddler Social and Emotional Assessment by Child Age and Sex ( $N = 1220$ )

BITSEA	Girls					Boys					<i>F</i> -Test Age	<i>F</i> -Test Sex
	<i>M</i>	<i>SD</i>	Cut point	Positive/Total	%Pos	<i>M</i>	<i>SD</i>	Cut point	Positive/Total	%Pos		
<i>Problems</i>												
12–17 months	9.7	(5.9)	13	34/134	25.4	10.2	(6.5)	15	35/142	24.7		
18–23 months	10.2	(6.4)	15	35/152	23.0	11.1	(6.4)	15	37/145	25.5		
24–29 months	9.6	(5.9)	13	41/173	23.7	10.1	(5.7)	14	38/149	25.5		
30–35 months	9.4	(5.6)	14	35/156	22.4	9.9	(5.5)	14	36/155	23.2	1.5	3.6
<i>Competence</i>												
12–17 months	15.9	(3.3)	11	18/134	13.4	15.2	(2.9)	11	19/142	13.4		
18–23 months	17.9	(2.3)	15	19/152	12.5	16.8	(2.8)	13	16/145	11.0		
24–29 months	18.4	(2.4)	15	21/173	12.1	17.2	(3.0)	14	23/149	15.4		
30–35 months	18.1	(2.2)	15	19/156	12.2	17.4	(3.0)	14	22/155	14.2	42.9*	35.32*
<i>Problems and/or competence</i>												
12–17 months				49/134	36.6				48/142	33.8		
18–23 months				47/152	30.9				44/145	30.3		
24–29 months				53/173	30.6				53/149	35.6		
30–35 months				45/156	28.9				48/155	31.0		

\*  $p < .0001$ .

$p < .05$ ), except the youngest ( $t = -0.74$ ). Thus, age by sex statistical cutpoints were indicated for *competence*. Statistical cutpoints for *problems* and *competence* were defined in 6-month age groups by child's sex (Table I).

Based on age and sex findings, cutpoints were set to identify approximately 25% of children in the at-risk range for *problems* and 10% to 15% as low in *competence*, a higher threshold than for *problems*, due to an expectation that significant social-emotional delays will be less common than significant problem behaviors.

### Criterion-Related Validity

The criterion-related validity of the BITSEA was examined relative to the CBCL/1.5-5, an established measure of emotional/behavioral problems (Table II). Correlations between the BITSEA/P and CBCL/1.5-5 internalizing, externalizing, and total problem scores were moderate and significantly higher than correlations between the BITSEA/C and the CBCL/1.5-5 (Fisher  $r$  to  $z'$  transformation = 21.2 to 25.6,  $p < .01$ ). BITSEA/P correlated significantly with evaluator ratings of social-emotional/behavioral problems and negatively with competence ratings. BITSEA/C correlated significantly with evaluator ratings of competence.

In addition, dichotomous sensitivity-specificity analyses were employed to examine criterion-related validity. *Sensitivity* refers to the proportion of true "positives" according to an external validating criterion (e.g., the CBCL) that are BITSEA positives. *Specificity* is the proportion of true "negatives" that are BITSEA

negatives. As unequal margin totals can negatively affect accuracy estimates (Cicchetti, 2001), analyses included all criterion positives and a random sample of an equal number of criterion negatives. To ensure comparability of data when comparing the performance of different BITSEA cutpoints, only subjects with complete data on both cutpoints were used ( $N = 1206$ ). The following guidelines were employed in interpreting sensitivity-specificity estimates: below 70% = poor; 70–79% = fair; 80–89% = good; and 90–100% = excellent (Cicchetti et al., 1995).

The BITSEA/P cutpoint and combined problem and competence cutpoints (BITSEA/PC) had good to excellent sensitivity and good specificity relative to the CBCL/1.5-5 (Table III). Notably, the BITSEA/PC successfully detected 95% of children with clinical CBCL/1.5-5 scores and 85% of those with subclinical/clinical CBCL/1.5-5 scores. Despite the BITSEA's broader symptom coverage, false-positive rates were acceptable (i.e., <30%). One exception was a high false-positive rate compared with the clinical CBCL. This pattern was to be expected, given that the BITSEA/PC includes a competence domain, whereas the CBCL does not.

