Cornerstones

Identification of Infant and Toddler Social-Emotional Disorders Using the *DC: 0-3 Diagnostic Classification System*

Carl J. Dunst
Andrea Storck
Donna Snyder

The reliability, validity, and accuracy (diagnostic efficiency) of the *Diagnostic Classification: 0-3* (DC: 0-3) *Diagnostic Classification of Mental Health and Developmental Disorders of Infancy and Early Childhood* was the focus of this research synthesis. Fifteen studies including 2,065 young children administered the DC: 0-3 were examined to ascertain interrater and test/retest reliability, convergent/divergent validity, and the sensitivity and specificity of the assessment and diagnostic procedures. Information on the psychometric properties of the DC: 0-3 were found to be quite limited, and the information that was available suggested that better designed and implemented psychometric studies are needed to justify the use of the diagnostic classification system for early identification or eligibility determination purposes.

Purpose

The primary purpose of this practice-based research synthesis was to ascertain the reliability and validity of the Zero-to-Three *Diagnostic Classification of Mental Health and Developmental Disorders of Infancy and Early Childhood* (Lieberman, Wieder, & Fenichel, 1997; Zero to Three: National Center for Infants, 1994). The secondary purpose was to ascertain the accuracy of classificatory assignments of children to different diagnostic categories (Cantwell, 1996). As stated in the administration manual, the DC: 0-3 classification system gives a "timely assessment and accurate diagnosis . . . of mental health and developmental difficulties in the first four years of life . . . can provide the foundation for effective intervention before early deviations become consolidated into maladaptive patterns of functioning" (Zero to Three: National Center for Infants, 1994, p. 3).

The conduct of the synthesis was guided by a framework that focused on the degree to which operationally defined characteristics of a practice are related to desired outcomes (Dunst, Trivette, & Cutspec, 2002). More specifically, we examined available research to discern the reliability and validity of DC: 0-3 diagnostic classifications as well as the diagnostic efficiency (accuracy) of the assignment of children to different diagnostic categories. The particular reliability, validity, and accuracy data constituting the focus of analysis were considered minimally necessary to warrant the use of the DC: 0-3 for early identification of infant and toddler social-emotional disorders as part of IDEA Part C early intervention eligibility determination (Shackelford, 2006).

Background

The development of the DC: 0-3 classification system was begun in 1987 by a multidisciplinary task force established by the Zero to Three: National Center for Infants, Toddlers, and Families. The task force met twice a year for 3 years and communicated with one another during this time to "build a data base…[that] served as the foundation for case discussion and for the identification of recurring patterns of behavioral problems" (Zero to Three: National Center for Infants, 1994, p. 4) that became the basis for descriptive categories of mental-health-related diagnostic classifications. The task force continued to meet and further developed the diagnostic categories that became the classification system first published in 1994 (Lieberman et al., 1997; Zero to Three: National Center for Infants, 1994) and that has recently been revised (Zero to Three: National Center for Infants, 2005).

The DC: 0-3 is a multiaxial classification system that includes five diagnostic and developmental categories: Axis I–Primary classification, Axis II–Relationship classification, Axis III–Physical, neurological, developmental, and mental health disorders or conditions, Axis IV–Psychosocial stressors, and Axis V–Functional emotional developmental level. The organization of the DC: 0-3 is similar to other multiaxial classification systems although as noted in the administration manual "the axes in [the DC: 0-3] are not intended to be entirely symmetrical with other such systems as DSM IV and ICD-10 because this system, in dealing with infants and young children, focuses on developmental issues. Dynamic processes, such as relationship and developmentally based conceptualizations of adaptive patterns (i.e., functional emotional developmental level) are therefore of central importance" (Zero to Three: National Center for Infants, 1994, pp. 14-15).

Description of the Practice

The DC: 0-3 includes both assessment procedures and diagnostic classification based on evaluation findings. As stated in the administration manual, "A full evaluation usually requires a minimum of three to five sessions of 45 or more minutes each. A complete evaluation will usually involve taking the history, direct observation of functioning (i.e., of family and parental dynamics, caregiver-infant relationship and interaction patterns, the infant's constitutional—maturational characteristics, and language, cognitive and affective patterns), and hands on interactive assessment of the infant, including assessment of sensory reactivity and processing, motor tone and planning, language, cognition, and affective expression" (Zero to Three: National Center for Infants, 1994, p. 13). The use of standardized instruments is recommended only to the extent that they yield information about specific concerns. A single instrument, the Parent-Infant Relationship Global Assessment Scale (PIR-GAS), is included in the administration manual and is used to discern whether an Axis II relationship disorder classification is warranted. Clinical guidelines are included throughout the administration manual for aiding in the assessment and classification process for the other axes.

Guidelines for using assessment findings and evaluation results for diagnostic classification are included in both the DC: 0-3 administration manual (Zero to Three: National Center for Infants, 1994) and casebook (Lieberman et al., 1997; see also Weston et al., 2003). Table 1 summarizes the major diagnostic categories for each DC: 0-3 axis and the focus of assessment on each axis. The reader is referred to the administration manual and casebook for more detailed descriptions of the criteria used to make classifications. See especially *Guidelines to Selecting the Appropriate Diagnosis* (pp. 16-18) in the administration manual (Zero to Three: National Center for Infants, 1994).

Psychometric Properties

Information on the reliability and validity of the DC: 0-3 as well as the accuracy of diagnostic classifications is noticeably missing in both the administration manual and casebook. There is, however, reference to these aspects of assessment and classification in both sources. For example, it is noted that "expert clinicians can reliably agree on primary diagnosis using the system proposed in DC: 0-3" (Zero to Three: National Center for Infants, 1994, p. 18). Similarly, the claim is made that "extensive use [of the DC: 0-3]…should add to the system's already established clinical usefulness and teach us more about the validity and reliability of the specific diagnostic categories" (Lieberman et al., 1997, p. 15). The reason for the lack of reliability, validity, and accuracy data can be found in the administration manual. In the introductory chapter to the manual it is noted that:

> In any scientific enterprise, but particularly in a new field, a healthy tension exists between the desire to analyze findings from systematic research before offering even initial conceptualizations, and the need to disseminate preliminary conceptualizations so that they can serve as a basis for collecting systematic data, which can lead to more empirically based efforts.… The development of *Diagnostic Classification: 0-3* represents an important first step: the presentation of expert consensus-based categorizations of mental health and developmental disorders in the early years of life. (Zero to Three: National Center for Infants, 1994, p. 11)

A tremendous amount of work and effort went into de-

veloping the DC: 0-3 classification system. Notwithstanding the task force's accomplishments, the need for psychometric data is especially important when a diagnostic classification is used as the basis of "a comprehensive treatment or preventive intervention plan" (Zero to Three: National Center for Infants, 1994, p. 13) and for making decisions about people's lives (Becker, n.d.). Based on available research at the time, a special issue of the *Infant Mental Health Journal* was published on the DC: 0-3, Emde and Wise (2003) noted that it is "disappointing that we have not been able to go further with information about its research related criteria of reliability and validity" (p. 437) in the 10 years since the classification system was first published (see also Skovgaard, Houmann, Landorph, & Christiansen, 2004).

