The Infant–Toddler Social and Emotional Assessment (ITSEA): Factor Structure, Reliability, and Validity

Alice S. Carter,[1,4] **Margaret J. Briggs-Gowan,**[2] **Stephanie M. Jones,**[2] **and Todd D. Little**[3]

In this paper the refinement and psychometric properties of the Infant–Toddler Social and Emotional Assessment (ITSEA) are described. Results from a sociodemographically diverse birth cohort sample of 1,235 parents of children between the ages of 12 and 36 months are presented. Confirmatory factor analyses supported the hypothesized Internalizing, Externalizing, Regulatory, and Competence domains as well as the 17 individual scales that comprise the ITSEA. Findings for 3 additional indices useful in identifying significant psychopathology are presented. Subgroup analyses revealed structural invariance and expected mean level differences across both child sex and 6-month age bands. Child sex differences emerged for some problem and most competence scales, with boys rated as higher on Activity/Impulsivity and girls rated higher on Anxiety and most Competence scales. All competence scores increased across age groups. Problem behaviors showed a more mixed developmental pattern. Test–retest and interrater reliability were acceptable. Associations between the ITSEA and independent evaluator ratings and parental ratings of child behavior problems, temperament, and parental distress support the validity of the instrument.

In this paper, the refinement of a recently developed comprehensive adult-report measure of social–emotional problems and competencies in 1- to 3-year-olds, the Infant–Toddler Social and Emotional Assessment (ITSEA), is presented. Until recently, the assessment of very young children's problem behaviors and competencies has been constrained by a lack of age-appropriate instruments. Although systematic study of social–emotional problems and competencies in a normative context was compromised by the lack of age-appropriate instruments, clinical observations and empirical studies of at-risk children have documented the presence of serious and persistent social–emotional problems in infants and toddlers (Radke-Yarrow, Nottelmann, Martinez, Fox, & Belmont,

1992; Sameroff & Emde, 1989; Zeanah, 2000; Zero to Three, 1994). Moreover, many of these early problems have shown remarkable stability into the school-age years (Campbell & Ewing, 1990; Fischer, Rolf, Hasazi, & Cummings, 1984; Prior, Smart, Sanson, Pedlow, & Oberklaid, 1992; Shaw, Keenan, & Vondra, 1994).

Normative data for problem behaviors are available for children from 2- to 3-years of age (Achenbach, Edelbrock, & Howell, 1987; Briggs-Gowan, Carter, Skuban, & Horwitz, 2001; Prior et al., 1992; Richman, Stevenson, & Graham, 1982; Thompson et al., 1996) and more recently for children as young as 18 months of age (Achenbach & Rescorla, 2000; Mathiesen & Sanson, 2000). However, to fully understand the social–emotional development of infants and toddlers, it is critical to assess social–emotional and behavioral competencies as well as problems. Competence in stage salient tasks increases the likelihood of later competence and likely minimizes the emergence and maintenance of maladaptive patterns. In contrast, lags in social–emotional competence will likely increase risk for subsequent social–emotional and behavioral problems (Carter, 2002; Cicchetti & Cohen, 1995;

[1]Department of Psychology, University of Massachusetts Boston, Boston, Massachusetts.

[2]Yale University, New Haven, Connecticut.

[3]University of Kansas, Lawrence, Kansas.

[4]Address all correspondence to Alice S. Carter, PhD, Department of Psychology, University of Massachusetts Boston, 100 Morrissey Boulevard, Boston, Massachusetts 02125; e-mail: alice.carter@umb.edu.

Keenan & Shaw, 1997; Masten & Coatsworth, 1995, 1998). For example, secure infant attachment status is related to lower risk for later peer and behavior problems (Easterbrooks & Goldberg, 1990; Erickson, Sroufe, & Egeland, 1985).

There has been some controversy regarding the presence of sex differences in problem behaviors in very young children. Several studies of problem behaviors in very young children have not reported differences between boys and in girls (Anderson, 1983; Briggs-Gowan et al., 2001; Cornely & Bromet, 1986; Newth & Corbett, 1993), whereas other studies report more externalizing behaviors in boys than in girls (Briggs-Gowan & Carter, 1998; Earls, 1980; Koot & Verhulst, 1991; Luk, Leung, Bacon-Shone, & Lieh-Mak, 1991; Richman, Stevenson, & Graham, 1975) and greater fears and sleep problems in girls than in boys (Earls, 1980; Richman et al., 1975). For example, in one study, parents of 2-year-olds reported higher rates of individual externalizing problems on the CBCL/1.5-5 for boys than for girls (Koot & Verhulst, 1991). Similarly, parents of 1- and 2-year-olds involved in a pediatric study, in which an earlier version of the ITSEA was employed, rated boys as more active and less socially and emotionally competent than girls (Briggs-Gowan & Carter, 1998). Beyond sex differences in rates of problem behaviors and competencies, there are also some data to support sex differences in correlates of problem behaviors and competencies, with boys showing greater vulnerability to maternal depression (Carter, Garrity-Rokous, Chazan-Cohen, Little, & Briggs-Gowan, 2001; Earls & Jung, 1987; Murray, 1992; Shaw & Vondra, 1995) and girls showing more vulnerability to disturbances in the quality of the early mother–infant interaction (Carter et al., 2001). In response to the lack of age-appropriate measures to assess problems and competencies in 1- to 3-year-olds, the ITSEA (Carter & Briggs-Gowan, 1993) was developed. The initial process for selecting items, preliminary psychometric properties in a pediatric sample of 214 parents, and observational validity of an earlier version of the ITSEA were reported previously (Briggs-Gowan & Carter, 1998; Carter, Little, Briggs-Gowan, & Kogan, 1999). Results in a pediatric sample of 1- and 2-year-olds (Briggs-Gowan & Carter, 1998) supported the ITSEA's acceptability to parents and documented acceptable internal consistency for domains (.92 for Externalizing, .75 for Internalizing, .78 for Dysregulation, and .84 for Competencies), strong internal consistency for eight of nine problem scales ($\alpha = .69$–.86, mean $= 0.76$), and marginal to acceptable internal consistency for competence scales (Cronbach's $\alpha = .60$–.79, mean $= 0.69$). Two-week test–retest reliability was good to excellent for all scales (.61–.91, mean $= 0.79$). With respect to validity, moderate correlations in expected di-

rections were observed with parent reports of child temperament, parenting stress, and the CBCL behavior problem domains. In addition, in a separate sample of 12-month-olds (Carter et al., 1999), the ITSEA was shown to account for unique variance in several observed indices of social–emotional functioning (e.g., attachment security, mastery motivation, emotion regulation), after controlling for multiple dimensions of parent-rated temperament.

The current refinement, presented here, includes the addition of several problem items that reflect symptoms included as diagnostic criteria within the DC:0–3 Alternative Diagnostic System for infants and toddlers (Zero to Three, 1994) and the *DSM-IV* (American Psychiatric Association, 1994) nosologies. Although neither the DC:0–3 nor *DSM-IV* has been validated with children in the age range covered by the ITSEA, it is hoped that the inclusion of these items will facilitate empirical efforts that can begin to support or refute the clinical significance of these symptoms in this age range.

The ITSEA is designed to assess a wide array of social–emotional and behavioral problems and competencies. Our operational definition of problems includes behaviors that have traditionally been termed "internalizing" and "externalizing" problems (Achenbach, 1966), as well as regulatory problems, and serious maladaptive behaviors that, even in isolation, may be of concern (e.g., head-banging). Behaviors represented on the ITSEA include both (1) those behaviors that are otherwise part of typical development, but become problems when exhibited either in excess or too infrequently, and (2) infrequently occurring problem behaviors that represent deviations from a normative developmental course. In infants and toddlers, externalizing behaviors include high activity, impulsivity, aggression, and defiance. Internalizing problems include depression, social withdrawal, anxiety, separation distress, and extreme inhibition/shyness. Dysregulation includes problems in sleeping and eating, problems regulating negative emotional states with respect to reactivity and regulation, and unusual sensory sensitivities. The *a priori* decision to include negative emotionality as part of the dysregulation domain was based on conceptualizing negative emotionality within a developmental regulatory framework that builds on the work of Rothbart (e.g., Rothbart, 1981; Rothbart & Putnam, 2002). Within this normative temperamental frame, infants are expected to vary in the latency and intensity of their affective responses to a range of negative stimuli (i.e., reactivity). In addition, as part of a larger developmental task of achieving state regulation, infants are expected to acquire increasingly sophisticated and effective emotion regulation skills. Greater emotional negative reactivity and lower emotion regulation capacities are viewed as risk factors

for later internalizing and externalizing problems and disorders as well as problems that can cause significant impairment for the infant/toddler and family. However, emotional negativity, which encompasses a range of specific emotions (e.g., sadness, anger), is not viewed as having specificity for either internalizing or externalizing disorders until later in development. Consistent with this conceptualization, longitudinal studies have shown that negative emotionality, a component of difficult temperament, predicts both internalizing and externalizing problems later in childhood (e.g., Earls & Jung, 1987; Prior et al., 1992; Rende, 1993; Shaw & Vondra, 1995; Shaw, Owens, Giovannelli, & Winslow, 2001).

Social–emotional competencies include compliance, attention regulation, imitation and pretend play skills, mastery motivation, empathy, emotional awareness, and prosocial peer behaviors (Eisenberg & Mussen, 1989; Radke-Yarrow & Zahn-Waxler, 1984; Saarni, 1988; Zahn-Waxler, Radke-Yarrow, Wagner, & Chapman, 1992). Problems and competencies are viewed as related, but independent, constructs. For example, although problems may show some normative developmental variation, they are not expected to increase over time. In contrast, competencies are expected to increase over time. Thus, attention regulation is included in the competence, rather than externalizing, domain. Further, although certain competence items address behaviors that may appear to be the opposite of problem behaviors (e.g., "Tries to do as you ask" vs. "Is disobedient or defiant, for example refuses to do as you ask"), competencies are viewed as reflecting the presence of age-appropriate skills, not simply the lack of problem behaviors. For example, a socially withdrawn child who is not disobedient/defiant, yet also does not attempt to "do as you ask," may be evidencing a lag in competence without displaying related behavior problems.

Similarly, although aspects of social–emotional problems and competence are expected to be associated with temperament, they are viewed as conceptually distinct from temperament. Whereas both temperament and social–emotional problems and competencies may share genetic and environmental determinants, temperament is conceptualized as an internal, enduring set of attributes, whereas social–emotional problems and competencies reflect adaptations between the child's temperamental predispositions and the affordances in the environment. For a more extensive discussion of this distinction, Goldsmith and Lemery (2000) present 10 etiological explanations for the observed relationship between measures of temperament and child psychopathology.

Several issues complicate the task of defining and assessing problems that emerge in infancy and toddlerhood (Del Carmen-Wiggins & Carter, 2001). Developmental issues must be considered, as many behaviors thought to be "clinically relevant" at older ages may be manifestations of normal development in early childhood, such as the self-assertion that typically emerges in the 2nd year of life (Campbell, 1990). In addition, some behaviors may be context-specific or reflect a mismatch between a child's developmental level and the situational demands and supports in the environment. For example, a child may be highly aggressive in the presence of a particular child who has bitten her in childcare, but not evidence unusual rates of aggression in other settings. Thus, it is optimal to identify "symptoms" as clusters of behaviors that occur consistently across contexts and time, rather than as individual behaviors (Achenbach, 2000; Campbell, 1995). Furthermore, although parents are often the most intimately involved and available informants about very young children, it is optimal to obtain information from multiple sources (e.g., from child care providers or by observation). Problems may be context-specific and informants' perceptions may differ (Achenbach, McConaughy, & Howell, 1987; Briggs-Gowan, Carter, & Schwab-Stone, 1996; Offord, Boyle, & Racine, 1989). At the same time, parent perceptions of child problems may, in and of themselves, confer risk for later problems. Therefore, assessment of child problems is optimized by considering the possible influences of developmental level, context, and informant, and focusing on clusters of behaviors, rather than on individual behaviors.