### Discriminant Validity

Given the documented presence of social-emotional/behavioral problems and lower social competence among toddlers with language delays (e.g., Irwin, Carter, & Briggs-Gowan, 2002), it is important to examine overlap between positive BITSEA scores and low

Table II. Validity of the Problem and Competence Scales of the Brief Infant-Toddler Social and Emotional Assessment

	Year 1						Year 2						
	Evaluator Ratings		Parental Worry		CBCL/1.5-5		CBCL/1.5-5		ITSEA				
	Problems	Competence	Emot./Behav.	Language	Internal	External	Internal	External	Total	Internal	External	Disregulation	Competence
Year 1 BITSEA <i>problems</i>	.34****	-.25**	.34****	.17****	.64****	.63****	.45****	.48****	.53****	.45****	.57****	.55****	-.24****
Year 1 BITSEA <i>competence</i>	-.22*	.36****	-.21****	-.20****	-.23****	-.31****	-.12****	-.20****	-.18****	-.11****	-.21****	-.13****	.63****

CBCL = Child Behavior Checklist; ITSEA = Infant-Toddler Social and Emotional Assessment; BITSEA = Brief Infant-Toddler Social and Emotional Assessment; Emot/Behav = Emotions/Behavior. Reduced sample size for Year 1 CBCL ( $n = 911-915$ ) reflects the 18-month+ age requirement of this measure. Evaluator ratings were collected only for the home visit sample ( $n = 113-124$ ). Sample sizes for other correlations varied due to random missing data (evaluator ratings  $n = 113-124$ ; parental worry  $n = 1,209-1,211$ ; Year 2 CBCL  $n = 1,094-1,107$ ; Year 2 ITSEA  $n = 1,121-1,123$ ). \*  $p < .05$ ; \*\*  $p < .01$ ; \*\*\*  $p < .001$ ; \*\*\*\*  $p < .0001$ .

language skills. Supporting the BITSEA's discriminant validity, 22.9% of children with positive BITSEA/PC scores had low MCDI vocabulary (85 of 372). Of 179 with low vocabulary, 85 (47.5%) had low BITSEA/PC scores. Further supporting discriminant validity, small correlations were observed between BITSEA ratings and parent reports of worry about language skills (Table II). Thus, BITSEA scores do not simply reflect parental worry, such as might be associated with developmental delay.

Predictive Validity

Additionally, Year 1 BITSEA scores significantly predicted both CBCL/1.5-5 and ITSEA *problem* and *competence* scores one year later (Table II). The BITSEA/P and BITSEA/C showed domain specificity, with each predicting scores in the same domain more strongly than scores in the other domain (Fisher  $r$  to  $z' = 7.5$  to  $11.8$ ,  $p < .01$ ).

Associations with Scores on the Infant-Toddler Social-Emotional Assessment

Finally, the relationship between the BITSEA and ITSEA was examined ( $n = 1,216$ ). BITSEA/P correlated positively with ITSEA internalizing, externalizing, and dysregulation domains ( $r = 0.58, 0.75$ , and  $0.75$ , respectively,  $p < .01$ ) and negatively with ITSEA competence ( $r = -0.20$ ,  $p < .0001$ ). BITSEA/C and ITSEA competence correlated highly ( $r = 0.82$ ,  $p < .01$ ), with age partialled from the correlation. BITSEA/C had low negative correlations with the ITSEA internalizing ( $r = -0.06$ ,  $p < .05$ ), externalizing ( $r = -0.23$ ,  $p < .01$ ), and dysregulation domains ( $r = -0.16$ ,  $p < .01$ ). In addition, the BITSEA/P and BITSEA/PC demonstrated fair to good sensitivity and good specificity in detecting children with high ITSEA internalizing, externalizing, and/or dysregulation domains (Table III). Although the BITSEA/C alone was not adequately sensitive to low ITSEA competence, the BITSEA/PC had good sensitivity and acceptable specificity in measuring this domain.