The research synthesis described in this *Cornerstones* report examined available data on the reliability and validity of the DC: 0-3 assessment and classification system. Reliability was examined in terms of two (or more) persons making the same diagnostic classification (interrater reliability) and whether the same classificatory assignments were made on different occasions (test/retest reliability or stability). Validity was examined in terms of the extent to which diagnostic classifications were related to some external standard or criterion (convergent validity) and whether children with or without a mental health disorder, or children with different disorders, in fact were assigned to the correct groups (discriminant validity). These particular reliability and validity measures were considered the minimum necessary for the DC: 0-3 to be used for diagnostic, clinical, and intervention purposes (Anastasi & Urbina, 1997). Accuracy, or diagnostic efficiency, was examined in terms of the sensitivity and specificity of classificatory assignments (Medical Algorithms Project, 2006; Meyer, 2003). Sensitivity was determined by examining the extent to which children with different presenting conditions were correctly assigned to a diagnostic category. Specificity was determined by examining the extent to which children with no presenting conditions were not given a diagnosis.

The need for a research synthesis of the psychometric properties of the DC: 0-3 is based, in part, on the recommendations of a number of early childhood intervention specialists who have suggested that the diagnostic system should prove useful for IDEA Part C early intervention eligibility determination and the early identification of social-emotional disorders in infants and toddlers (e.g., Bagnato & Neisworth, 1999; Evangelista & McLellan, 2004; Gomez, Baird, & Jung, 2004; Thomas & Clark, 1998; Weston et al., 2003). The need for a systematic review and analysis of the psychometric properties of the DC: 0-3 is further warranted by the current interest in early childhood mental health and the provision of IDEA Part C early intervention to infants and toddlers with mental health problems (Foley & Hochman, 2006; Oser, 2004; Osofsky & Fitzgerald, 2000; Thomas, Benham, & Guskin, 2000) and the implicit or explicit endorsement of the use of the DC: 0-3 by IDEA Part C early intervention programs (e.g., Waisman Center, 2006).

Search Strategy

Search Terms

Identification of relevant studies was accomplished using the key words *diagnostic classification*, *DC: 0-3 classification system*, *DC: 0-3*, *diagnostic classification: 0-3*, and *diagnostic classification of mental health and developmental disorders of infancy and early childhood*, *Zero to Three diagnostic classification*. Subject heading searches were conducted as Boolean searches AND conditions, using developmental disabilities OR disabilities OR mental health AND infants OR toddlers OR children OR preschoolers as search terms.

Sources

The primary databases searched for relevant studies were: Psychological Abstracts online (PsycINFO), Social Sciences Citation Index (SSCI), Educational Resource Information Center (ERIC) database, MEDLINE, Cumulative Index to Nursing and Allied Health Literature (CINAHL), Health Source: Nursing/Academic Edition, Cochrane databases, Academic Search Elite, InfoTrac Expanded Academic ASAP, Dissertation Abstracts International, National Technical Information Services, REHABDATA, CIRRIE, WorldCat and Ingenta. Online searches using the Google search engine and Google Scholar were also conducted. In addition, we reviewed the reference lists in the DC: 0-3 administration manual (Zero to Three: National Center for Infants, 1994, 2005) and casebook (Lieberman et al., 1997), and the reference sections of all articles located through the primary search, including those in a special issue of the *Infant Mental Health Journal* (Guédeney & Maestro, 2003), to identify additional studies.

Selection Criteria

Studies were included in this synthesis if the DC: 0-3 was used to assign children to diagnostic classifications on any of the five diagnostic axes and sufficient information was included to discern assessment and classification procedures. Studies were also included if any of the psychometric properties constituting the focus of this synthesis or any other reliability or validity data were included in a study. Case studies were excluded because of the particular psychometric properties that were the focus of analysis.

Search Results

Fifteen (15) research reports were located that included 18 samples of children administered the DC: 0-3. Four samples (22%) had no presenting problems or concerns (nonclinical samples) and 14 samples (78%) had some type of presenting problems or referral concerns at the time the DC: 0-3 was administered (clinical samples). One of the four nonclinical samples (Reams, 1999) had no mental health related presenting conditions but was considered at risk for socioenvironmental reasons. Table 2 shows selected characteristics of the study participants.

Participants

The 15 studies included 2,065 children birth-to-58 months of age where the majority (91%) of the children were birth to 48 months of age.[1] Children's gender was reported in all but three studies. Sixty (60%) percent of the children were male and 40% were female.

The samples were from studies conducted in the United States (55%), Israel (16%), Portugal (11%), Austria (6%), Denmark (6%), and France (6%). The ethnic or racial backgrounds of the study participants were reported in all but three studies and included children who were Caucasian or White (34%), Israeli (26%), Portuguese (25%), African American (5%), multiracial or multiethnic (5%), and Latino (2%). A small percentage (3%) of the study participants had other ethnicities or racial backgrounds.

Assessment Procedures

Table 3 shows the assessment procedures that were used for making diagnostic classifications, the diagnosticians, and the degree to which the assessment procedures described in published papers were consistent with those recommended in the DC: 0-3 manual (Zero to Three: National Center for Infants, 1994). The latter was discerned by determining how many of the 14 recommended methods and procedures described in the DC: 0-3 manual were used for diagnostic classification. The use of eight or more assessment procedures was considered highly consistent, the use of four to seven procedures was considered moderately consistent, and the use of three or fewer procedures was considered minimally consistent with the recommended approach to assessment and diagnosis. The assessment procedures were considered highly consistent with the DC: 0-3 recommended procedures in 3 studies (20%), moderately consistent in 7 studies

(47%), and minimally consistent in 5 studies (33%).

A clinical or multidisciplinary team made the diagnoses and classifications in five studies; and child psychiatrists, psychologists, or other mental health professionals made the diagnoses and classifications in seven studies. The diagnosticians were not identified in two studies.

Diagnostic Classifications

Tables 4 and 5 show the percentages of children in each sample that were assigned to a diagnostic category or given a diagnosis on the axes constituting the focus of assessment. Axis I primary diagnoses were made for all 18 samples[2], Axis II relationship disorder classifications were made for 10 samples, Axis III medical and developmental disorders and conditions were identified in 8 samples, Axis IV psychosocial stressors were assessed in 6 samples, and Axis V functional emotional developmental levels were ascertained for only one sample.

Synthesis Findings

Table 6 shows which studies included reliability, validity, and accuracy data. Six studies included reliability data, 10 studies included validity data, and no study included accuracy data. The reliability and validity data included in the research reports were mostly for Axis I primary diagnoses, except in the Keren et al. (2001) and Skovgaard et al. (2005) studies. Therefore, findings pertaining to this axis constitute the main focus of synthesis and analysis.

Reliability

Interrater reliability was reported in five studies (33%) and test/retest reliability was reported in only two studies (13%). Interrater reliability was reported on 295 or just 14% of the 2,065 children included in the studies, and test/retest reliability was reported on only 30 or just 1% of the study participants.

Interrater reliability. DeGangi et al. (2000), Frankel et al. (2004), Keren et al. (2001), and Skovgaard et al. (2005) reported percent of agreement indices for ascertaining interrater reliability.[3] Percent agreement of at least 85% is generally considered the minimum for ascertaining interrater reliability calculated as the number of agreements divided by the number of agreements plus

---

[1]The number of participants in Table 2 include those children administered the DC: 0-3 and given an Axis I diagnosis or classificatory assignment. For several samples, these numbers differ from those in the original reports because not all the children who were identified as study participants were subsequently assessed by using the DC: 0-3.

[2]The percentages reported in Table 4 were recalculated in a number of studies to make the classificatory assignments comparable across samples. This involved primarily the exclusion of children given Axis I diagnoses not included in the DC: 0-3 classification system and by combining subcategories of Axis I diagnoses.

[3]Weston et al. (2003) reported interrater reliability where agreement was by consensus of the study investigators rather than by independent assessors and, therefore, did not meet the criteria for this type of reliability.

the number of disagreements multiplied by 100 (Fleiss, Levin, & Paik, 2003).