It is also important to emphasize that identifying social–emotional and behavioral problems within the child neither minimizes the contribution of the family environment to the development and continuity of early problems nor negates the need for relational interventions with young children. Indeed, studies that assess both the child and the family environment (e.g., Shaw et al., 1994) reveal that each uniquely contributes to the prediction of later problems. The strategy adopted here, that of focusing on the child, is based on the assumption that parents will provide more reliable information about their young children's behaviors than about dyadic or triadic disturbances, or their own maladaptive parenting behaviors.

At the time that we initiated the development of the ITSEA, a review of existing measures revealed no comprehensive instrument that would be suitable for an epidemiologic study of social–emotional problems and competencies in 1- and 2-year-old children. Existing measures were limited either in age range or in scope. For example, the Child Behavior Checklist for 2- and 3-year-olds (CBCL/2-3; Achenbach, 1992), although valuable for studying behavior problems, had not been adapted for use under the age of 2 years. Even with the recent publication of the CBCL/1.5-5 (Achenbach & Rescorla, 2000), the CBCL

is not appropriate below 18 months of age. Moreover, the CBCL was not designed to address social–emotional competencies. Other extant measures focus on specific aspects of typical social–emotional development, but do not comprehensively assess social–emotional problems and competencies (e.g., The Child Development Inventories, Ireton, 1992; and the Vineland Adaptive Behavior Scales; Sparrow, Balla, & Cicchetti, 1984). Thus, the ITSEA was developed on the basis of the need for a comprehensive measure of problems and competencies for 1- to 3-year-old children.

Because of the lack of existing validated instruments and/or a validated diagnostic nosology, demonstrating the validity of the ITSEA represents a significant challenge. With a limited set of criterion measures, the presence of expected associations with related constructs can be used to support the validity of the measure being evaluated (Carter, Grigorenko, & Pauls, 1995). The approach taken in this paper is to examine associations with the CBCL/1.5-5 and with measures of maternal distress, which have consistently been associated with the early emergence of social–emotional problems (Carter et al., 2001; Downey & Coyne, 1990; Gelfand & Teti, 1990). Moreover, consistent with our earlier work (Carter et al., 1999), we expected moderate associations between ITSEA domains and parent ratings of temperament. Obtaining moderate associations indicates that these are related but independent constructs. In addition, to provide evidence that other informants and methods corroborate parent reports on the ITSEA, associations between parental ITSEA ratings and independent observer ratings are presented for a subsample of children.

To establish the psychometric properties of the refined ITSEA, five general hypotheses were examined. First, we hypothesized that parents would be successful in rating their child's behavior in a coherent, yet differentiated, manner. Specifically, we hypothesized that parental ratings would lend support to the existence of the hypothesized ITSEA scales and domains of social–emotional/behavior problems and competencies. Second, we hypothesized that similar items would be appropriate for boys and girls and for children across the 12- to 36-month age range. Third, we hypothesized that sex differences in ITSEA ratings would be observed, with girls rated higher on competence than boys, and boys rated higher on overactivity and aggression. Fourth, we hypothesized that the ITSEA would demonstrate acceptable test–retest and interrater reliability. Fifth, in the area of validity, we hypothesized that ITSEA domains would correlate highly with CBCL scores, reflecting shared construct coverage, but correlate moderately with parent ratings of temperament and parental affective symptoms. We

also hypothesized that parent reports of competence would correlate with standardized assessments of development. Finally, we hypothesized that parent reports would be corroborated by independent observational ratings.

METHOD

Participants

Reports from parents of children between 12 and 36 months of age ($N = 1,235$; 605 boys and 630 girls) were included in the analyses for this report. An age- and sex-stratified sample of 8,404 was randomly drawn from all children born at Yale-New Haven Hospital who lived in the 15 towns/cities that comprise the New Haven-Meriden Standard Metropolitan Statistical Area from the 1990 Census. Birth records were provided by the State of Connecticut Department of Public Health.

Because of the goal of refining the ITSEA in a developmentally healthy sample of children, children who were likely to have developmental delays were excluded from the sample. Birth record data were used to exclude children with the following conditions: low birth weight ($<2,200$ g); preterm (less than 36 weeks gestation); low Apgar scores (1- and 5-min scores below 5); birth complications likely to be associated with developmental delay (e.g., need for resuscitation and severe anoxia); and known genetic or congenital anomalies (e.g., Down's syndrome). A total of 675 children (8%) were excluded on the basis of these criteria. Eighteen children who were deceased or adopted prior to sampling were excluded. One child per mother was selected, resulting in 277 sibling exclusions. One child of an investigator also was excluded. After randomly sampling 1,788 families from the remaining sample ($N = 7,433$), families were excluded if (1) neither parent spoke English well enough to answer the surveys ($n = 50$) and (2) the biological mother had lost custody of the child ($n = 17$) or the family had moved out of state ($n = 116$). These 183 excluded families did not differ from remaining families ($n = 1,605$) in terms of maternal education, child race, 1-min Apgar scores, gestational age, birth weight or number of years at the birth address. Parental age and 5-min Apgar scores were slightly higher in the eligible sample ($t = 2.04$–4.03, $p < .05$).

The response rate among the 1,605 eligible participants was 79.8%, with 1,280 out of 1,605 eligible families participating. Nonparticipants tended to be younger, less educated, and to have lower birth weight infants than did participants. There was also a lower response rate among minority families (69.9%) than among Caucasian families (83.8%), $\chi^2(1, 1603) = 38.6$, $p < .01$. However,

minority participants and non-participants did not differ on any birth status or sociodemographic variables. Moreover, the effect sizes for age, education, and minority status were small. Four percent of parents ($n = 45$) who delayed participation until children were older than 36 months were excluded from this report.

Children were approximately equally distributed in terms of child sex (50.9% boys) and age group (47.9% 1-year-olds and 52.1% 2-year-olds). The mean age of children was 23.8 months ($SD = 6.8$). Participants were ethnically and socioeconomically diverse. Approximately 34% of children were members of ethnic minority groups. Most children were first- (41.2%), second- (38.1%), or third-born (14.3%), whereas 5.4% were fourth-born and 0.9% were fifth- or sixth-born. Respondent education varied, including 8.0% who had not completed a high school education, 18.2% with a high school degree or equivalent, 32.3% with some education beyond high school, and 41.5% with at least a college degree. Most respondents (96.2%) were biological mothers. Approximately 19% of households were single parent homes. The mean household size was 4.1 people ($SD = 1.1$, range $= 2$–13), with a mean of 2.0 adults ($SD = 0.6$) and 2.1 children ($SD = 1.0$). Poverty estimates, calculated against the Consumer Price Index and on the basis of household income and family composition, indicated that 18.1% of families were living below the poverty line and 15.7% of families were living in borderline poverty (with incomes above 100% but below 185% of the poverty line). Most families (87.6%) used nonparental care at least 1 hour per week. Among families using childcare, the mean number of hours per week in childcare was 22.1 ($SD = 18.7$).

Procedures

Parents were mailed a letter describing their selection for a study about young children's social–emotional development and behaviors. The next week, a questionnaire booklet and a children's book were mailed. The booklet included questions about children's development, family demographic characteristics, parental affective symptoms, parenting styles and stress, and family functioning. Families who did not participate within 1 month were mailed a second questionnaire booklet and children's book. Families who did not participate within 2 months of receiving the initial booklet were contacted by phone and/or in-person to offer interviews, babysitting, and other assistance needed to facilitate participation. Informed consent procedures were followed. The entire booklet could be completed in approximately 90 min. Parents who participated received $25.

A subset of families who completed the questionnaire booklet and whose children were in daycare a minimum of 15 hours per week were invited to participate in a subsequent methodological substudy ($n = 173$). For married and cohabiting families, both parents were asked to complete questionnaire booklets. In all families, one parent was asked to participate in a home visit with their child, in which (1) parent–child play, teaching, and cleanup interactions were filmed; (2) the child was assessed with the Mullen Developmental Scales (Mullen, 1989); and (3) parents were interviewed with the Vineland Adaptive Behavior Scales—Expanded Edition (Sparrow et al., 1984). Evaluators were blind to all parental questionnaire ratings. Following home visits with children, evaluators rated children's problem behaviors and competencies, on the basis of their observations during the entire visit with the child and family. Home visit families 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 status.

Measures

Sociodemographic Variables

Participants answered several questions about sociodemographic factors including child sex, age, ethnicity, and birth order, and maternal age, parental education, marital status, before-tax household income, and receipt of governmental assistance. Additional information, including infant birth weight, gestational age, APGAR scores, parental age, and maternal education was obtained from birth records provided by the State of Connecticut Department of Public Health.

Infant–Toddler Social and Emotional Assessment
(Carter & Briggs-Gowan, 1993, 2000)

The ITSEA assesses four broad domains of behavior (i.e., Externalizing, Internalizing, Dysregulation, and Competencies). In addition, Maladaptive, Atypical Behavior, and Social Relatedness indices are included to assess more serious problems, which tend to have low base rates of occurrence. The Externalizing domain is composed of Activity/Impulsivity, Aggression/Defiance, and Peer Aggression scales. The Internalizing domain includes Depression/Withdrawal, General Anxiety, Separation Distress, and Inhibition to Novelty scales. The Dysregulation domain includes Sleep, Negative Emotionality, Eating, and Sensory Sensitivity scales. Competencies

include Compliance, Attention, Imitation/Play, Mastery Motivation, Empathy, and Prosocial Peer Relations scales. The Maladaptive index includes symptoms of Tourette's syndrome, Posttraumatic stress disorder, toileting problems, sexualized behavior, and pica. Two additional indices, Social Relatedness and Atypical Behavior, are included to assess behaviors that may be indicative of the presence of Pervasive Developmental Delay (PDD)/ Autism. A final set of items is composed of low base-rate behaviors of possible clinical significance. These behaviors were rated as clinically informative by an "expert" panel of psychiatrists and psychologists with clinical and research experience with infants and toddlers, but did not belong conceptually and/or statistically on any domain or index.

The core components of the ITSEA (all scales excluding Atypical and Social Relatedness indices, and the individual items of clinical significance) comprise 139 items. The complete ITSEA includes 166 items. Selected items within scales and domains are listed in the Appendix. Items are rated on the following 3-point scale: (0) *Not true/rarely*, (1) *Somewhat true/sometimes*, and (2) *Very true/often*. A "No opportunity" code allows parents to indicate that they have not had the opportunity to observe certain behaviors (e.g., behavior with peers). Multiple reading level indices demonstrate that the ITSEA requires a fourth- to sixth-grade reading level. The version presented in this report requires approximately 30 min to complete. To minimize response set biases, problem and competence items are interwoven. Psychometric properties are presented in the Results section.

Child Behavior Checklist for 1.5-5 (Achenbach and Rescorla, 2000)

The CBCL/1.5-5 is composed of 113 items and consists of Internalizing, Externalizing, and Total Problem domains. This measure has demonstrated very good 8-day test–retest reliability ($r = .68$–$.92$, mean $r = .84$), cross-informant agreement (mean mother–father $r = .61$, mean parent–child care $r = .65$), and success in discriminating between referred and nonreferred children (Achenbach & Rescorla, 2000). The CBCL/1.5-5 was administered to parents of children older than 17 months.

Colorado Child Temperament Inventory (CCTI; Buss & Plomin, 1975)

The CCTI is a parent-report measure of child temperament, which contains five subscales (i.e., emotionality, sociability, soothability, attention, and activity). Each

subscale has demonstrated good internal consistency (α from .73 to .88). With the exception of the soothability subscale ($r = .43$), the CCTI has shown acceptable test–retest reliability, with correlations from .58 to .80 (Lyon & Plomin, 1981; Plomin, 1976; Rowe & Plomin, 1977). To minimize respondent burden, only the sociability, soothability, and emotionality scales were included in the study. Each scale is composed of five items.