Discussion

Routine pediatric screening using brief screeners has been shown to be feasible and to significantly improve the identification of at-risk infants and toddlers whose difficulties may warrant additional follow-up or intervention (Baird et al., 2000; Jellinek et al., 1999; Murphy et al., 1996). There is currently a need for a comprehensive screener for detecting social-emotional/



**Table III.** Sensitivity and Specificity of the Brief Infant-Toddler Social and Emotional Assessment (BITSEA) Relative to the Child Behavior Checklist (CBCL)/1.5-5 and the Infant-Toddler Social and Emotional Assessment (ITSEA)

	Status on BITSEA (B)/ Criterion Test (C)				Sensitivity (%)	Specificity (%)	Overall Accuracy (%)	Predicted Positive Accuracy (%)	Predicted Negative Accuracy (%)	False Positives <sup>a</sup> (%)	False Negatives <sup>a</sup> (%)
	B-/C-	B+/C-	B-/C+	B+/C+							
BITSEA problems											
Clinical CBCL (t-score ≥63)	46	13	4	55	93.2	78.0	85.6	80.9	92.0	22.0	2.7
Subclinical/clinical CBCL (t-score ≥60)	81	17	19	79	80.6	82.7	81.6	82.3	81.0	17.3	11.9
ITSEA intern, exterm, dysreg <sup>b</sup>	207	26	51	182	78.1	88.8	83.5	87.5	80.2	11.2	19.6
BITSEA competence											
Low ITSEA competence	116	6	38	84	68.9	95.1	82.0	93.3	75.3	4.9	24.9
Combined problem and/or competence											
Clinical CBCL (t-score ≥63)	40	19	3	56	94.9	67.8	81.4	74.7	93.0	32.2	2.0
Subclinical/clinical CBCL (t-score ≥60)	73	25	15	83	84.7	74.5	79.6	76.9	83.0	25.5	8.9
ITSEA intern, exterm, dysreg <sup>b</sup>	188	45	47	186	79.8	80.7	80.3	80.5	80.0	19.3	17.7
Low ITSEA competence	87	35	24	98	80.3	71.3	75.8	73.7	78.4	28.7	13.5

*Overall accuracy* is the proportion of all subjects correctly classified by the BITSEA; *predicted positive accuracy* is the proportion of BITSEA positives that are accurate relative to the criterion; *predicted negative accuracy* refers to the proportion of BITSEA negatives that are accurate according to the criterion.

As described in the Methods, each sensitivity-specificity analysis was conducted on a restricted subset of the sample. This subset included all positives on the criterion measure and a equal number of randomly selected negatives on the criterion.

<sup>a</sup> False positives and false negatives are calculated as the percentage of the criterion negatives and positives, respectively.

<sup>b</sup> ITSEA Internalizing excludes the *inhibition to novelty* scale, in order to maximize comparability with the areas addressed by the BITSEA.

behavioral problems and delays in competence in infants and toddlers.

Findings provide preliminary support for the BITSEA as a reliable and valid brief screener for infant-toddler social-emotional and behavioral problems and delays in competence. When used in a socioeconomically and ethnically diverse community-based population, the BITSEA demonstrated excellent test-retest reliability and good interrater agreement between parents. Moreover, it detected 80% to 95% of infants and toddlers identified by the longer CBCL/1.5-5 as having social-emotional and/or behavioral problems. This finding supports the validity of the BITSEA, given the established success of the CBCL in discriminating referred and nonreferred children (Achenbach & Rescorla, 2000). Of note, despite covering a broader range of symptoms than the CBCL (e.g., inclusion of dysregulation), the BITSEA maintained acceptable rates of false positives and false negatives relative to the CBCL. Further evidence of validity was documented in the small to modest correlations between the BITSEA and evaluator ratings of problems and competence, which were well within the range of commonly reported parent/observer correlations (Achenbach et al., 1987). This observational corroboration is notable because, relative to parents, who had experienced children's behavior in multiple settings and over time, evaluators had relatively limited access to children's behavioral repertoire, having observed them during a fairly brief (1- to 2-hour) and structured assessment protocol. The BITSEA also demonstrated discriminant validity, as low scores on the BITSEA showed low to moderate overlap with delayed vocabulary skills, assessed via parent report. Combined with low correlations with parental worry, this suggests that BITSEA scores do not simply reflect diffuse parental concern or a selective negative view of the child.

In addition, although not an indication of validity per se, the BITSEA performed well compared with the longer ITSEA from which it was drawn, demonstrating good sensitivity and specificity relative to the problem and competence domains. The BITSEA was most sensitive to competence when both problem and competence cutpoints were used, a pattern that may reflect the inclusion in the BITSEA/P scale of atypical behaviors that may co-occur with delay, thus boosting sensitivity to delays in competence.