DeGangi et al. (2000) reported that "two independent child psychiatrists, blind to the group classification of subjects" (p. 164) had only 50% agreement for "the regulatory disordered group" (p. 164). Reliability for other DC: 0-3 classificatory assignments was not reported. Frankel et al. (2004) had two psychologists conduct retrospective chart reviews and assign "a maximum of three diagnoses to each case, based on the totality of the chart information independent of, but not blinded to, the original clinical diagnosis" (p. 581). The investigators reported "interrater reliability calculated by percentage agreement" (p. 581) of 68%, 76%, and 98% for three different primary diagnoses. Of the four interrater reliability estimates found in both research reports, only one was 85% or higher.

Keren et al. (2001) reported interrater reliability between a child and adolescent psychiatrist and clinical psychologist for Axis I and Axis II diagnosis. Interrater agreement was 100% for Axis I classificatory assignments and 92% for Axis II relationship diagnoses.

Skovgaard et al. (2005) were the only investigators who reported percent agreement interrater reliability estimates for more than one Axis I diagnostic classification and for more than one axis. For five Axis I diagnostic categories, percent agreement ranged between 83% and 100% for no diagnosis, adjustment disorders, regulatory disorders, multisystem developmental disorders, and emotional disorders. Interrater reliability for sleep disorders was 66%. For two Axis II diagnostic categories, percent agreement was 100% for both.

Guedeney et al. (2003) and Skovgaard et al. (2005) both calculated kappa for ascertaining interrater reliability. A kappa of .60 or higher is generally considered the minimum for ascertaining interrater reliability (Stemler, 2004). Guedeney et al. (2003) reported a kappa of .73 for two child psychiatrists not blinded to the children's diagnoses but kappa's of only .49 and .56 for raters blinded to the children's diagnoses. Skovgaard et al. (2005) had three child psychiatrists blind to 18 children's previously assigned classifications diagnose "each 18 cases by reviewing the case material…[and observing] 10 minutes of videorecording per case" (p. 473). They reported kappa's of .72 for Axis I primary diagnoses, 1.0 for Axis II relationship classifications, .71 for Axis III developmental conditions, .47 and .55 for Axis IV stressors, and .71 for Axis V functional emotional development levels. Taken together, the interrater reliability estimates in these two studies where raters were blind to classificatory assignments yielded inconsistent results.

The five studies including interrater reliability estimates included 10-percent-of-agreement estimates and four kappa estimates. Only four of the 10-percent-of-

agreement estimates were 85% or higher and only two of the four kappas were .60 or higher.

Test/retest reliability. Cordeiro et al. (2003) reported the classificatory assignments of 24 children administered the DC: 0-3 six or more months apart. The children's initial diagnoses were regulatory disorders (N = 15), affect disorders (N =5), and multisystem developmental disorders (N = 4). For the children diagnosed with either regulatory disorders or multisystem developmental disorders, only half of each group was given the same diagnosis at the follow-up assessment. The changes in diagnostic classifications may be due, in part, to the fact that the children were participants in a therapeutic intervention designed to improve mental health outcomes. Findings showed, however, that at the time of the second DC: 0-3 administration, only 6 (25%) of the 24 children were not given an Axis I diagnosis.

Skovgaard et al. (2005) reported test/retest reliability for six children administered the DC: 0-3 3-to-12 months apart. The test/retest reliability estimates were for children involved in an epidemiological study of mental health problems that did not involve any specified interventions. The kappa's were .74 for Axis I, 1.0 for Axis II, .72 for Axis III, .57 and .68 for Axis IV, and .84 for Axis V classifications and diagnoses.

Validity

Some type of convergent validity was reported or could be discerned in nine studies (60%). Discriminant validity was reported in only one study. Secondary analyses were performed to estimate the discriminant validity of the DC: 0-3.

Convergent validity. The criterion measures for ascertaining convergent validity in the largest number of studies were not well-established scales or were measurement procedures not frequently used with young children except the *Child Behavior Checklist* (Achenbach, 1991, 1992). The criterion measures in other studies included DSM III-R or DSM IV (American Psychiatric Association, 1987, 1994), the ICD-10 (World Health Organization, 1992), and the PIR-GAS (Zero to Three: National Center for Infants, 1994). The latter three measures are problematic for a number of reasons, including the lack of reliability and validity data on young children administered these procedures (Cantwell, 1996). Therefore, the extent to which they can be considered adequate criterion measures is somewhat questionable. At least one other problem warrants consideration in interpreting convergent validity data using these particular scales. The clinicians administering the criterion measures were typically the same persons administering the DC: 0-3. In studies where this was the case, convergent validity estimates may be confounded by administration procedures.

The DC: 0-3 classificatory assignments were compared to DSM III-R or DSM IV classificatory assignments in seven studies (Dunitz, Scheer, Kvas, & Macari, 1996; Frankel et al., 2004; Minde & Tidmarsh, 1997; Perez, Newman, Bruton, & Peifer, 2003; Reams, 1999; Thomas & Clark, 1998; Thomas & Guskin, 2001). For the same diagnoses included in both classification systems there was, not unexpectedly, relatively high degrees of correspondence. There was, however, generally no direct correspondence. For example, Reams (1999) reported that 40% of his sample received a DSM IV adjustment disorder diagnosis but that only 23% of these same children received a DC: 0-3 adjustment disorder diagnosis. Minde and Tidemarsh (1997) and Thomas and Clark (1998) reported low levels of diagnostic compatibilities whereas Dunitz et al. (1996) and Perez et al. (2003) reported high levels of correspondence.

Wright et al. (2004) in a study comparing DSM IV classifications of children seen in a community health program with published data on DC: 0-3 classifications (Cordeiro et al., 2003; Guédeney et al., 2003; Keren, Feldman, & Tyano, 2003; Maldonado-Duran et al., 2003) found both correspondences and incompatibilities. For diagnoses included on both classification systems, there were both similar (e.g., Adjustment disorders: DC: 0-3 = 7% and DSM IV = 6%) and dissimilar (e.g., Relational disorders: DC: 0-3 = 55% and DSM IV = 33%) percents of children receiving the same diagnoses.

A single study compared the DC: 0-3 classificatory assignments to ICD-10 classifications (Skovgaard et al., 2005). For 18 children assessed using both diagnostic systems, exactly the same numbers of children received (N = 8) and did not receive primary diagnosis (N = 10) using both procedures. For those children receiving a diagnosis, there was nearly a one-to-one correspondence in terms of classificatory assignments.

The relationship between DC: 0-3 Axis I primary diagnoses and PIR-GAS scores was examined in three studies (Guédeney et al., 2003; Thomas & Clark, 1998; Thomas & Guskin, 2001). Thomas and Clark (1998) reported no significant relationship between Axis I primary diagnoses (traumatic stress disorders, disorders of affect, and regulatory disorders) and PIR-GAS scores, whereas Guédéney et al. (2003) and Thomas and Guskin (2001) reported significant relationships between these same three diagnostic categories and PIR-GAS scores. In the latter two studies, PIR-GAS scores indicative of disordered parent/child relationships were more likely to be associated with Axis I disorders of affect classification.

Thomas and Clark (1998) and Thomas and Guskin (2001) compared DC: 0-3 classifications to the *Child Behavior Checklist* (Achenbach, 1991, 1992). These are the *only* studies where the criterion measure for ascertaining convergent validity was obtained independent of and unconfounded by DC: 0-3 administration procedures. The children in both studies had DC: 0-3 primary diagnoses of traumatic stress disorders, disorders of affect, or regulatory disorders. Thomas and Clark (1998) reported no significant relationship between child diagnosis and CBCL subscale scores, whereas Thomas and Guskin (2001) reported a significant relationship between child diagnosis and CBCL internalizing symptoms. Findings showed that a larger percentage of children with a disorder of affect had internalizing symptoms.