Center for Epidemiologic Studies Depression Inventory (CESD; Radloff, 1977)

The CES-D is a 20-item self-report scale that assesses depressive symptoms in adults. It has demonstrated high internal consistency (coefficient α from .84 to .90) and modest test–retest reliability for 2- to 4-week intervals (r from .51 to .67; Radloff, 1977).

Beck Anxiety Inventory (BAI; Beck, Epstein, Brown & Steer, 1988)

This self-report measure consists of statements that describe common symptoms of anxiety. The individual indicates how much he/she has been bothered by each symptom on a 4-point scale from "*not at all*" to "*severely bothered*." Psychometric properties are adequate.

Vineland Adaptive Behavior Scales for Children— Expanded Form (Sparrow et al., 1984) is a measure of personal and social sufficiency. The Vineland was designed to examine the domains of Communication, Daily Living Skills, Socialization, and Motor Skills. The Adaptive Behavior Composite provides a summary score across domains. This semistructured interview was normed to a carefully selected national standardization sample, and has excellent levels of internal reliability for each domain (ICC > .95). The Vineland was included to evaluate the validity of ITSEA Competence domain.

The Infant Mullen Scales of Early Learning (Infant MSEL; Mullen, 1989) is a comprehensive scale of mental and motor ability for children from birth to 36 months of age. It has a strong theoretical base in neurodevelopment and intrasensory and intersensory learning. Further, it assesses both visual and language abilities at both receptive and expressive levels. The Infant MSEL is divided into the following five scales: Gross Motor Base; Visual Receptive Organization; Visual Expressive Organization; Language Receptive Organization; and Language Expressive Organization. The MSEL also provides a Composite Score that is based on the four mental scales. The MSEL was standardized on a nationally representative sample and psychometric properties appear adequate.

Evaluator Ratings of Child Problems and Competencies

Following home-based assessments, evaluators rated children's problems and competencies on the basis of behavior that was observed during the visit. Evaluators rated the presence of 17 potential problem behaviors on a 4-point scale (0 = *No Problem*, 1 = *Possible Problem*, 2 = *Probable Problem*, and 3 = *Definite Problem*). Nine competencies were rated along a 7-point scale (1 = *Definite Problem* to 7 = *Definite Competence*). Ratings were averaged to form scores within the Externalizing, Internalizing, Dysregulation, Competence, and Maladaptive areas. The Externalizing score was computed as the mean of ratings on the following behavior problem descriptions: *Overactivity*; *Peer Aggression*; *Aggression/Defiance (General)*; and *Behavior Problems*. The Internalizing score was based on ratings of *Depression/Social Withdrawal*, *Inhibition/Extreme Shyness*, and *Anxiety/Fears*. Dysregulation was composed of ratings of *Negative Emotionality* (e.g., excessive crying), *Unusual Sensory Sensitivities*, and *Eating*. *Separation Problems* and *Sleep Problems* were included on the rating form, but excluded from summary scores because evaluators rarely had the opportunity to observe these behaviors in the home visit context. The Competence score was the mean of ratings on *Compliance*, *Attention Skills*, *Persistence/Enjoyment of Challenging Activities*, *Imitation/Pretend Play*, *Prosocial Peer Interactions*, *Empathy*, and *Awareness of Others' Emotions*. Interrater reliability of evaluator ratings was evaluated in a separate study involving 40 children referred for early intervention. The mean interval between ratings made by early intervention service providers and ratings by members of the research team was approximately 5 weeks ($M = 35.7$ days, $SD = 48.5$ days). Results indicated good interrater reliability ($r = .56$ for Internalizing to .72 for Dysregulation, Mean $r = .66$).

ANALYTIC PROCEDURES

To assess the hypothesized structure of the ITSEA measure, confirmatory factor analysis (CFA) using LISREL 8.30 (Joreskog & Sorbom, 1999) was employed. Both the theoretical background and the prior empirical work on the ITSEA dimensions warranted a CFA approach. In addition, CFA is sensitive to departures from the expected structure, allowing us to discern any deviations from the hypothesized model.

In modeling efforts, cross-validation procedures were utilized to minimize chance findings. Specifically, the final models for each scale comprising a domain were derived on one randomly selected split-half of the total sample and confirmed on the other. Although the CFA approach with cross-validation is particularly well-suited for scale refinement, we were prepared to employ exploratory analyses if the expected structure of the scales deviated substantially and theoretically from our a priori expectations, or failed to cross-validate.

Both empirical and rational approaches were employed to identify items to be retained in the final models. With respect to empirical efforts, individual item loadings were inspected and only items with a communality of at least 10% were retained (Achenbach, 1992). In addition, tolerance statistics (i.e., residuals, standard errors, and modification indices) were examined to determine if additional estimates were needed. Finally, the clinical relevance of any changes from the initial hypothesized structure was considered. For example, when it was possible to avoid compromising the statistical integrity of scales, attempts were made to retain those items that correspond to symptoms necessary for diagnosing childhood *DSM-IV* and/or DC: 0–3 disorders. Because of the hierarchical nature of the dimensions captured by the ITSEA, we were also sensitive to the fact that some scales may in fact be divisible into subdimensions. In such cases, all subsets of items were theoretically derived and then tested using the cross-validation procedures employed for the primary dimensions. Using these criteria, 166 of the possible 229 items were retained.

After establishing the basic structure of the ITSEA dimensions, the scales were examined to determine if they were appropriate for both boys and girls across the four age groups (11–17 months, 18–23 months, 24–29 months, and 30–36 months). This was done using multiple-group mean and covariance structure (MACS) methods. In addition to other psychometric advantages, such as correcting for measurement unreliability, MACS analyses allow the distinction between the measurement and the construct levels of analysis. At the measurement level, MACS analyses test the equivalence of the items' factor loadings and the items' intercepts across groups. At the construct level, MACS analyses test the equivalence of the constructs' variances, covariances, and means across groups (Browne & Arminger, 1995; Little, 1997; Meredith, 1993). Model fit was assessed using three standard fit indices: the root-mean-squared error of approximation (RMSEA), for which values of .08 or less are deemed acceptable, the Comparative Fit Index (CFI), for which values greater than .90 are deemed acceptable, and the Non-normed Fit Index, which also has a criterion of .90 (see Bollen & Long, 1993; Browne & Cudeck, 1993; Marsh, Balla, & McDonald, 1988). Because the maximum likelihood chi-square value is highly sensitive to sample size, it was not employed to evaluate overall model fit.

For both gender and age comparisons, we followed a sequence of three nested models. First, the loadings and intercepts were free to vary across groups. Second, we tested the equality of the loadings. Third, we added equality of the intercepts or item means. This third model was deemed a good representation, provided its fit remained within the criteria of the relative fit statistics, and the drop in fit from the first model was within the tolerance criterion of .05 suggested by Little (1997) and Tucker and Lewis (1973). As a supplement to testing the main effects of the latent means across child age and sex, we examined interactions between sex and age via a standard analysis of variance. To examine the degree of construct independence of ITSEA domains and scales, intercorrelations among the resulting constructs were examined.

Test–retest and interrater reliability were examined in a subset of families who were involved in the methodologic substudy. To establish criterion validity, we examined the associations of the ITSEA constructs with (a) the CBCL/1.5-5 Total Problems, Internalizing, and Externalizing domains; (b) the CCTI Soothability, Emotionality, and Sociability temperament scales; (c) parental distress ratings based on CESD and BAI scores; (d) evaluator ratings of children's behavior based on observations during home-based evaluations; and (e) standardized developmental assessment based on the Vineland and Mullen scales.

RESULTS

The results of the final confirmatory factor analytic models for each domain and scale are presented in Table I. Model fit indices and the range of item loadings are presented (Table I). Of the 17 hypothesized primary scales expected from the ITSEA, the following 12 were well-validated as unidimensional constructs: Depression/Withdrawal; Separation Distress; Inhibition to Novelty; Sleep; Negative Emotionality; Sensory Sensitivity; Compliance; Attention; Imitation/Play; Mastery Motivation; Empathy; and Prosocial Peer Relations. In contrast, for five scales, analyses indicated that the scales were not unidimensional, and that items could be efficaciously divided into subdimensions (see Fig. 1). Each subdimension was both theoretically meaningful and empirically supported through cross-validation techniques. In addition, in every case the chi-square difference test comparing the initial unidimensional models to their final multidimensional counterparts revealed a significant increase in overall model fit ($p < .05$). The Activity/Impulsivity scale revealed two dimensions differentiating between Activity

(e.g., "Is constantly moving") and Impulsivity (e.g., "Gets hurt so often you can't take your eyes off him/her"). The Peer Aggression scale evidenced two dimensions characterized as Relational Aggression (e.g., "Teases other children") and Overt Aggression (e.g., "Hurts other children on purpose"). The General Anxiety scale comprised two dimensions distinguishing between Worry (e.g., "Seems nervous, tense, or fearful") and Obsessionality (e.g., "Is very worried about getting dirty"). The Eating scale also showed two dimensions, Typical Eating (e.g., "Is a picky eater") and Pathology (e.g., "Holds food in cheeks"). Finally, the Aggression/Defiance scale revealed four subdimensions: Defiance (e.g., "Is disobedient or defiant, refuses to do as you ask"), Relational Defiance (e.g., "Acts bossy"), Dispositional Aggression (e.g., "Acts aggressive when frustrated"), and Conduct related items (e.g., "Hurts animals on purpose," "Swears").

Child Age and Sex

When the latent means for the ITSEA scales and dimensions were tested across age and sex by using MACS methods, we found that structural invariance was tenable for each ITSEA scale and sub-dimension. As expected, several constructs evidenced mean-level differences across sex and age. In the MACS analyses, latent means were higher for girls than for boys on Compliance, Attention, Imitation/Play, Empathy, and Prosocial Peer Relations. There were no gender or age effects for the three clinical indices (i.e., Maladaptive, Social Relatedness, or Atypical Index). In the latent mean analyses for age, all competence scores increased across age groups. Problem behaviors showed a more mixed pattern. Activity, Eating Problems, and Separation Distress appear to diminish across this age range. In contrast, within the Internalizing domain, General Anxiety and Depression/Withdrawal appear to increase with age.