Further, the BITSEA appears to measure difficulties that, for some children, are fairly enduring, rather than transient problems, such as those often associated with

the "terrible twos." Not only did initial BITSEA scores significantly predict CBCL/1.5-5 and ITSEA *problem* and *competence* scores one year later, but 59% of BITSEA screen positives remained positive on the BITSEA at follow-up. Such evidence of persistence is particularly important given the availability of early intervention services for at-risk infants and toddlers and recent indications of a substantial unmet need for early intervention services in the very young (Horwitz et al., 2003).

Given the intended use of the BITSEA as a first-stage screener for the early identification of children and families who may be experiencing difficulties that merit additional follow-up, an understanding of the practical implications of implementing the BITSEA in routine screening is critical. Our results indicate that when used in a diverse suburban and urban population, the BITSEA is likely to identify about one in three children as at-risk for problems or delays in competence. This rate appears to be reasonable, based on an expectation that 15% of infants and toddlers would have clinically significant social-emotional/behavioral problems and/or delays in competence (Briggs-Gowan et al., 2001; Roberts et al., 1998) and that an additional 15% would have problems in the at-risk range that, while meriting follow-up, are unlikely to require clinical referral. This latter 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.

It is therefore 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). In addition to discussing BITSEA scores with parents, pediatricians may obtain a more detailed profile of infant-toddler strengths and weaknesses by having parents complete the remaining ITSEA questions. This approach offers the clinical benefit of a complete ITSEA scoring profile across the internalizing, externalizing, dysregulation, and competence domains. Alternatively, pediatricians may follow up with more specific measures that focus on the areas of concern that arise from discussions with parents (e.g., autism-specific measures, sleep-problem measures). Only those children and

families who, after further inquiry, evidence impairment would require immediate referral for more extensive evaluation.

### Limitations of the Study

Although this study was an unusual opportunity to develop a screener in a large and sociodemographically diverse representative healthy birth cohort sample, additional psychometric research with the BITSEA is warranted. Due to the study design, the BITSEA items were embedded and answered within the longer ITSEA, a strategy that is often used when developing short forms of longer measures. However, parents may answer BITSEA questions differently when they are not asked in the context of the longer ITSEA. Also, two BITSEA questions were calculated as the composite of ITSEA questions. For these reasons, it would be worthwhile to assess the BITSEA's reliability and validity when answered as a stand-alone measure. 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. For example, the BITSEA scales include several behaviors that are low base rate in a community sample and thus may have higher internal consistency in more symptomatic or delayed samples. It also is probable that a different proportion of children would be identified as positive using the statistically at-risk cutpoints developed in this study if the BITSEA were used to screen early intervention samples. Finally, the work presented herein did not evaluate the psychometrics of the BITSEA with respect to autism spectrum disorders or other early emerging developmental or psychiatric disorders (e.g., feeding disorders). Thus, it is important to establish the BITSEA's clinical validity and to develop cutpoints that reflect clinically significant problems.

To address these issues, the clinical validity of the BITSEA is currently being evaluated in a study in which parent BITSEA ratings, answered independently of the ITSEA, will be examined in relation to consensus diagnoses about the presence of child-onset psychiatric disturbances or parent-child relationship disturbances in infants and toddlers referred to infant mental health clinics. The BITSEA also will be used in a study of children diagnosed with autism. While awaiting clinical cutpoints, when the explicit goal is to identify children with more extreme psychopathology and/or delays in competence, one may employ more stringent cutpoints, based on this representative sample. For example, one

may set cutpoints at 1.5 or 2 standard deviations from the mean on each scale.

### Conclusions

Use of screeners, such as the BITSEA, may improve significantly the identification of infants and toddlers with possible social-emotional problems or delays, thereby aiding efforts to provide early intervention services to young children with early social-emotional/behavioral problems and/or delays in competence. The BITSEA evidences strong psychometric properties and appears to be an appropriate tool for screening for these types of problems.

### Acknowledgement

This research was supported by National Institute of Mental Health grant R01 MH55278 and the Astra Foundation, Acton, Massachusetts.