Discriminant validity. DeGangi et al. (2000) and Keren et al. (2001) were the only investigators who reported discriminant validity of DC: 0-3 classificatory assignments (although it was possible to estimate validity in most studies as described below). The DeGangi et al. study included three samples of infants, one that was normally developing (nonclinical sample) and two with different severities of regulatory disorders. The discriminant validity of the DC: 0-3 was ascertained by findings showing that only 3% of the normally developing infants received a primary diagnosis, 40% of the mild regulatory disordered children received a primary diagnosis, and 95% of the moderate regulatory disorder children received a primary diagnosis. Keren et al. (2001) compared the diagnostic classifications of clinical and nonclinical groups of infants on four DC: 0-3 axes. Fewer than expected clinically referred infants received diagnoses and more than expected comparison group infants did receive diagnoses (see Tables 4 and 5).

The findings reported in Table 4 for Axis I primary diagnoses were used to calculate discriminant validity estimates of the DC: 0-3 by comparing the percentages of children in the clinical and nonclinical samples that were assigned to different classifications. Discriminant validity would be demonstrated by classificatory assignments that were considerably different in the two samples of children. The results are shown in Figure 1. Contrary to expectation, 51% of children in the nonclinical samples received an Axis I primary diagnosis, whereas 30% of children in the clinical samples did not receive a primary diagnosis. Further inspection of the nonclinical samples finds that the percent receiving a primary diagnosis ranged from as low as 3% (DeGangi et al., 2000) to as high as 77% (Keren et al., 2001). While it is the case that one would expect a certain percent of children in the nonclinical samples to have mental-health-related problems or presenting conditions, the failure of the DC: 0-3 to differentiate between clinical and nonclinical cases indicates that the assessment procedures may lack discriminant validity. The contention that the DC: 0-3 may lack discriminant validity is bolstered by the fact that the percentages of children in the clinical and nonclinical samples receiving Axis I regulatory and adjustment disorder diagnoses
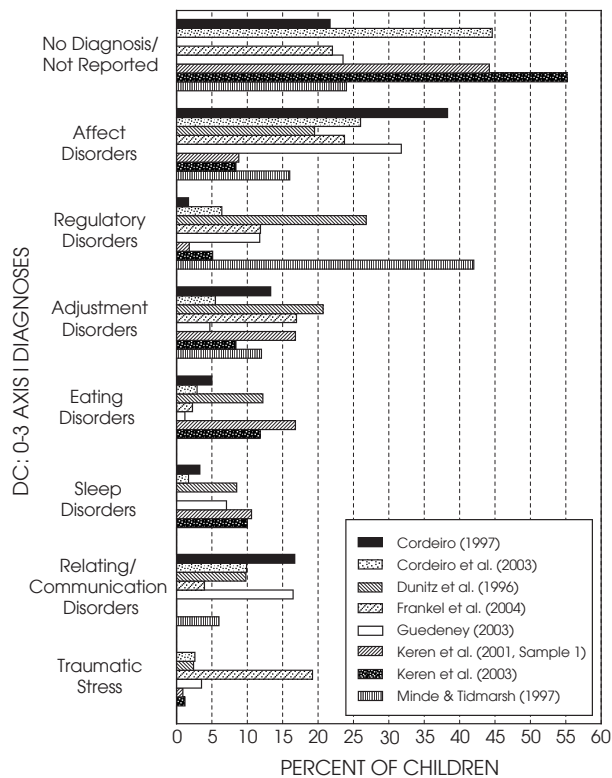
Figure 1. Percentage of study participants assigned to different DC: 0-3 Axis I primary diagnostic categories.

were more similar than different (Figure 1).

The discriminant validity of the DC: 0-3 can also be partly ascertained by examining the pattern of distribution of classifications in the clinical samples under the assumption that the distributions of Axis I diagnoses should be more alike than different. This would be expected to a large degree in those studies where the children were referred for similar concerns or problems. Figure 2 shows the prevalence of the Axis I diagnoses in the clinical samples that were referred for similar presenting problems or concerns. Patterns of assignments to the different DC: 0-3 diagnostic categories varied considerably. Children that received no diagnosis ranged from zero (0) to 60% of the study participants. For the two classifications assigned the largest numbers of children, the ranges were 8% to 38% for disorders of affect and 2% to 42% for regulatory disorders. These data indicate that the DC: 0-3 may lack discriminant validity as well.

Accuracy

None of the 15 research reports included any sensitivity or specificity data. Keren et al. (2003) were the only investigators that explicitly noted the likelihood of the false positives and false negatives using the DC: 0-3. Some simple calculations of the percentages of children

receiving and not receiving Axis I primary diagnoses (see Figure 1) found a larger than expected percentage of nonclinical sample children receiving a diagnosis (possible false positives) and a larger than expected percentage of clinical sample children not receiving a diagnosis (possible false negatives). This suggests that the DC: 0-3 may have both poor sensitivity and specificity, and, therefore, poor diagnostic efficiency.

Discussion

Claims about the reliability and validity of the DC: 0-3 run the gamut from those who assert that it has adequate psychometric properties (e.g., Evangelista & McLellan, 2004; Weston et al., 2003) to those who assert that the psychometric properties have not yet been determined (e.g., Delcarmen-Wiggins & Carter, 2001; Lyons-Ruth, Zeanah, & Benoit, 2003). Findings from this research synthesis indicate that the paucity of psychometric data on the DC: 0-3 does not support its use for early identification or eligibility determination of infants and toddlers for Part C early intervention (e.g., Gomez et al., 2004; Thomas & Tidmarsh, 1997). Reliability and validity data are very limited, and what data are available indicate that the procedure generally does not meet psychometric standards. Moreover, the failure of most study investigators to describe exactly how reliability and validity were determined makes interpretation of available data difficult at best.

Examination of the studies constituting the focus of analysis in this *Cornerstones* report found that very few studies included interrater reliability data and only two studies included test/retest data. Interrater reliability is
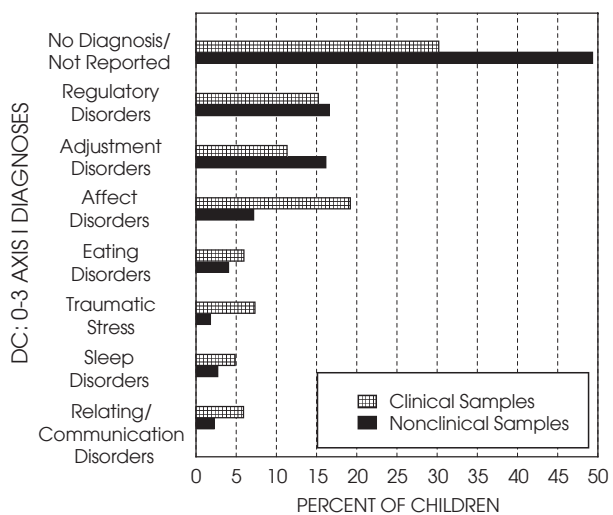

Figure 2. Percentages of study participants in different samples referred for similar reasons assigned to different DC: 0-3 Axis I primary diagnoses.

important because it provides a way of ascertaining the agreement of assignments to diagnostic classifications. Test/retest reliability is important because it provides a way of ensuring that classificatory assignments were not influenced by factors other than the true presence of a condition warranting a diagnosis. The need for both types of reliability data for the DC: 0-3 is clearly indicated. Only a single study included reliability estimates that were uniformly psychometrically acceptable (Skovgaard et al., 2005).