In addition, Analyses of Variance were employed to test for age and sex differences in raw means on the ITSEA domains and scales (see Table II for raw means and standard deviations by age group and by child sex). Only main effects are presented because there were no significant interactions between age and sex for any domain or scale. In the Competence domain, the pattern of age and sex differences was the same as that revealed in the MACs analyses, with age-related increases in competence and greater competence among girls than among boys. Also consistent with the MACs analyses, General Anxiety was reported to increase with age, whereas Separation Distress appears to decrease with age. In addition, Aggression/Defiance

Table I. Confirmatory Factor Analysis Model Fit Statistics for ITSEA Scales ($N = 1,235$)

	# of items	Alpha coefficients	Item loadings	Model fit statistics			
				Chi-square	RMSEA	CFI	NNFI
1. **Externalizing Symptoms**	24	0.87		$\chi^2(17) = 86.74$	0.058	0.976	0.960
Activity/impulsivity	6	0.73	0.42–0.72	$\chi^2(8) = 28.83$	0.047	0.983	0.970
Aggression/defiance	12	0.79	0.31–0.67	$\chi^2(46) = 215.06$	0.055	0.943	0.919
Peer aggression	6	0.79	0.54–0.77	$\chi^2(8) = 47.09$	0.064	0.980	0.963
2. **Internalizing Symptoms**	30	0.80		$\chi^2(38) = 168.77$	0.053	0.951	0.930
Depression/withdrawal	9	0.74	0.36–0.73	$\chi^2(25) = 84.75$	0.045	0.975	0.965
General anxiety	10	0.71	0.32–0.72	$\chi^2(84) = 119.09$	0.045	0.951	0.935
Separation distress	6	0.73	0.39–0.70	$\chi^2(7) = 9.03$	0.015	0.999	0.997
Inhibition to novelty	5	0.77	0.50–0.73	$\chi^2(4) = 12.68$	0.042	0.994	0.984
3. **Dysregulation**	34	0.86		$\chi^2(60) = 298.87$	0.057	0.945	0.929
Sleep	5	0.78	0.46–0.76	$\chi^2(4) = 14.20$	0.045	0.994	0.986
Negative emotionality	13	0.84	0.30–0.70	$\chi^2(63) = 220.14$	0.046	0.959	0.950
Eating	9	0.78	0.32–0.85	$\chi^2(26) = 122.77$	0.055	0.962	0.948
Sensory sensitivity	7	0.63	0.29–0.67	$\chi^2(13) = 46.63$	0.045	0.959	0.933
4. **Competence**	37	0.90		$\chi^2(61) = 325.25$	0.059	0.957	0.945
Compliance	8	0.74	0.39–0.74	$\chi^2(20) = 94.94$	0.056	0.954	0.936
Attention	5	0.70	0.45–0.69	$\chi^2(5) = 25.05$	0.056	0.979	0.959
Imitation/play	6	0.59	0.27–0.55	$\chi^2(9) = 12.80$	0.018	0.993	0.988
Mastery motivation	6	0.62	0.28–0.70	$\chi^2(9) = 33.94$	0.046	0.964	0.939
Empathy	7	0.82	0.45–0.75	$\chi^2(14) = 38.53$	0.038	0.990	0.984
Prosocial peer relations	5	0.66	0.42–0.75	$\chi^2(5) = 30.30$	0.060	0.971	0.942
5. **Additional Indices**	31						
Maladaptive	13	0.56	0.12–0.61	$\chi^2(63) = 284.79$	0.060	0.902	0.879
Social relatedness	10	0.56	0.33–0.71	$\chi^2(33) = 112.61$	0.044	0.921	0.893
Atypical	8	0.45	0.17–0.82	$\chi^2(20) = 86.50$	0.051	0.951	0.931

increases in the early part of the age range, between 12 and 24 months. Inhibition to Novelty appears to peak in the 18- to 23-month period. Finally, boys were rated higher than girls in Activity/Impulsivity, whereas girls were rated higher than boys on General Anxiety.

Reliability

In addition to examining confirmatory models, traditional psychometric properties, including coefficient alpha, test–retest, and cross-method reliabilities, were assessed. Internal consistency for each domain and scale is presented in Table I. Test–retest reliability was evaluated in 93 families who completed the ITSEA within a 44-day time interval ($M = 26.81$, $SD = 7.83$). Test–retest coefficients for domains ranged from 0.82 to 0.90 and from 0.69 to 0.85 for scales. Further, a subsample of 90 parents completed the ITSEA as both an interview and a questionnaire. The magnitude of test–retest reliability coefficients for scales and domains did not differ significantly between parents who were interviewed versus those who completed the ITSEA as a questionnaire

on both occasions. Information on interrater agreement was available for 100 mother–father pairs. Agreement between mothers and fathers based on intraclass correlation coefficients (ICC) ranged from 0.58 to 0.79 for domains (Mean ICC = 0.71) and from 0.43 to 0.78 for scales (Mean ICC = 0.64).

Domain Specificity

Pearson correlations among the ITSEA domains were computed and revealed low to moderate associations among the ITSEA problem domains ($r = .35$ for the Internalizing and Externalizing domains, $r = .50$ for the Internalizing and Dysregulation domains, and $r = .54$ for the Externalizing and Dysregulation domains). There were small negative correlations between Competence and two of the problem domains ($r = -.18$, and $-.13$, $p < .001$, for Externalizing and Dysregulation, respectively). The correlation between the Internalizing and Competence domains was not significant. Information about correlations among individual scales is available in the technical manual (Carter & Briggs-Gowan, 2000).

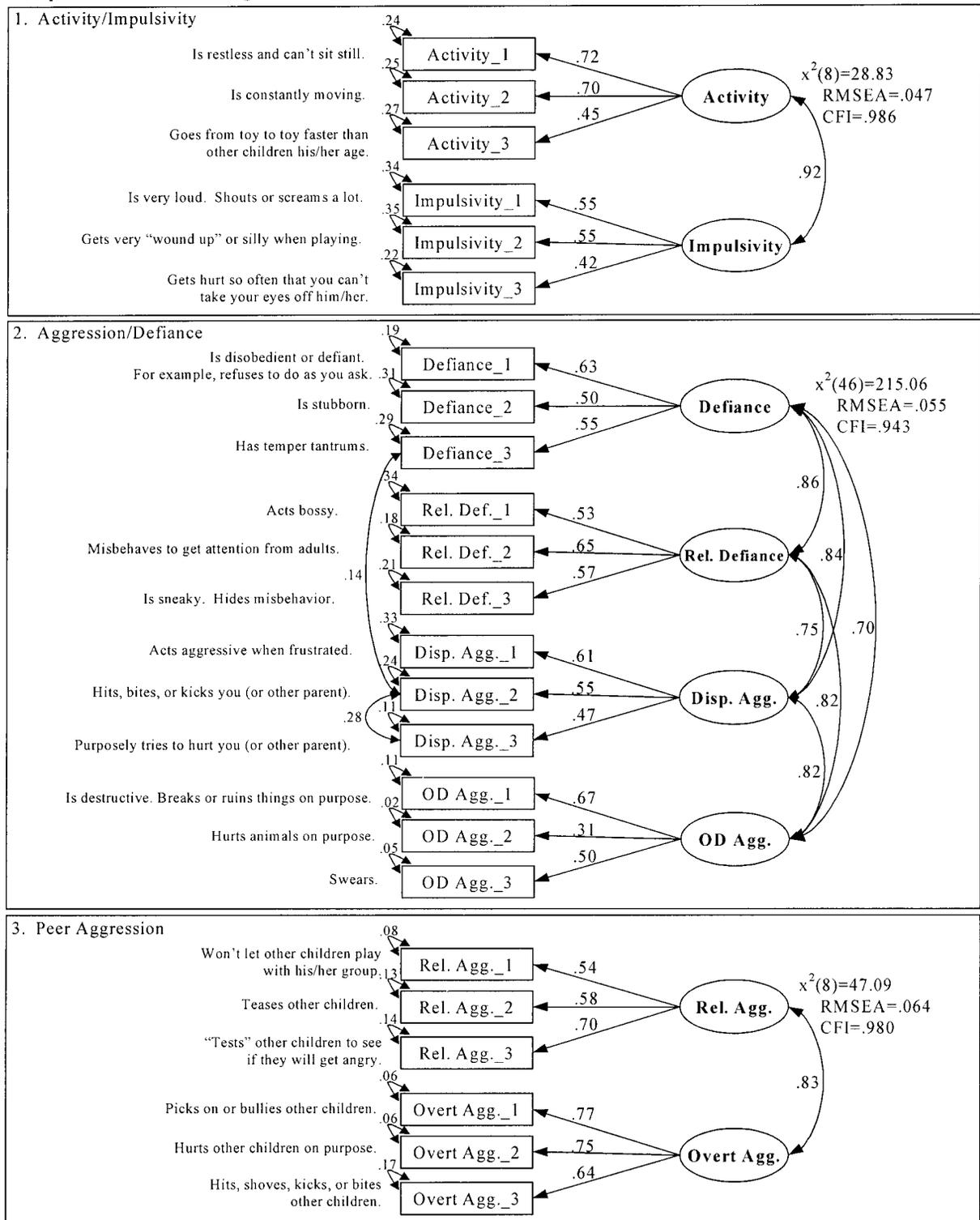**Components of Externalizing Domain.**

Fig. 1. Scale models with confirmatory factor analysis for (1) Activity/Impulsivity scale, (2) Aggression/Defiance, (3) Peer Aggression, (4) General Anxiety, and (5) Eating scale.
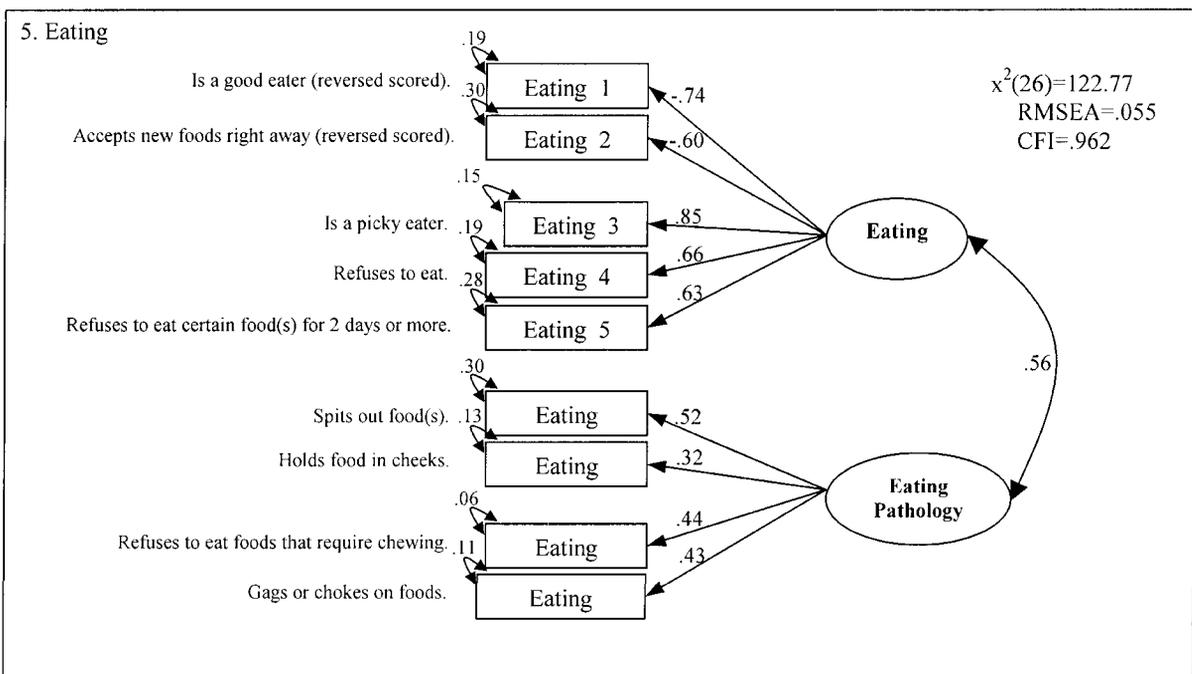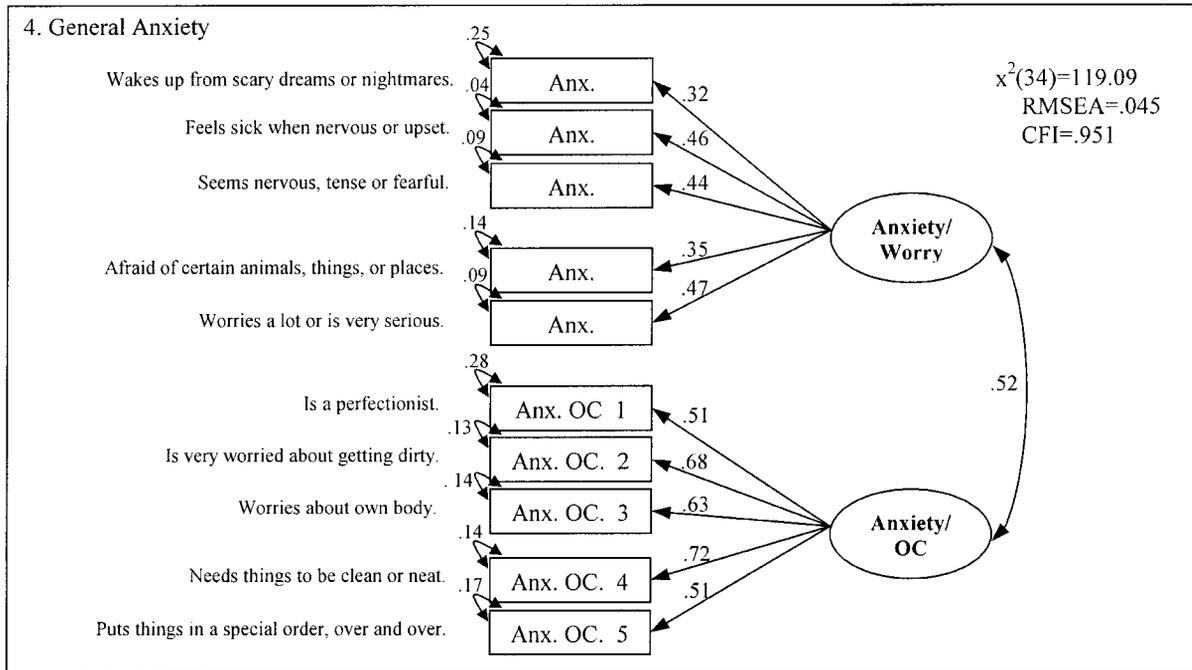
4. General Anxiety

Wakes up from scary dreams or nightmares.
Feels sick when nervous or upset.
Seems nervous, tense or fearful.