Received December 23, 2002; revisions received April 1, 2003; accepted June 6, 2003

### References

- Achenbach, T., & Rescorla, L. (2000). *Manual for the ASEBA preschool forms and profiles*. Burlington: University of Vermont.
- Achenbach, T. M., McConaughy, S. H., & Howell, C. T. (1987). Child/adolescent behavioral and emotional problems: Implications of cross-informant correlations for situational specificity. *Psychological Bulletin*, 101, 213–232.
- AAP [American Academy of Pediatrics] (2001). Committee on Children with Disabilities. Developmental surveillance and screening of infants and young children. *Pediatrics*, 108, 192–196.
- 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, 94–702.
- Blumenthal, D., Causino, N., Chang, Y. C., Culpepper, L., Marder, W., Saglam, D., et al. (1999). The duration of ambulatory visits to physicians. *Journal of Family Practice*, 48, 264–271.
- Briggs-Gowan, M., & Carter, A. S. (1998). Preliminary acceptability and psychometrics of the

- Infant-Toddler Social and Emotional Assessment (ITSEA): A new adult-report questionnaire. *Infant Mental Health*, 19, 422–445.
- 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., & Schwab-Stone, M. (1996). Discrepancies among mother, child, and teacher reports: Examining the contributions of maternal depression and anxiety. *Journal of Abnormal Psychology*, 105, 749–765.
- Briggs-Gowan, M. J., Carter, A. S., Skuban, E. M., & Horwitz, S. M. (2001). Prevalence of social-emotional and behavioral problems in a community sample of 1- and 2-year-old children. *Journal of the American Academy of Child and Adolescent Psychiatry*, 40, 811–819.
- 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. (2000). *Manual of the Infant-Toddler Social-Emotional Assessment*. New Haven, CT: Yale University.
- Carter, A. S., Briggs-Gowan, M. J., Jones, S. M., & Little, T. D. (2003). The Infant-Toddler Social and Emotional Assessment: Factor structure, reliability, and validity. *Journal of Abnormal Child Psychology*, 31, 495–514.
- Carter, A. S., Little, C., Briggs-Gowan, M. J., & Kogan, N. (1999). The Infant-Toddler Social-Emotional Assessment (ITSEA). Comparing parent ratings to laboratory observations of task mastery, emotion regulation, coping behaviors, and attachment status. *Infant Mental Health*, 20, 375–392.
- Cicchetti, D. (1993). Developmental psychology: Reactions, reflections, and projections. *Developmental Review*, 13, 471–502.
- Cicchetti, D. V. (2001). *On the evaluation of screeners in behavioral and biomedical research*. Paper presented at the Yale Child Study Center, New Haven, CT.
- Cicchetti, D. V., & Sparrow, S. S. (1981). Developing criteria for establishing interrater reliability of specific items: Applications to assessment of adaptive behavior. *American Journal of Mental Deficiency*, 86, 127–137.
- 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.
- Costello, E. J., Edelbrock, C., Costello, A. J., Dulcan, M. K., Burns, B. J., & Brent, D. (1988). Psychopathology in pediatric primary care: The new hidden morbidity. *Pediatrics*, 82, 415–424.
- Dale, P., Reznick, J., & Thal, D. (1998). *A parent report measure of language development for three-year-olds*. Poster presented at the International Congress for Infant Studies, Atlanta, GA.
- Dale, P., Reznick, J., Thal, D., & Newton, T. (2000). *A parent-report measure of language development for three-year-olds*. Preliminary report.
- Dulcan, M. K., Costello, E. J., Costello, A. J., Edelbrock, C., Brent, C., & Janiszewski, S. (1990). The pediatrician as gatekeeper to mental health care for children: Do parents' concerns open the gate? *Journal of the American Academy of Child and Adolescent Psychiatry*, 29, 453–458.
- 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.
- Eyberg, S. (1980). The Eyberg Child Behavior Inventory. *Journal of Clinical Child Psychology*, 9, 22–28.
- FedStats (2002). MapStats: New Haven County, Connecticut. Website: <http://www.fedstats.gov/qf/states/09/09009.html>. Retrieved March 6, 2003.
- Fenson, L., Dale, P., Reznick, J., Thal, D., Bates, E., Hartung, J., et al. (1993). *The MacArthur Communicative Development Inventories: User's guide and technical manual*. San Diego: Singular Publishing Group.
- Fenson, L., Pethick, S., Renda, C., Cox, J., Dale, P., & Reznick, J. (2000). Short-form versions of the MacArthur Communicative Development Inventories. *Applied Psycholinguistics*, 21, 95–115.
- Forrest, C. B., Glade, G. B., Starfield, B., Baker, A. E., Kang, M., & Reid, R. J. (1999). Gate-keeping and referral of children and adolescents to specialty care. *Pediatrics*, 104, 28–34.
- Glascow, F. P. (2000). Early detection of developmental and behavioral problems. *Pediatric Review*, 21, 272–280.