Information on the convergent validity of the DC: 0-3 was quite limited, and what information was available, in most cases, was confounded by administration procedures (i.e., the same persons administering both the DC: 0-3 and the criterion measures). The one exception was studies examining the relationship between DC: 0-3 classificatory assignments and the CBCL. Results from these studies, however, produced contradictory findings. Divergent validity data were entirely lacking, and an attempt to estimate discriminant validity found that the DC: 0-3 generally fared poorly in its ability to differentiate between children with and without mental- health-related problems. The need to show a relationship between the DC: 0-3 and some external criterion (convergent validity) and the need to be sure that a child is assigned to the correct diagnostic category (divergent validity) is important because both establish the fact that the DC: 0-3 is measuring what it purports to measure.

No accuracy (sensitivity or specificity) data on the DC: 0-3 was found in any study, and a modest attempt to ascertain diagnostic efficiency found the indices poor at best. The need for diagnostic efficiency data is especially warranted given the fact that the DC: 0-3 generally fares poorly in terms of differentiating between clinical and nonclinical samples of children.

Implications for Early Identification and Eligibility Determination

Results from this synthesis indicate that the DC: 0-3 is not warranted as either an early identification or eligibility determination assessment tool. The research synthesis nonetheless has a number of implications for early identification and eligibility determination to establish whether or not the DC: 0-3 can be recommended for use by IDEA Part C early intervention practitioners for identifying infant and toddler social-emotional disorders and mental-health-related problems. The following constitute five recommendations for further study of the psychometric properties of the DC: 0-3.

Interrater reliability studies of classificatory assignments are needed for all five DC: 0-3 axes diagnoses and especially for Axis I primary diagnoses. Such studies should ensure that researchers or clinicians making classificatory assignments are blind to one another's diagnosis or an already established diagnosis serving as the standard against which interrater reliability is being judged. There is also a need for interrater reliability studies of subcategories of Axis I diagnoses (disorders of affect, regulatory disorders, and disorders of relating and communicating). The latter is especially important given the fact that reliability studies of these subcategories are essentially nonexistent.

Test/retest reliability studies are needed to ensure classificatory assignments are truly based on the presence of a mental-health-related disorder or condition. Inasmuch as infant and toddler behavior is often influenced by many factors outside the child, and the fact that so few studies employed the recommended DC: 0-3 administration procedures, the possibility exists that classificatory assignments were state rather than trait based. Only test/retest reliability analyses would permit one to discern if a DC: 0-3 diagnosis truly reflects the presence of a mental health condition. These studies should have the children retested within a month or so of an initial diagnosis inasmuch as DC: 0-3 classificatory assignments appear age related (Maldonado-Duran et al., 2003). Additionally, the persons administering the retests should be different from those administering the first tests to be sure prior knowledge of a child's diagnosis does not influence a subsequent classificatory assignment.

Convergent validity studies are needed to ascertain if the DC: 0-3 classificatory assignments correlate with an external criterion. Much of the available information on the convergent validity of the DC: 0-3 is confounded by administration procedures used to measure the criterion. Studies are needed that use well-established criterion measures administered by persons other than the researchers or clinicians making DC: 0-3 classificatory assignments to ascertain convergent validity. The validity of these criterion measure studies would be bolstered considerably by the conduct of convergent and divergent validity statistical analyses (Campbell & Fiske, 1959).

Discriminant validity studies are needed to ascertain the extent to which children with the same presenting problems or concerns and assessment results are assigned to the same DC: 0-3 diagnostic category by two or more clinicians or researchers. Such studies would need to have researchers or clinicians blind to each other's assignments but where they make their classifications using the same assessment findings and results. These studies are sorely needed in light of the fact that DC: 0-3 classificatory assignments vary tremendously across studies.

Diagnostic efficiency studies are needed to ascertain the sensitivity and specificity of the DC: 0-3. Such studies would establish the false positive and false negative rates of the DC: 0-3 in terms of differentiating between children truly having and not having mental-health-related problems or disorders. These studies are especially

needed if the DC: 0-3 is to be used for eligibility determination for IDEA Part C early intervention to be sure a child assigned a diagnostic classification in fact has a "real" mental health problem or social-emotional disorder.

The particular psychometric properties for which recommendations were made are not the only reliability, validity, and accuracy data needed on the DC: 0-3. The reader is referred to Cantwell (1996) for a description of other psychometric data needed for diagnostic classification systems to justify their use for diagnostic and intervention purposes. The major conclusion from this practice-based research synthesis is that 12 years after the publication of the DC: 0-3 classification system, the reliability, validity, and diagnostic efficiency of the assessment procedures has yet to be adequately established. Caution is therefore warranted in using the DC: 0-3 for early identification or eligibility determination purposes by IDEA Part C early intervention practitioners.

Findings from this research synthesis do not warrant the development of a TRACE practice guide. The reader is referred to a nontechnical summary of this synthesis (*Endpoints*, Volume 1, Number 2) written specifically for practitioners and parents for an explanation of the reasons for this conclusion. The reader is also referred to Delcarmen-Wiggins and Carter (2001) and Lyons-Ruth et al. (2003) for other concerns involving the DC: 0-3. The diagnosis of social-emotional disorders early in a child's life is a serious enterprise and the children and their families need to be sure that the procedures used to make such diagnoses meet acceptable psychometric standards. This does not seem to be the case at this time for the DC: 0-3.

References

Achenbach, T. M. (1991). *Manual for the Child Behavior Checklist/4-18 and 1991 Profile*. Burlington: University of Vermont, Department of Psychiatry.

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

American Psychiatric Association. (1987). *Diagnostic and statistical manual of mental disorders: DSM-III-R* (3rd Rev. ed.). Washington, DC: Author.

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

Anastasi, A., & Urbina, S. (1997). *Psychological testing* (7th ed.). Upper Saddle River, NJ: Prentice Hall.

Bagnato, S. J., & Neisworth, J. T. (1999). Normative detection of early regulatory disorders and autism: Empirical confirmation of DC:0-3. *Infants and Young Children, 12*(2), 98-106.

Becker, L. A. (n.d.). *Reliability and validity, Part I and Part II*. Retrieved January 2, 2006, from http://web.uccs.edu/lbecker/Psy590/relval_I.htm

Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin, 56*, 81-105.

Cantwell, D. P. (1996). Classification of child and adolescent psychopathology. *Journal of Child Psychology and Psychiatry, 37*, 3-12.

Cordeiro, M. (1997). Treating infants and mothers in psychic distress: A mental health program for infancy. *Infant Mental Health Journal, 18*(2), 145-147.

Cordeiro, M. J., Da Silva, P. C., & Goldschmidt, T. (2003). Diagnostic classification: Results from a clinical experience of three years with DC: 0-3. *Infant Mental Health Journal, 24*, 349-364.

DeGangi, G. A., Breinbauer, C., Roosevelt, J. D., Porges, S., & Greenspan, S. (2000). Prediction of childhood problems at three years in children experiencing disorders of regulation during infancy. *Infant Mental Health Journal, 21*, 156-175.

Delcarmen-Wiggins, R., & Carter, A. S. (2001). Introduction: Assessment of infant and toddler mental health: Advances and challenges [Electronic version]. *Journal of the American Academy of Child and Adolescent Psychiatry, 40*, 8.

Dunitz, M., Scheer, P. J., Kvas, E., & Macari, S. (1996). Psychiatric diagnoses in infancy: A comparison. *Infant Mental Health Journal, 17*, 12-23.

Dunst, C. J., Trivette, C. M., & Cutspec, P. A. (2002). Toward an operational definition of evidence-based practices. *Centerscope, 1*(1), 1-10. Available at http://www.evidencebasedpractices.org/centerscope/centerscopevol1no1.pdf

Emde, R. N., & Wise, B. H. (2003). The cup is half full: Initial clinical trials of DC: 0-3 and a recommendation for revision. *Infant Mental Health Journal, 24*, 437-446.