Afraid of certain animals, things, or places.
Worries a lot or is very serious.

Is a perfectionist.
Is very worried about getting dirty.
Worries about own body.
Needs things to be clean or neat.
Puts things in a special order, over and over.

Anx.
Anx.
Anx.
Anx.
Anx.

Anx. OC 1
Anx. OC. 2
Anx. OC. 3
Anx. OC. 4
Anx. OC. 5

Anxiety/
Worry

Anxiety/
OC

.25 .04 .32
.09 .46
.44
.14 .35
.09 .47
.28 .51
.13 .68
.14 .63
.14 .72
.17 .51
.52

$x^2(34)=119.09$
RMSEA=.045
CFI=.951

5. Eating

Is a good eater (reversed scored).
Accepts new foods right away (reversed scored).

Is a picky eater.
Refuses to eat.
Refuses to eat certain food(s) for 2 days or more.

Spits out food(s).
Holds food in cheeks.
Refuses to eat foods that require chewing.
Gags or chokes on foods.

Eating 1
Eating 2
Eating 3
Eating 4
Eating 5
Eating
Eating
Eating
Eating

Eating

Eating
Pathology

.19 .30 -.74
-.60
.15 .85
.19 .66
.28 .63
.30 .52
.13 .32
.06 .44
.11 .43
.56

$x^2(26)=122.77$
RMSEA=.055
CFI=.962

Fig. 1. *Continued.*

Total Problem Model

Consistent with our approach of confirming scales and then domains, to evaluate the full theoretical model for problem behaviors a Total Problem Model (Fig. 2) was constructed that included all scales from each of the problem domains assessed by the ITSEA (i.e., Internalizing, Externalizing, and Dysregulation). In the initial model,

Table II. Means (*M*) and Standard Deviations (*SD*) of ITSEA Domains and Scales by Age Group and Child Sex (*N* = 1,235)

	Total sample, *M* (*SD*)	12–17 months (*n* = 286), *M* (*SD*)	18–23 months (*n* = 306), *M* (*SD*)	24–29 months (*n* = 328), *M* (*SD*)	30–36 months (*n* = 315), *M* (*SD*)	*F* value age	Boys (*n* = 605), *M* (*SD*)	Girls (*n* = 630), *M* (*SD*)	*F* value sex
Total problems	0.44 (0.18)	0.43 (0.18)	0.45 (0.18)	0.44 (0.18)	0.45 (0.17)	1.68	0.44 (0.17)	0.44 (0.18)	0.0
1. Externalizing symptoms	0.47 (0.28)	0.45 (0.28)	0.49 (0.28)	0.47 (0.28)	0.49 (0.27)	1.46	0.49 (0.29)	0.46 (0.27)	3.27*
Activity/impulsivity	0.73 (0.43)	0.75 (0.44)	0.74 (0.42)	0.71 (0.43)	0.71 (0.43)	0.63	0.76 (0.43)	0.69 (0.42)	7.52**
Aggression/defiance	0.47 (0.30)	0.39^A (0.30)	0.49^B (0.31)	0.48^B (0.29)	0.53^B (0.30)	10.26***	0.48 (0.32)	0.47 (0.28)	0.01
Peer aggression	0.22 (0.30)	0.19 (0.31)	0.23 (0.31)	0.22 (0.29)	0.23 (0.31)	1.13	0.23 (0.31)	0.21 (0.30)	1.05
2. Internalizing symptoms	0.52 (0.22)	0.49^A (0.21)	0.55^B (0.23)	0.52^{AB} (0.21)	0.51^{AB} (0.22)	3.82**	0.50 (0.21)	0.52 (0.22)	3.43*
Depression/withdrawal	0.06 (0.14)	0.05 (0.13)	0.08 (0.18)	0.06 (0.11)	0.07 (0.13)	1.64	0.06 (0.14)	0.06 (0.13)	0.0
General anxiety	0.26 (0.24)	0.15^A (0.17)	0.26^B (0.24)	0.28^B (0.24)	0.34^C (0.26)	29.59***	0.23 (0.24)	0.27 (0.26)	8.67**
Separation distress	0.86 (0.42)	0.94^A (0.41)	0.93^A (0.44)	0.83^B (0.43)	0.75^B (0.38)	14.16***	0.85 (0.43)	0.87 (0.42)	1.07
Inhibition to novelty	0.88 (0.46)	0.82^A (0.45)	0.93^B (0.46)	0.89^{AB} (0.45)	0.88^{AB} (0.49)	3.07*	0.87 (0.46)	0.90 (0.46)	0.84
3. Dysregulation	0.46 (0.25)	0.45 (0.26)	0.46 (0.25)	0.45 (0.24)	0.48 (0.25)	0.98	0.46 (0.24)	0.46 (0.25)	0.06
Sleep	0.45 (0.47)	0.46 (0.50)	0.44 (0.47)	0.45 (0.44)	0.47 (0.45)	0.21	0.47 (0.48)	0.43 (0.45)	2.30
Negative emotionality	0.52 (0.33)	0.51 (0.31)	0.53 (0.34)	0.51 (0.33)	0.54 (0.33)	0.62	0.52 (0.32)	0.53 (0.33)	0.17
Eating	0.49 (0.35)	0.49 (0.34)	0.50 (0.34)	0.47 (0.35)	0.48 (0.35)	0.35	0.49 (0.33)	0.48 (0.36)	0.01
Tactile sensitivity	0.39 (0.31)	0.36^A (0.31)	0.39^{AB} (0.31)	0.36^A (0.29)	0.43^B (0.33)	3.76*	0.38 (0.31)	0.40 (0.31)	1.27
4. Competence	1.38 (0.29)	1.14^A (0.27)	1.37^B (0.24)	1.47^C (0.26)	1.51^C (0.24)	120.30***	1.33 (0.29)	1.42 (0.28)	33.53***
Compliance	1.26 (0.35)	1.02^A (0.35)	1.29^B (0.32)	1.35^B (0.32)	1.33^B (0.33)	65.98***	1.22 (0.36)	1.29 (0.35)	12.63**
Attention	1.47 (0.41)	1.24^A (0.43)	1.47^B (0.38)	1.56^C (0.38)	1.59^C (0.36)	51.63***	1.44 (0.42)	1.50 (0.40)	8.32**
Imitation/play	1.51 (0.35)	1.31^A (0.35)	1.50^B (0.33)	1.59^C (0.34)	1.63^C (0.31)	55.33***	1.44 (0.37)	1.58 (0.32)	56.38***
Mastery motivation	1.64 (0.30)	1.54^A (0.33)	1.66^B (0.28)	1.68^B (0.29)	1.70^B (0.26)	18.64***	1.63 (0.31)	1.66 (0.28)	2.56
Empathy	1.19 (0.48)	0.83^A (0.45)	1.15^B (0.43)	1.32^C (0.43)	1.41^D (0.38)	106.66***	1.11 (0.48)	1.25 (0.46)	27.84***
Prosocial peer relations	1.20 (0.42)	0.94^A (0.37)	1.13^B (0.40)	1.31^C (0.38)	1.39^C (0.38)	79.61***	1.14 (0.41)	1.25 (0.42)	21.96***
5. Maladaptive	0.11 (0.13)	0.12 (0.13)	0.11 (0.13)	0.10 (0.12)	0.10 (0.12)	1.39	0.10 (0.12)	0.11 (0.13)	2.57
6. Social relatedness	1.71 (0.21)	1.71^{AB} (0.22)	1.73^A (0.21)	1.71^{AB} (0.21)	1.68^B (0.21)	3.08*	1.68 (0.23)	1.73 (0.20)	16.41***
7. Atypical	0.32 (0.25)	0.31^{AB} (0.24)	0.26^A (0.24)	0.32^B (0.24)	0.39^C (0.26)	13.66***	0.32 (0.25)	0.32 (0.25)	0.0

Note: Different letter superscript indicate that row means differ significantly by age.

*$p < .05$. **$p < .01$. ***$p < .001$.

scales were constrained to load only on their individual domain. Because of the general co-occurrence of problem behaviors in this age group, the initial model revealed an expected positive manifold among the correlations (i.e., high intercorrelations among the individual scales across domains resulted in a poorly fitting total model). To remedy this problem, we introduced a General Problems latent construct to the overall model that would draw out the variance shared across all scales, which contrasts with the variance unique to each domain. For this model, we allowed each scale to load on both the General Problems construct and its respective domain construct. By doing so, the common and unique relations among the ITSEA dimensions could be evaluated. Moreover, the model fit was very good when the scales were free to load on both the General Problem latent construct and their respective hypothesized ITSEA problem domain (e.g., Anxiety/Worry was free to load on both the General Problem latent construct and the Internalizing Domain). This model revealed that, after controlling for the shared variance across problem scales with the General Problems latent construct, the ITSEA domains are largely distinct and unique dimensions of early childhood psychopathology.

In contrast to the moderate associations observed between the latent Externalizing and Internalizing and Externalizing and Dysregulation domains, the association between the Internalizing and Dysregulation domains were quite high, possibly because of differences in loading patterns of scales in the individual domains versus Total Problem model (See italicized latent correlations versus underlined Pearson correlations in Fig. 2). For example, in the Total Problem model the relative contribution of Sensory Sensitivities to Dysregulation is inflated as compared with the relative contribution in the individual Dysregulation domain model.