- Horwitz, S. H., Gary, L. A., Briggs-Gowan, M. J., & Carter A. S. (2003). Do needs drive services or do services drive needs? *Pediatrics*, 112, 1373–1378.
- Horwitz, S. M., Leaf, P. J., Leventhal, J. M., Forsyth, B., & Speechley, K. N. (1992). Identification and management of psychosocial and developmental problems in community-based, primary care pediatric practices. *Pediatrics*, 89, 480–485.
- Horwitz, S. M., Leaf, P. J., & Leventhal, J. M. (1998). Identification of psychosocial problems in pediatric primary care: Do family attitudes make a difference? *Archives of Pediatrics and Adolescent Medicine*, 152, 367–371.
- Irwin, J. R., Carter, A. S., & Briggs-Gowan, M. J. (2002). The social-emotional development of “late-talking” toddlers. *Journal of the American Academy of Child & Adolescent Psychiatry*, 41, 1324–1332.
- 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.
- Jellinek, M. S., & Murphy, J. M. (1988). Screening for psychosocial disorders in pediatric practice. *American Journal of Diseases in Children*, 142, 1153–1157.
- Masten, A. S., & Coatsworth, J. D. (1995). Competence, resilience, and developmental psychopathology. In D. Cicchetti & D. J. Cohen (Eds.), *Developmental psychopathology: Vol. 2. Theory and methods* (pp. 715–752). New York: John Wiley and Sons, Inc.
- McCain, A. P., Kelley, M. L., & Fishbein, J. (1999). Behavioral screening in well-child care: Validation of the Toddler Behavior Screening Inventory. *Journal of Pediatric Psychology*, 24, 415–422.
- Mouton-Simien, P., McCain, A. P., & Kelley, M. L. (1997). The development of the toddler behavior screening inventory. *Journal of Abnormal Child Psychology*, 25, 59–64.
- 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.
- NCES [National Center for Education Statistics] (2001). *Digest of education statistics. Chapter 1: All levels of education, Table 11: Educational attainment of persons 18 years old and over, by state: 1990–2000*. Website: <http://nces.ed.gov/pubs2002/digest2001/tables/dt011.asp>. Retrieved March 6, 2003.
- Regier, D. A., Goldberg, I. D., & Taube, C. A. (1978). The de facto US mental health services system: A public health perspective. *Archives of General Psychiatry*, 35, 685–693.
- Roberts, R. E., Attkisson, C. C., & Rosenblatt, A. (1998). Prevalence of psychopathology among children and adolescents. *American Journal of Psychiatry*, 155, 715–725.
- 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.
- Squires, J., Bricker, D., & Twombly, E. (2002a). *Ages & Stages Questionnaires–Social-Emotional: A parent-completed, child-monitoring system for social-emotional behaviors*. Baltimore: Paul H. Brookes Publishing Co., Inc.
- Squires, J., Bricker, D., & Twombly, E. (2002b). *The ASQ:SE User’s Guide*. Baltimore: Paul H. Brookes Publishing Co., Inc.
- Stancin, T., & Palermo, T. M. (1997). A review of behavioral screening practices in pediatric settings: Do they pass the test? *Journal of Developmental and Behavioral Pediatrics*, 18, 183–194.
- 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.
- U.S. Public Health Service (2000). *Report of the surgeon general’s conference on children’s mental health: A national action agenda*. Washington, DC: Department of Health and Human Services.
- Zuckerman, B., Moore, K. A., & Gleib, D. (1996). Association between child behavior problems and frequent physician visits. *Archives of Pediatric and Adolescent Medicine*, 150, 146–153.