Evangelista, N., & McLellan, M. J. (2004). The Zero to Three Diagnostic System: A framework for considering emotional and behavioral problems in young children. *School Psychology Review, 33*, 159-173.

Fleiss, J. L., Levin, B., & Paik, M. C. (2003) The measurement of interrater agreement. In *Statistical methods for rates and proportions* (3rd ed.). Hoboken, NJ: Wiley.

Foley, G. M., & Hochman, J. D. (Eds.). (2006). *Mental health in early intervention: Achieving unity in principles and practice*. Baltimore: Brookes.

Frankel, K. A., Boyum, L. A., & Harmon, R. J. (2004). Diagnoses and presenting symptoms in an infant psychiatry clinc: Comparison of two diagnostic sys-

tems [Electronic version]. *Journal of the American Academy of Child and Adolescent Psychiatry, 43*, 578-587.

Gomez, C. R., Baird, S., & Jung, L. A. (2004). Regulatory disorder identification, diagnosis, and intervention planning: Untapped resources for facilitating development. *Infants and Young Children, 17*, 324-339.

Guédeney, A., & Maestro, S. (2003). The use of the diagnostic classification 0-3. *Infant Mental Health Journal, 24*(4), 309-446.

Guédeney, N., Guédeney, A., Rabouam, C., Mintz, A.-S., Danon, G., Huet, M. M., & Jacquemain, F. (2003). The zero-to-three diagnostic classification: A contribution to the validation of this classification from a sample of 85 under-threes. *Infant Mental Health Journal, 24*, 313-336.

Keren, M., Feldman, R., & Tyano, S. (2001). Diagnoses and interactive patterns of infants referred to a community-based infant mental health clinic. *Journal of the American Academy of Child and Adolescent Psychiatry, 40*, 27-35.

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

Lieberman, A., Wieder, S., & Fenichel, E. (Eds.). (1997). *DC: 0-3 casebook. A guide to the use of Zero to Three's Diagnostic Classification of Mental Health and Developmental Disorders of Infancy and Childhood in assessment and treatment planning*. Washington, DC: Zero to Three.

Lyons-Ruth, K., Zeanah, C. H., & Benoit, D. (2003). Disorder and risk for disorder during infancy and toddlerhood. In E. J. Mash & R. A. Barkley (Eds.), *Child psychopathology* (2nd ed., pp. 589-631). New York: Guilford Press.

Maldonado-Duran, M., Helmig, L., Moody, C., Fonagy, P., Fulz, J., Lartigue, T., Sauceda-Garcia, J. M., Karacostas, V., Millhuff, C., & Glinka, J. (2003). The zero-to-three diagnostic classification in an infant mental health clinic: Its usefulness and challenges. *Infant Mental Health Journal, 24*, 378-397.

Medical Algorithms Project. (2006). *Decisional analysis: Overall accuracy, or diagnostic efficiency (Chapter 39)*. Houston, TX: Institute for Algorithmic Medicine. Retrieved August 17, 2006, from http://www.medal.org/visitor/www/Active/ch39/ch39.01/ch39.01.09.aspx

Meyer, G. J. (2003). Guidelines for reporting information in studies of diagnostic test accuracy: The STARD Initiative. *Journal of Personality Assessment, 81*, 191-193.

Minde, K., & Tidmarsh, L. (1997). The changing practices of an infant psychiatry program: The McGill experience. *Infant Mental Health Journal, 18*, 135-144.

Oser, C. (2004, October). *Infant/early childhood mental health: A concern for all*. Presentation made at the HRSA/MCHB Grantee meeting, Arlington, VA.

Osofsky, J. D., & Fitzgerald, H. E. (2000). *WAIMH handbook of infant mental health: Vol. 2. Early intervention, evaluation, and assessment*. New York: Wiley.

Perez, L. M., Newman, M. C., Bruton, N., & Peifer, K. L. (2003). Infant mental health evaluation process; Evaluating, diagnosing, and treating infant mental health in community practice. *Zero to Three, 23*(5), 55-64.

Reams, R. (1999). Children birth to three entering the state's custody. *Infant Mental Health Journal, 20*, 166-174.

Shackelford, J. (2006, February). State and jurisdictional eligibility definitions for infants and toddlers with disabilities under IDEA. *NECTAC Notes,* (No. 20.), 1-16. Chapel Hill: National Early Childhood Technical Assistance Center, University of North Carolina.

Skovgaard, A. M., Houmann, T., Christiansen, E., & Andreasen, A. H. (2005). The reliability of the ICD-10 and the DC 0-3 in an epidemiological sample of children 1 1/2 years of age. *Infant Mental Health Journal, 26*, 470-480.

Skovgaard, A. M., Houmann, T., Landorph, S. L., & Christiansen, E. (2004). Assessment and classification of psychopathology in epidemiological research of children 0-3 years of age: A review of the literature. *European Child and Adolescent Psychiatry, 13*, 337-346.

Stemler, S. E. (2004). A comparison of consensus, consistency, and measurement approaches to estimating interrater reliability. *Practical Assessment, Research and Evaluation, 9*(4). Retrieved August 14, 2006, from http://pareonline.net/getvn.asp?v=9&n=4

Thomas, J. M., Benham, A., & Guskin, K. A. (2000). Intervention-centered assessment: Opportunity for early and preventive intervention. In J. D. Osofsky & H. E. Fitzgerald (Eds.), *WAIMH handbook of infant mental health: Vol. 2. Early intervention, evaluation, and assessment* (pp. 403-440). New York: Wiley.

Thomas, J. M., & Clark, R. (1998). Disruptive behavior in the very young child: Diagnostic Classification: 0-3 guides identification of risk factors and relational interventions. *Infant Mental Health Journal, 19*, 229-244.

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

Thomas, J. M., & Tidmarsh, L. (1997). Hyperactive and disruptive behaviors in very young children: Diagnosis and intervention. *Infants and Young Children, 9*(3), 46-55.

Waisman Center. (2006). *Guidelines for best practices in determining eligibility based on children's social and emotional development*. Retrieved November 20, 2006, http://www.waisman.wisc.edu/birthto3/Txt/social.html

Weston, D. R., Thomas, J. M., Barnard, K. E., Wieder, S., Clark, R., Carter, A. S., & Fenichel, E. (2003). DC: 0-3 assessment protocol project: Defining a comprehensive information set to support DC: 0-3 diagnostic formulation. *Infant Mental Health Journal, 24*, 410-427.

World Health Organization. (1992). *The ICD-10 classification of mental and behavioural disorders: Clinical descriptions and diagnostic guidelines*. Geneva, Switzerland: Author.

Wright, H., Holmes, G., Stader, S., Penny, R., & Wieduwilt, K. (2004). Psychiatric diagnoses of infants and toddlers referred to a community mental health system. *Psychological Reports, 95*, 495-503.

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

Zero to Three: National Center for Infants, Toddlers, and Families. (2005). *Diagnostic classification of mental health and developmental disorders of infancy and early childhood: DC: 0-3R* (Rev. ed.). Washington, DC: Zero to Three Press.

Authors

Carl J. Dunst, Ph.D., is Co-Principal Investigator at TRACE and Co-Director of the Orelena Hawks Puckett Institute, Asheville, North Carolina (dunst@puckett.org). Andrea Storck, B.S., is a Research Associate and Donna Snyder, B.A., is a Research Assistant at the Puckett Institute.