Validity of Problem Domains

To confirm the validity of the ITSEA problem domains and the General Problem latent construct included in the Total Problem Model, individual criterion constructs were added to the Total Problem model to assess their correlations with each of the ITSEA problem domains (see Table III). The following single indicator variables were included as criterion constructs: CBCL/1.5-5 Total
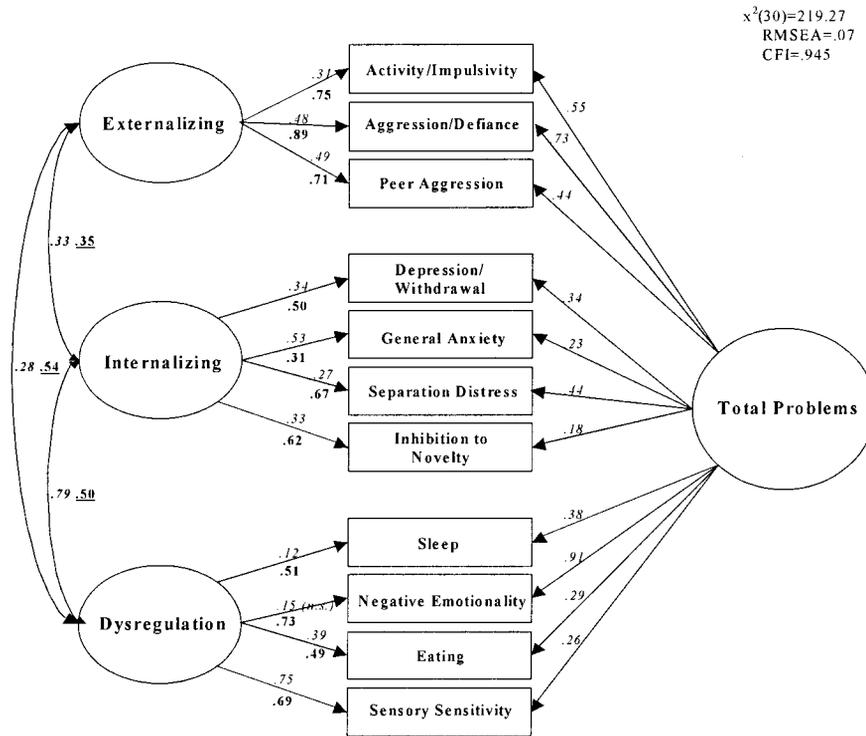
$x^2(30)=219.27$
RMSEA=.07
CFI=.945

Fig. 2. Total ITSEA Problems Model. *Note.* Loadings between observed ITSEA Scales and latent ITSEA Domains are presented in bold (below the arrows) to refer to the loadings from individual domain models whereas loadings in italics refer to the loadings observed in this Total Problems Model. In addition, the correlations among the problems domains in italic are those from the Total Problems Model, whereas those that are underlined are basic Pearson correlations. The fit presented in the figure is the fit for the Total Problems model. The fit for the individual domain models can be found in Table I.

Problem Score; CBCL/1.5-5 Internalizing; CBCL/1.5-5 Externalizing; CCTI Soothability; CCTI Emotionality; CCTI Sociability; BAI; and CESD. Of interest, the latent and measured ITSEA Externalizing domains show highest concordance with the comparable CBCL/1.5-5 domain. In contrast, the latent and measured ITSEA Internalizing domains show high concordance with the CCTI Sociability scale and the CBCL/1.5-5 Internalizing domain. In

Table III. Validity Indices: Associations Between the Ratings on the ITSEA and Ratings on the Child Behavior Checklist/1.5-5, Colorado Child Temperament Inventory (CCTI), Beck Anxiety Scale, and Center for Epidemiology Scale for Depression

	ITSEA domains: Latent correlations				ITSEA domains: Pearson correlations			
Criterion measures	Total problems	Intern.	Extern.	Dysreg.	Intern.	Extern.	Dysreg.	Comp.
CBCL Total[a]	.85**	.39**	.38**	.33**	.47***	.67***	.62***	−.30***
CBCL Internalizing[a]	.74**	.64**	.25**	.51**	.57***	.46***	.57***	−.21***
CBCL Externalizing[a]	.86**	.15**	.44**	.15**	.33***	.73***	.52***	−.33***
CCTI Soothability[b]	−.73**	−.33*	−.16*	−.30**	−.21***	−.31***	−.30***	.28***
CCTI Emotionality	.88**	.27**	.14**	.25**	.37***	.44***	.53***	−.13***
CCTI Sociability	−.47**	−.72**	−.08**	−.55**	−.57***	−.08**	−.23***	.18***
Beck Anxiety Inventory[b]	.63**	.22*	.29*	.18*	.18***	.27***	.26***	−.12***
CES-D[b]	.65**	.37*	.31*	.30*	.26***	.29***	.29***	−.20***

Note. Intern. = ITSEA Internalizing, Extern. = ITSEA Externalizing, Dysreg. = ITSEA Dysregulation, and Comp = ITSEA Competence.
[a] $N = 959$, n's vary slightly because of missing data, and the significance of the latent correlations varies according to the size of their standard errors.
[b] $N = 1,235$, n's vary slightly because of missing data.
* $p < .05.$ ** $p < .01.$ *** $p < .001.$

Table IV. Validity Indices: Pearson Correlations for the ITSEA Ratings With Evaluator Ratings and Standardized Developmental Assessment Data

Criterion measures	N	M	SD	ITSEA domains			
				Internalizing	Externalizing	Dysregulation	Competence[a]
Evaluator ratings							
Internalizing ratings[b]	139	0.10	0.31	.20*	−.01	.18*	−.05
Externalizing ratings	134	0.25	0.53	.07	.28**	.28**	−.18*
Dysregulation ratings[b]	134	0.08	0.24	.19*	.36***	.31**	.12
Competence ratings	121	1.15	1.25	−.16	−.13	−.26**	.29**
Vineland standard scores							
Socialization domain	158	91.3	1.0	−.05	−.21**	−.13	.48***
Communication domain	158	104.5	13.7	−.04	−.30**	−.09	.56***
Vineland adaptive composite	158	97.6	12.2	−.02	−.26**	−.07	.58***
Mullen Standard Scores							
Visual reception domain	166	52.8	12.6	−.02	−.13	−.01	.39***
Expressive language domain	166	51.9	11.2	.01	−.09	−.04	.41***
Composite t score	166	105.5	18.6	−.01	−.16*	−.04	.47***

[a]Child age was covaried from correlations between ITSEA Competence and Evaluator Competence, Mullen scores, and Vineland scores.
[b]Evaluator ratings of Internalizing Problems and Dysregulation excluded separation anxiety and sleep problems, respectively, because evaluators were unable to observe these problems during home visits. Therefore, in correlations between evaluator ratings and ITSEA reports, Separation Distress was excluded from the ITSEA Internalizing score and Sleep was excluded from the ITSEA Dysregulation score.
*$p < .05$. **$p < .01$. ***$p < .001$.

addition, the Total Problems construct showed little differentiation across child temperament, or child and parent distress criterion measures. Indeed, the BAI and CESD were most highly associated with the Total Problems construct, indicating that method variance that is shared across these measures is accounted for by the Total Problems construct, such that the uniquely modeled ITSEA Internalizing, Externalizing, and Dysregulation scores may be independent of such method biases. The pattern of correlations between ITSEA scales and the criterion measures was slightly different when latent and measured ITSEA variables were examined.

Validity of Competence Domain

Pearson correlations were utilized to examine the validity of the ITSEA Competence domain (see Table III). Consistent with expectations, the ITSEA Competence domain correlated negatively with CBCL problem scales, maternal distress, and high emotionality, but correlated positively with CCTI soothability and sociability. None of these correlations were of a magnitude to suggest construct redundancy. In addition, supporting the developmental nature of competence, ITSEA Competence domain correlated significantly with standard scores on the Vineland Socialization Domain, Communication Domain, and Adaptive Behavior Composite, as well as with the Mullen Scales of Early Learning Visual Reception,

Expressive Language, and Composite Standard Scores (see Table IV).

Validity Relative to Evaluator Ratings

As a final validity analysis, ITSEA ratings were compared with evaluator ratings and standardized developmental assessment data from the methodologic substudy (see Table IV). Within-domain correlations between evaluator ratings and parent reports on the ITSEA were significant for every ITSEA domain. Notably, evaluator ratings of Dysegulation correlated significantly with ITSEA reports of problems in all areas and ITSEA Dysregulation correlated with evaluator ratings of Internalizing and Externalizing problems.

DISCUSSION

The results of both confirmatory factor models and traditional psychometric approaches consistently supported the internal consistency of the hypothesized structure of the ITSEA. The strong support for the Internalizing, Externalizing, and Dysregulation problem domains, individually and when allowing for a general problem factor, argues for the early differentiation in parent ratings of problem behaviors. Clearly, by 12–36 months of age,

parents are reporting coherent patterns of internalizing, externalizing, and regulatory problems that cannot be explained on the basis of general concerns about children's behavior. This is in marked contrast to those who have argued that infant and toddler behaviors may be undifferentiated when compared with pre-school-aged and older children (Buss & Plomin, 1975; McGuire & Richman, 1986) and is consistent with more recent studies of young children's behavior (e.g., Mathieson & Sanson, 2000). It is worth noting that, although the correlation between the Internalizing and Dysregulation domains is quite high in the Total Problems model, the corresponding correlation between their measured counterparts (i.e., the Pearson correlation between repeated scores) is lower, supporting the independence of these two dimensions of child behavior problems.

In addition to supporting the presence of distinct and coherent domains of problem behaviors, the best fitting overall model including the three problem domains also included a total problem latent construct that most likely includes variance that can be attributed to overall problem severity, as well as to parental rating distortions. For example, to the extent that parents employ either "yea saying" or "nay saying" biases in descriptions of their children, variance across all problem item ratings will load on a total problem factor. Further support for this interpretation of the total problem construct is gained from evidence that it was strongly associated with maternal depression and anxiety symptoms. Moreover, the general lack of differentiation in the magnitude of correlations between the total problems construct and parent distress, child problem behaviors, and child temperament attributes further supports the interpretation that this construct represents variance due to overall severity of perceived child problems and rating biases. This may present a novel approach for controlling for the effects of rating biases in examining the role of discrete aspects of problem behaviors. However, a limitation of this approach is that in controlling for parental distortion, one also removes variance associated with severity. Further, factor loadings of ITSEA scales on the latent domains change considerably between the individual and Total Problem model, leading to difficulties interpreting the variance that remains within the domain level latent constructs. Scales that are highest loading on the total problems construct (e.g., Negative Emotionality and Aggression/Defiance) show the greatest reduction in loading on their respective individual domain, such that other scales increase in importance. For example, in contrast to the relatively even loadings in the individual Dysregulation model, the Total Problem Dysregulation scale is most strongly influenced by Sensory

Sensitivities. Thus, more work is needed prior to recommending this approach as a solution to the rater bias problem.

The integrity of the three hypothesized problem domains, Externalizing, Internalizing, and Dysregulation, after removing variance attributable to parental rating biases and overall problem severity, provides compelling support for these latent constructs. Moreover, lower correlations among these latent constructs (as well as among bivariate correlations between measured variables) than were observed between the Externalizing and Internalizing domains of the CBCL suggest that they are tapping independent aspects of children's problem behaviors. Additional support for the validity of the ITSEA problem domains emerged in the form of significant correlations with parental reports on the CBCL and parental affective symptoms, as well as independent evaluator ratings of observed child behavior. Similarly, the validity of the ITSEA Competence domain was supported by evidence that ITSEA Competence increased with child age, was correlated significantly with developmental skills assessed with Vineland and Mullen scales, and was associated with evaluator ratings of competence.

Further, with the exception of the large negative correlation observed between the temperament dimension of sociability and the ITSEA Internalizing domain, the low to moderate associations between the ITSEA domains and the CCTI temperament scales provide additional support that the ITSEA is measuring more than temperamental variation. The high correlation between sociability and ITSEA Internalizing is likely due to the inclusion of the Inhibition Scale in this domain. This scale appears to be more temperamentally based than the other problem scales, evidencing a more normal, rather than skewed, distribution. Inhibition is retained on the ITSEA because clinical decisions and longitudinal prediction may be improved through an understanding of the profile of scores across Depression Withdrawal, General Anxiety, Separation Distress, and Inhibition to Novelty.