Table 1
DC: 0-3 Diagnostic and Classification Categories

Axis I: Primary Diagnosis
      Traumatic stress disorder
      Disorders of affect
      Adjustment disorder
      Regulatory disorders
      Sleep behavior disorder
      Eating behavior disorder
      Disorders of relating/communicating

Axis II: Relationship Disorder Classification
      Over involved
      Under involved
      Anxious/tense
      Angry/hostile
      Mixed
      Abusive

Axis III: Medical/Developmental Disorders
      Physical
      Mental health
      Developmental

Axis IV: Psychosocial Stressors
      Sources
      Severity
      Duration
      Impact
      Supports (caregivers)

Axis V: Functional Emotional Developmental Level
      Mutual attention
      Mutual engagement
      Interactive intentionality/reciprocity
      Affective communication
      Elaboration
      Differentiation

Table 2

Characteristics of the Study Participants at the Time of the Administration of the DC: 0-3

Study	Type of Sample	Number of Participants	Child Age (Months)	Child Gender		Child Ethnicity		Country
				Male	Female	Ethnicity	Percent	
Cordeiro (1997)	Clinical	60	0–36	35	25	Portuguese	100	Portugal
Cordeiro et al. (2003)	Clinical	343	0–48	206	137	Portuguese	100	Portugal
DeGangi et al. (2000, Sample 1)	Nonclinical	38	7–30	23	15	Caucasian Latino	97 3	USA
DeGangi et al. (2000, Sample 2)	Clinical	10	7–30	8	2	Caucasian	100	USA
DeGangi et al. (2000, Sample 3)	Clinical	22	7–30	12	10	Caucasian Asian Latino African American	85 5 5 5	USA
Dunitz et al. (1996)	Clinical	82	1–24	38	44	Caucasian Filipino Hamite African	96 1 1 1	Austria
Frankel et al. (2004)	Clinical	177	0–58	99	78	Caucasian African American Latino Multiracial Other Nor reported	73 10 6 4 4 3	USA
Guedeney et al. (2003)	Clinical	85	0–36	45	40	Not reported	–	France
Keren et al. (2001, Sample 1)	Clinical	113	0–36	68	45	Western Jews Eastern Jews Russian Immigrants	– – –	Israel
Keren et al. (2001, Sample 2)	Nonclinical	30	0–36	22	8	Western Jews Eastern Jews Russian Immigrants	– – –	Israel
Keren et al. (2003)	Clinical	431	0–36	254	177	Israeli Jews	100	Israel
Maldonado-Duran et al. (2003)	Clinical	167	0–36	98	69	Caucasian Multiethnic African American Latino Other	59 23 11 5 2	USA
Minde & Tidmarsh (1997)	Clinical	57	15–48	45	12	Not reported	–	USA
Perez et al. (2003)	Clinical	142	0–48	NR	NR	Not reported	–	USA
Reams (1999)	Nonclinical (at risk)	144	0–48	76	68	European American Multiracial African American Native American Latino Asian American	59 25 11 3 1 1	USA
Skovgaard et al. (2005)	Nonclinical	18	18	NR	NR	Danish	100	Denmark
Thomas & Clark (1998)	Clinical	64	12–47	46	18	Caucasian African American Latino Multiracial	64 25 5 5	USA
Thomas & Guskin (2001)	Clinical	82	18–47	63	19	Caucasian African American Asian or Biracial	71 23 6	USA

NOTES: NR = Not reported. The sample sizes in this table in some cases differ from those in the published papers because of inconsistencies in investigator reporting.

Table 3

Methods Used for Assigning Children DC: 0-3 Diagnosis and the Degree of Consistency of the Assessment Methods with Recommended Procedures

Study	Assessment Procedures

Cordeiro (1997)
Diagnostician:
  Child psychiatrist
  Clinical psychologist
 	
Degree of Consistency with DC: 0-3 Procedures:
  High
	Presenting symptoms and behaviors
	Developmental history — past and current affective, language, cognitive, motor, sensory, family, and interactive functioning
	Parents and individuals
	Caregiver/infant (child) relationship and interactive patterns
	Affective, language, cognitive, motor, and sensory patterns
	Family's psychosocial and medical history
	Current environmental conditions and stressors
	Taking the history
	Direct observation of functioning (i.e., of family and parental dynamics, caregiver-infant relationship and interactive patterns, the infant's constitutional-maturational characteristics, and language, cognitive, and affective patterns

Cordeiro et al. (2003)
Diagnostician:
  Multidisciplinary team
 	
Degree of Consistency with DC: 0-3 Procedures:
  High
	Presenting symptoms and behaviors
	Developmental history — past and current affective, language, cognitive, motor, sensory, family, and interactive functioning
	Family functioning and cultural community patterns
	Parents and individuals
	Caregiver/infant (child) relationship and interactive patterns
	Infant's constitutional-maturational characteristics
	Affective, language, cognitive, motor, and sensory patterns
	Family's psychosocial and medical history
	History of the pregnancy and delivery
	Current environmental conditions and stressors
	Taking the history
	Direct observation of functioning (i.e., of family and parental dynamics, caregiver-infant relationship and interactive patterns, the infant's constitutional-maturational characteristics, and language, cognitive, and affective patterns
	Hands-on interactive assessment of the infant
	Assessment of sensory reactivity and processing, motor tone and planning, language, cognition, and affective expression

DeGangi et al. (2000)
Diagnostician:
  Two child psychiatrists blind to the group classification of subjects, conducted a preliminary clinical diagnosis at 36 months
 	
Degree of Consistency with DC: 0-3 Procedures:
  High
	Presenting symptoms and behaviors
	Developmental history — past and current affective, language, cognitive, motor, sensory, family, and interactive functioning
	Affective, language, cognitive, motor, and sensory patterns
	Family's psychosocial and medical history
	History of the pregnancy and delivery
	Current environmental conditions and stressors
	Direct observation of functioning (i.e., of family and parental dynamics, caregiver-infant relationship and interactive patterns, the infant's constitutional-maturational characteristics, and language, cognitive, and affective patterns
	Assessment of sensory reactivity and processing, motor tone and planning, language, cognition, and affective expression

Dunitz et al. (1996)
Diagnostician:
  Clinical team
 	
Degree of Consistency with DC: 0-3 Procedures:
  Low
	Presenting symptoms and behaviors
	Caregiver/infant (child) relationship and interactive patterns
	History of the pregnancy and delivery

Frankel, et al. (2004)
Diagnostician:
  Two postdoctoral psychologists
 	
Degree of Consistency with DC: 0-3 Procedures:
  Low
	Presenting symptoms and behaviors
	Developmental history — past and current affective, language, cognitive, motor, sensory, family, and interactive functioning
	Family's psychosocial and medical history

Table 3, continued

Study	Assessment Procedures

Guedeney et al. (2003)
Diagnostician:
    Consensus team — used
    full assessment materials
 	
Degree of Consistency with
DC: 0-3 Procedures:
    Moderate
	Parents and individuals
	Caregiver/infant (child) relationship and interactive patterns
	Affective, language, cognitive, motor, and sensory patterns
	Family's psychosocial and medical history
	Current environmental conditions and stressors
	Direct observation of functioning (i.e., of family and parental dynamics, caregiver-infant relationship and interactive patterns, the infant's constitutional-maturational characteristics, and language, cognitive, and affective patterns)

Keren et al. (2001)
Diagnostician:
    Child-Adolescent Psychiatrist
 	
Degree of Consistency with
DC: 0-3 Procedures:
    Low
	Presenting symptoms and behaviors
	Developmental history — past and current affective, language, cognitive, motor, sensory, family, and interactive functioning
	Current environmental conditions and stressors

Keren et al. (2003)
Diagnostician:
    Not Specified
 	
Degree of Consistency with
DC: 0-3 Procedures:
    Moderate
	Developmental history — past and current affective, language, cognitive, motor, sensory, family, and interactive functioning
	Parents and individuals
	Caregiver/infant (child) relationship and interactive patterns
	Current environmental conditions and stressors