Another scale that is likely strongly influenced by temperamental predispositions is Negative Emotionality. In previous work, an earlier version of this scale correlated highly with observed negative reactivity and regulation (Carter et al., 1999). Of note, Negative Emotionality was the highest loading scale on the Total Problems latent construct. Moreover, emotional negativity is the only scale on the ITSEA that is often highly, and at a minimum moderately, correlated with all other ITSEA scales. This may be in part due to the fact that young children's behavioral repertoire is more limited and thus

a wide range of psychopathology (i.e., both internalizing and externalizing) is likely to be accompanied by undifferentiated emotional distress (e.g., crying). However, as behavioral repertoires expand with development, or when measured more discretely, the manner in which emotional distress is expressed is likely to be linked to specific forms of psychopathology (e.g., anger with oppositional defiant disorder, sadness with depression; cf Eisenberg et al., 2001). Such a pattern of differentiation with development may be supported by prior evidence that that negative emotionality and/or difficult temperament in infancy and toddlerhood predict later psychopathology in both the internalizing and externalizing domains (Earls & Jung, 1987; Prior et al., 1992; Rende, 1993; Shaw et al., 2001; Shaw & Vondra, 1995).

The next most highly loading scale on the Total Problem construct was Aggression/Defiance. This finding is consistent with prior reports indicating that parental distress is associated with elevated reports of children's behavior problems, and particularly externalizing problems (Briggs-Gowan et al., 1996; Najman et al., 2001). In the absence of appropriate measures of temperament that tap aspects of attention and persistence, we were not able to assess expected associations between this aspect of temperament and the ITSEA Attention scale. In summary, we view the Negative Emotionality, Inhibition, and Attention scales as the most likely to show high associations with existing infant and toddler temperament scales.

Conclusions about the validity of the ITSEA must be tempered until adequate clinical samples can be carefully assessed with a variety of methods that are contrasted with the results of parent and childcare reports on the ITSEA. However, the consistency of reporting found across informants within method (i.e., mothers, fathers), as well as moderate correlations between competencies and standardized developmental assessments, and small correlations between parents and independent observers, provide strong preliminary validity for parent ratings on the ITSEA. In addition to clinical data, further longitudinal analyses within this cohort will strengthen our understanding of the validity of these constructs in this early age range. For example, while there is considerable evidence of stability within the domain of externalizing problems, particularly for aggression (Cummings, Ianotti, & Zahn-Waxler, 1989; Moffit, 1993; Nagin & Tremblay, 1999; Pierce, Ewing, & Campbell, 1999), less is known about the stability of early emerging internalizing and regulatory problem behaviors. In addition, multi method, multi-informant modeling (Loehlin, 1992) will illuminate the utility of the Total Problem Factor with respect to diminishing the problem of bias when using parental reports.

In addition to examining psychometric properties of the ITSEA, several important child age and sex patterns emerged. Specifically, along with latent and raw mean differences across child sex and the four 6-month age bands examined here, the factor structure of the ITSEA scales were invariant. Thus, the scales are appropriate throughout this age range. This is critical for studies designed to address continuity and change in the early course of problem behaviors and competencies. For example, this structural invariance across age indicates that the ITSEA would be appropriate to document change associated with early intervention programs. At the same time, the observed age and sex mean level differences reflect the importance of comparing young children within relatively narrow age bands, as well as to children of their own sex, when attempting to profile problems and competencies relative to normative patterns of functioning. That girls appear to acquire competence earlier than boys may have implications for understanding later pathways with respect to disruptive disorders (Keenan & Shaw, 1997) as well as other social–emotional and behavioral problems. Similarly, this study lends support to the early emergence of sex differences, with boys rated higher on activity and girls rated higher on anxiety.

The ITSEA offers a new parent-report instrument for evaluating the early emergence of problems and competencies. Although there has been some reluctance to acknowledge the presence of problems in very young children, many infants and toddlers exhibit serious mental health problems (cf. Del Carmen-Wiggins & Carter, 2001; Zeanah, 2000). In light of the fact that one in every five students in the United States has significant mental health problems that require attention (U.S. Department of Health and Human Services, 1999), early detection and intervention (i.e., prior to school entry) must be pursued to promote successful school adaptation (U.S. Department of Health and Human Services, 1999). The ITSEA provides a new tool for learning about the very early course of social–emotional and behavioral problems and competencies. Used in conjunction with other measures that assess the broader context of risk factors and resources in the child's ecology, the ITSEA can enhance (1) the identification of 1- to 3-year-old children currently suffering and those at highest risk for later disturbance; and (2) the evaluation of prevention and early intervention efforts. Assessment of competencies, as well as problems, is crucial to evaluate fully early intervention and prevention programs that adopt strength-based approaches.

APPENDIX

ITSEA Higher-Order Domains, Component Constructs and Selected Individual Items

1. Externalizing Domain (24 items)

Activity/impulsivity (6 items)
Is restless and can't sit still.
Gets hurt so often that you can't take your eyes off him/her

Aggression/defiance (12 items)
Is disobedient or defiant. For example, refuses to do
as you ask.
Acts bossy.
Misbehaves to get attention from adults.
Hits, bites, or kicks you (or other parent).
Is destructive. Breaks or ruins things on purpose.

Peer aggression (6 items)
Won't let other children play with his/her group.
Teases other children.
Hits, shoves, kicks, or bites other children.

2. Internalizing Domain (30 items)

Depression/withdrawal (9 items)
Seems withdrawn.
Avoids physical contact.
Has less fun than other children.
Seems very unhappy, sad, or depressed.

General anxiety (10 items)
Wakes up from scary dreams or nightmares.
Seems nervous, tense or fearful.
Afraid (composite of 3 fear items pertaining to
certain animals, things, or places).
Is a perfectionist.
Is very worried about getting dirty.

Separation distress (6 items)
Cries or hangs onto you when you try to leave.
Gets upset when left with a familiar babysitter
or relative
Hangs on you or wants to be in your lap when
with other people.

Inhibition to novelty (5 items)
Takes a while to feel comfortable in new places
(10 minutes or more).
Is shy with new children.
Is quiet or less active in new situations.

3. Dysregulation Domain (34 items)

Sleep (5 items)
Must be held to go to sleep.
Wakes up at night and needs help to fall asleep
again.
Has trouble falling asleep or staying asleep.

Negative emotionality (13 items)
Wakes up grouchy or in a bad mood.
Gets angry or pouts.
Is impatient or easily frustrated.
Has trouble adjusting to changes.
Is hard to soothe when upset.
Cries or tantrums until s/he is exhausted.

Eating (9 items)
Accepts new foods right away (Reversed).
Is a picky eater.
Refuses to eat certain food(s) for 2 days or more.
Spits out food(s).
Refuses to eat foods that require chewing.

Sensory sensitivity (7 items)
Is bothered by loud noises or bright lights.
Is bothered by being in motion (e.g., swinging, etc.).
Is bothered by how some things feel on his/her skin.
Is bothered by certain odors (smells).

4. Competence Domain (37 items)

Attention (5 items)
Can pay attention for a long time (Not including TV).
Plays with toys for 5 minutes or longer.
Pays careful attention when being taught
something new.

Compliance (8 items)
Helps with dressing (e.g., puts arm in sleeve).
Tries to do as you ask.
Puts toys away after playing.
Obeys when asked to stop being aggressive.

Imitation/play (6 items)
Hugs or feeds dolls or stuffed animals.
Imitates clapping or waving "bye-bye."
Pretends that objects are something else.

Mastery motivation (6 items)
Enjoys challenging activities.
Shows pleasure when s/he succeeds (e.g., claps
for self).
Keeps trying even when something is hard.

Empathy (7 items)
Tries to "make up" after misbehaving.
Tries to help when someone is hurt (e.g., gives a toy).
Is worried or upset when someone is hurt.
Talks about other people's feelings (like
"Mommy mad").

Prosocial peer relations (5 items)
Plays well with other children.
Takes turns when playing with others.
Asks for things nicely when playing with children.

5. Maladaptive Index (13 items)
Has body tic or twitch s/he seems unable to
control.

6. Social Relatedness Index (10 items)
Likes being cuddled, hugged or kissed by loved
ones.

Appendix (*continued.*)

Makes sounds s/he seems unable to control.	Smiles back at you from across a room.
When upset gets very still, freezes or doesn't move.	Is interested in other babies and children.
Acts out the same pretend theme over and over.	Reaches for you when you are not holding him/her.
Has bowel movements where s/he shouldn't (like on the floor).	Looks for you when upset.
Plays with own sex parts often and for a long time.	Looks right at you when you say his/her name.
Eats or drinks things that are not edible, like paper or paint.	

7. Atypical Index (8 items)

Puts things in a special order, over and over.
Repeats a particular movement, over and over.
Without looking at you puts your hand on objects, such as wind-up toys, to make them work.
"Spaces out." Is totally unaware of what's happening around him/her.
Points to show you something far away. (Reverse scored.)

8. Individual Items of Clinical Significance (currently not part of a scale) (10 items)

Wakes up screaming and doesn't respond to you for a few minutes.
Runs away in public places.
Is affectionate with strangers.
Does not react when hurt.
Is not afraid when should be.
Gets upset when asked to change activities.
Hurts self on purpose. For example, bangs head.

ACKNOWLEDGMENTS

REFERENCES

Achenbach, T. M. (1966). The classification of children's psychiatric symptoms: A factor-analytic study. *Psychological Monographs*, *80*, 615.

Achenbach, T. M. (1992). *Manual for the Child Behavior Checklist/2-3 & 1992 Profile*. Burlington: University of Vermont.

Achenbach, T. M. (2000). Assessment of psychopathology. In A. Sameroff & M. Lewis (Eds.), *Handbook of developmental psychopathology* (pp. 549–566). New York: Plenum.

Achenbach, T. M., Edelbrock, C., & Howell, C. (1987). Empirically-based assessment of the behavioral/emotional problems of 2–3 year old children. *Journal of Abnormal Child Psychology*, *15*, 629–650.

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.

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

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Anderson, D. R. (1983). Prevalence of behavioral and emotional disturbance and specific problem types in a sample of disadvantaged preschool-aged children. *Journal of Clinical Child Psychology*, *12*, 130–136.

Beck, A. T., Epstein, N., Brown, G., & Steer, R. A. (1988). An inventory for measuring clinical anxiety: Psychometric properties. *Journal of Consulting and Clinical Psychology*, *56*, 893–897.

Bollen, K. A., & Long, J. S. (Eds.). (1993). *Testing structural equation models*. Newbury Park, CA: Sage.

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 Journal*, *19*, 422–445.

Briggs-Gowan, M., 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 Child Psychology*, *24*, 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-olds. *Journal of the American Academy of Child and Adolescent Psychiatry*, *40*, 811–819.

Browne, M. W., & Arminger, G. (1995). Specification and estimation of mean- and covariance-structure models. In G. Arminger, C. C. Clogg, & M. E. Sobal (Eds.), *Handbook of statistical modeling for the social and behavioral sciences* (pp. 185–249). New York: Plenum.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Newbury Park, CA: Sage.

Buss, A. H., & Plomin, R. (1975). *A temperament theory of personality*. New York: Wiley.

Campbell, S. B. (1990). *Behavior problems in preschool children: Clinical and developmental issues*. New York: Guilford.

Campbell, S. B. (1995). Behavior problems in preschool children: A review of recent research. *Journal of Child Psychology and Psychiatry*, *36*, 113–149.

Campbell, S. B., & Ewing, L. J. (1990). Hard to manage preschoolers: Adjustment at age nine and predictors of continuing symptoms. Journal of Child Psychology and Psychiatry, *31*, 871–889.

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. (1993). *The Infant–Toddler Social and Emotional Assessment* (ITSEA). Unpublished Measure. Yale University Department of Psychology, New Haven, CT. University of Massachusets Boston Department of Psychology, Boston, MA. Available by request at alice.carter@umb.edu.

Carter, A.S., & Briggs-Gowan, M. (2000). *The Infant–Toddler Social and Emotional Assessment* (ITSEA). Unpublished Manual. University of Massachusetts Boston Department of Psychology, Boston, MA. Yale University, New Haven, CT. Available by request at alice.carter@umb.edu

Carter, A. S., Garrity-Rokous, F. E., Chazan-Cohen, R., Little, C., & Briggs-Gowan, M. J. (2001). Maternal depression and comorbidity: Predicting early parenting, attachment security, and toddler social–emotional problems and competencies. *Journal of the American Academy of Child and Adolescent Psychiatry, 40,* 18–36.

Carter, A. S., Grigorenko, E. L., & Pauls, D. L. (1995). A Russian adaptation of the Child Behavior Checklist: Psychometric properties and associations with child and maternal affective symptomatology and family functioning. *Journal of Abnormal Child Psychology, 23,* 661–684.

Carter, A. S., Little, C., Briggs-Gowan, M. J., & Kogan, N. (1999). The Infant - Toddler Social and Emotional Assessment (ITSEA): Comparing parent ratings to laboratory observations of task mastery, emotional regulation, coping behaviors and attachment status. *Infant Mental Health Journal, 20,* 375–392.

Cicchetti, D., & Cohen, D. J. (1995). Perspectives on developmental psychopathology. In D. Cicchetti & D. J. Cohen (Eds.), *Developmental psychopathology: Vol. I. Theory and methods* (pp. 3–20). New York: Wiley.

Cornely, P., & Bromet, E. (1986). Prevalence of behavior problems in three-year-old children living near Three Mile Island: A comparative analysis. *Journal of Child Psychology and Psychiatry, 27,* 489–498.

Cummings, E. M., Iannotti, R., & Zahn-Waxler, C. (1989). Aggression between peers in early childhood: Individual continuity and developmental change. *Child Development, 60,* 887–895.

DelCarmen-Wiggins, R., & Carter, A. S. (2001). Introduction: Assessment of infant mental health: Advances and challenges. *Journal of the American Academy of Child and Adolescent Psychiatry, 40,* 8–10.

Downey, G., & Coyne, J. C. (1990). Children of depressed parents: An integrative report. *Psychological Bulletin, 108,* 50–76.

Earls, F. (1980). Prevalence of behavior problems in 3-year-old children: A cross-national replication. *Archives of General Psychiatry, 37,* 1153–1157.

Earls, F., & Jung, K. G. (1987). Temperament and home environment characteristics as causal factors in the early development of childhood psychopathology. *Journal of the American Academy of Child and Adolescent Psychiatry, 26,* 491–498.

Easterbrooks, A., & Goldberg, W. (1990). Security of toddler–parent attachment: Relation to children's socio-personality functioning during kindergarten. In M. T. Greenberg, D. Cicchetti, & E. M. Cummings (Eds.), *Attachment in the preschool years* (pp. 221–224). Chicago: University of Chicago Press.

Eisenberg, N., Cumberland, A., Spinrad, T. L., Fabes, R. A., Shepard, S. A., Reiser, M., et al. (2001). The relations of regulation and emotionality to children's externalizing and internalizing problem behavior. *Child Development, 71,* 1112–1134.

Eisenberg, N., & Mussen, P. H. (1989). *The roots of prosocial behavior in children.* New York: Cambridge University Press.

Erickson, M. F., Sroufe, L. A., & Egeland, B. (1985). The relationship between quality of attachment and behavior problems in preschool in a high-risk sample. *Monographs of the Society for Research in Child Development, 50,* 147–166.

Fischer, M., Rolf, J. E., Hasazi, J. E., & Cummings, L. (1984). Follow-up of a preschool epidemiologic sample: Cross-age continuities and predictions of later adjustment with internalizing and externalizing dimensions of behavior. *Child Development, 55,* 137–150.

Gelfand, D. M., & Teti, D. M. (1990). The effects of maternal depression on children. *Clinical Psychology Review, 10,* 329–353.

Goldsmith, H. H., & Lemery, K. S. (2000). Linking temperamental fearfulness and anxiety symptoms: A behavior-genetic perspective. *Biological Psychiatry, 48,* 1199–1209.

Ireton, H. (1992). *Child Development Inventory.* Minneapolis, MN: Behavior Science Systems.

Joreskog, K. G. & Sorbom, D. (1999). *LISREL 8.3 user's reference guide.* Scientific Software, Inc., Chicago, IL.

Keenan, K., & Shaw, D. (1997). Developmental and social influences on young girls' early problem behavior. *Psychological Bulletin, 121,* 95–113.

Koot, H. M., & Verhulst, F. C. (1991). Prevalence of problem behavior in Dutch children aged 2–3. *Acta Psychiatric Scandinavica: Supplementum, 83,* 37.

Little, T. D. (1997). Mean and covariance structures (MACS) analyses of cross-cultural data: Practical and theoretical issues. *Multivariate Behavioral Research, 32,* 53–76.

Loehlin, J. C. (1992). *Latent variable models.* Hillsdale, NJ: Erlbaum.

Luk, S., Leung, P. W., Bacon-Shone, J., & Lieh-Mak, F. (1991). The structure and prevalence of behavioral problems in Hong Kong preschool children. *Journal of Abnormal Child Psychology, 19,* 219–232.

Lyon, M. E., & Plomin, R. (1981). The measurement of temperament using parental ratings. *Journal of Child Psychology and Psychiatry, 22,* 47–53.

Marsh, H. W., Balla, J. R., & McDonald, R. P. (1988). Goodness-of-fit indexes in confirmatory factor analysis: The effect of sample size. *Psychological Bulletin, 103,* 391–410.

Masten, A. S., & Coatsworth, J. D. (1995). Competence, resilience, and psychopathology. In D. Cicchetti & D. J. Cohen (Eds.), *Developmental psychopathology: Vol. I. Theory and methods* (pp. 23–56). New York: Wiley.

Masten, A. S., & Coatsworth, J. D. (1998). The development of competence in favorable and unfavorable environments: Lessons form research on successful children. *American Psychologist, 53,* 205–220.

Mathiesen, K. S., & Sanson, A. (2000). Dimensions of early childhood behavior problems: Stability and predictors of change from 18 to 30 months. *Journal of Abnormal Child Psychology, 28,* 15–31.

McGuire, J., & Richman, N. (1986). Screening for behaviour problems in nurseries: The reliability and validity of the Preschool Behaviour Checklist. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 27,* 455–472.

Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika, 58,* 525–543.

Moffit, T. E. (1993). Adolescence-limited and life-course persistent antisocial behavior: A developmental taxonomy. *Psychological Review, 100,* 674–701.

Mullen, E. M. (1989). *Infant Mullen Scales of early learning.* Circle Pines, MN: American Guidance Service.

Murray, L. (1992). The impact of postnatal depression on infant development. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 33,* 543–561.

Nagin, D., & Tremblay, R.E. (1999). Trajectories of boys' physical aggression, opposition, and hyperactivity on the path to physically violent and non-violent juvenile delinquency. *Child Development, 70,* 1181–1196.

Najman, J. M., Williams, G. M., Nikles, J., Spence, S., Bor, W., O'Callaghan, M., et al. (2001). Bias influencing maternal reports of child behaviour and emotional state. *Social Psychiatry and Psychiatric Epidemiology, 36,* 186–194.

Newth, S. J., & Corbett, J. (1993). Behavior and emotional problems in three-year-old children of Asian parentage. *Journal of Child Psychology and Psychiatry, 34,* 333–352.

Offord, D. R., Boyle, M. H., & Racine, Y. (1989). Ontario Child Health Study: Correlates of disorder. *Journal of the American Academy of Child and Adolescent Psychiatry*, *28,* 856–860.

Pierce, E. W., Ewing, L. J., & Campbell, S. B. (1999). Diagnostic status and symptomatic behavior of hard-to-manage preschool children in middle childhood and early adolescence. *Journal of Clinical Child Psychology*, *28,* 44–57.

Plomin, R. (1976). A twin and family study of personality in young children. *Journal of Psychology*, *94,* 233–235.

Prior, M., Smart, D., Sanson, A., Pedlow, R., & Oberklaid, F. (1992). Transient versus stable behavior problems in a normative sample: Infancy to school age. *Journal of Pediatric Psychology*, *17,* 423–443.

Radke-Yarrow, M., Nottelmann, E., Martinez, P., Fox, M. B., & Belmont, B. (1992). Young children of affectively ill parents: A longitudinal study of psychosocial development. *Journal of the American Academy of Child and Adolescent Psychiatry*, *31,* 68–77.

Radke-Yarrow, M., & Zahn-Waxler, C. (1984). Roots, motives, and patterns and children's prosocial behavior. In E. Staub, D. Bar-Tal, J. Karylowski, & J. Reykowski (Eds.), *The development and maintenance of prosocial behavior: International perspectives on positive morality* (pp. 81–99). New York: Plenum.

Radloff, L. S. (1977). The CES-D Scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*, *1,* 385–401.

Rende, R. D. (1993). Longitudinal relations between temperament traits and behavioral syndromes in middle childhood. *Journal of the American Academy of Child and Adolescent Psychiatry*, *32,* 287–290.

Richman, N., Stevenson, J. E., & Graham, P. J. (1975). Prevalence of behavior problems in 3-year-old children: An epidemiological study in a London borough. *Journal of Child Psychology and Psychiatry*, *16,* 277–287.

Richman, N., Stevenson, J., & Graham, P. J. (1982). *Pre-school to school: A behavioral study*. London: Academic Press.

Rothbart, M. K. (1981). Measurement of temperament in infancy. *Child Development*, *52,* 569–578.

Rothbart, M. K., & Putnam, S. P. (2002). Temperament and socialization. In L. Pulkkinen, & A. Caspi, (Eds.), *Paths to successful development: Personality in the life course* (pp. 19–45). New York: Cambridge University Press.

Rowe, D. C., & Plomin, R. (1977). Temperament in early childhood. *Journal of Personality Assessment*, *41,* 150–156.

Saarni, C. (1988). Emotional competence: How emotions and relationships become integrated. In R.A. Thompson (Ed.), *Nebraska Symposium on Motivation: Socioemotional development: Vol. 36 Current theory and research in motivation*, (pp. 115–182). Lincoln: University of Nebraska Press.

Sameroff, A. J., & Emde, R. N. (Eds.). (1989). *Relationship disturbances in early childhood: A developmental approach*. New York: Basic Books.

Shaw, D. S., Keenan, K., & Vondra, J. I. (1994). Developmental precursors of externalizing behavior: Age 1 to 3. *Developmental Psychology*, *30,* 355–364.

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

Shaw, D. S., & Vondra, J. I. (1995). Infant attachment security and maternal predictors of early behavior problems: A longitudinal study of low-income families. *Journal of Abnormal Child Psychology*, *23,* 335–357.

Sparrow, S. S., Balla, D. A., & Cicchetti, D. (1984). *Vineland Adaptive Behavior Scales: Expanded Form*. Circle Pines, MN: American Guidance Service.

Thompson, M. J. J., Stevenson, J., Sonuga-Barke, E., Nott, P., Bhatti, Z., Price, A., et al. (1996). Mental health of preschool children and their mothers in a mixed urban/rural population: I. Prevalence and ecological factors. *British Journal of Psychiatry*, *168,* 16–20.

Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika*, *38,* 1–10.

U. S. Department of Health and Human Services. (1999). *Mental health: A report of the Surgeon General—Executive summary*. Rockdale, MD: U.S. Department of Health and Human Services Administration, Center for Mental Health Services, National Institutes of Health, National Institute of Mental Health.

Zahn-Waxler, C., Radke-Yarrow, M., Wagner, E., & Chapman, M. (1992). Development of concern for others. *Developmental Psychology*, *28,* 126–136.

Zeanah, C. H. (Ed.). (2000). *Handbook of infant mental health* (2nd ed.). New York: Guilford.

Zero to Three, National Center for Infant, Toddlers, and Families. (1994). *Diagnostic classification: 0–3. Diagnostic classification of mental health and developmental disorders of infancy and early childhood*. Washington, DC: Author.