Maldonado-Duran et al. (2003)
Diagnostician:
    Multidisciplinary infant mental
    health team
 	
Degree of Consistency with
DC: 0-3 Procedures:
    Moderate
	Presenting symptoms and behaviors
	Developmental history — past and current affective, language, cognitive, motor, sensory, family, and interactive functioning
	Parents and individuals
	Caregiver/infant (child) relationship and interactive patterns
	Infant's constitutional-maturational characteristics
	Family's psychosocial and medical history
	Current environmental conditions and stressors

Minde et al. (1997)
Diagnostician:
    Multidisciplinary team
 	
Degree of Consistency with
DC: 0-3 Procedures:
    Moderate
	Presenting symptoms and behaviors
	Developmental history — past and current affective, language, cognitive, motor, sensory, family, and interactive functioning
	Caregiver/infant (child) relationship and interactive patterns
	Current environmental conditions and stressors

Perez (2003)
Diagnostician:
    Mental health high-risk
    clinician
 	
Degree of Consistency with
DC: 0-3 Procedures:
    Moderate
	Presenting symptoms and behaviors
	Developmental history — past and current affective, language, cognitive, motor, sensory, family, and interactive functioning
	parents as individuals
	Caregiver/infant (child) relationship and interactive patterns
	Family's psychosocial and medical history
	History of pregnancy and delivery
	Current environmental conditions and stressors

Reams (1999)
Diagnostician:
    MS level staff (Children's
    Assessment Service).
    Report/dx reviewed by doctoral
    psychologist experienced in
    infant/toddler mental health
    issues.
 	
Degree of Consistency with
DC: 0-3 Procedures:
    Moderate
	Presenting symptoms and behaviors
	Developmental history — past and current affective, language, cognitive, motor, sensory, family, and interactive functioning
	Caregiver/infant (child) relationship and interactive patterns
	Infant's constitutional-maturational characteristics
	Current environmental conditions and stressors
	Direct observation of functioning (i.e., of family and parental dynamics, caregiver-infant relationship and interactive patterns, the infant's constitutional-maturational characteristics, and language, cognitive, and affective patterns
	Hands-on interactive assessment of the infant

Table 3, continued

Study	Assessment Procedures

Skovgaard et al (2005)
Diagnostician:
  Child Psychiatrists
 	
Degree of Consistency with DC: 0-3 Procedures:
  Low
	Review of records/questionnaires
	Review of 10-minute videorecording

Thomas & Clarke (1998)
Diagnostician:
  Not Specified
 	
Degree of Consistency with DC: 0-3 Procedures
  Moderate
	Caregiver/infant (child) relationship and interactive patterns
	Infant's constitutional-maturational characteristics
	Current environmental conditions and stressors
	Direct observation of functioning (i.e., of family and parental dynamics, caregiver-infant relationship and interactive patterns, the infant's constitutional-maturational characteristics, and language, cognitive, and affective patterns

Thomas & Guskin (2001)
Diagnostician:
  Senior psychiatrist with 14 years of infant and early childhood experience and child psychiatry resident
 	
Degree of Consistency with DC: 0-3 Procedures:
  Low
	Clinical interviews with the family
	Observations of child and primary caregiver during free play time
	Videotaped assessments were reviewed prior to PIR-GAS scoring

Table 4

DC: 0-3 Diagnosis on Axes I and II for the 18 Samples of Children

Study	Axis I: Primary Diagnosis		Axis II: Relationship Disorder	
	Diagnosis	Percent	Diagnosis	Percent
Cordeiro (1997)	No diagnosis	22	No diagnosis	35
	Affect disorders	38	Under involved	28
	Disorders of relating and communicating	17	Mixed	8
	Adjustment disorder	13	Anxious/tense	12
	Eating behavior disorder	5	Over involved	12
	Sleep behavior disorder	3	Angry/hostile	5
	Regulatory disorder	2	Abusive	0
Cordeiro et al. (2003)	No diagnosis	45	No diagnosis/not reported	36
	Affect disorders	26	Nonapplicable	3
	Disorders of relating and communicating	10	Under involved	29
	Regulatory disorder	6	Mixed	12
	Adjustment disorder	5	Anxious/tense	10
	Traumatic stress disorder	3	Over involved	4
	Eating behavior disorder	3	Angry/hostile	3
	Sleep behavior disorder	2	Abusive	3
Degangi et al. (2000, Sample 1)	No diagnosis	97	Not reported	–
	Regulatory disorder	3		
Degangi et al. (2000, Sample 2)	No diagnosis	60	Not reported	–
	Regulatory disorder	30		
	Disorders of relating and communicating	10		
Degangi et al. (2000, Sample 3)	No diagnosis	5	Not reported	–
	Regulatory disorder	50		
	Sleep behavior disorder	23		
	Disorders of relating and communicating	22		
Dunitz et al. (1996)	No diagnosis	0	Not reported	–
	Regulatory disorder	27		
	Adjustment disorder	21		
	Affect disorders	19		
	Eating behavior disorder	12		
	Disorders of relating and communicating	10		
	Sleep behavior disorder	9		
	Traumatic stress disorder	2		
Frankel et al. (2004)	No diagnosis	22	Not reported	–
	Affect disorder	24		
	Traumatic stress disorder	19		
	Adjustment disorder	17		
	Regulatory disorder	12		
	Disorders of relating and communicating	4		
	Eating behavior disorder	2		
Guedeney et al. (2003)	No diagnosis	23	No diagnosis	0
	Affect disorders	32	Under involved	31
	Disorders of relating and communicating	16	Anxious/tense	25
	Regulatory disorder	12	Mixed	25
	Sleep behavior disorder	7	Over involved	11
	Adjustment disorder	5	Angry/hostile	8
	Traumatic stress disorder	3		
	Eating behavior disorder	1		
Keren et al. (2001, Sample 1)	No diagnosis	44	No diagnosis	53
	Adjustment disorder	17	Mixed	21
	Eating behavior disorder	17	Anxious/tense	14
	Sleep behavior disorder	10	Over involved	6
	Affect disorders	9	Angry/hostile	3
	Regulatory disorder	2	Under involved	2
	Traumatic stress disorder	1	Abusive	1

Table 4, continued

Study	Axis I: Primary Diagnosis		Axis II: Relationship Disorder	
	Diagnosis	Percent	Diagnosis	Percent
Keren et al. (2001, Sample 2)	No diagnosis	23	No diagnosis	17
	Eating behavior disorder	30	Mixed	37
	Sleep behavior disorder	17	Anxious/tense	17
	Affect disorders	13	Over involved	23
	Adjustment disorder	7	Abusive	7
	Regulatory disorder	7	Angry/hostile	0
	Traumatic stress disorder	3	Under involved	0
Keren et al. (2003)	No diagnosis	55	No diagnosis	48
	Eating behavior disorder	12	Mixed type	18
	Sleep behavior disorder	10	Anxious tense type	12
	Adjustment disorder	8	Over involved type	11
	Affect disorders	8	Under involved type	7
	Regulatory disorder	5	Angry/hostile type	3
	Traumatic stress disorder	1	Abusive type (physical)	1
Maldonado-Duran et al. (2003)	No diagnosis	10	No diagnosis	63
	Regulatory disorder	42	Under involved	23
	Traumatic stress disorder	12	Over involved	7
	Adjustment disorder	11	Anxious/tense	2
	Disorders of relating and communicating	10	Angry/hostile	2
	Affect disorder	8	Mixed	2
	Eating behavior disorder	4	Abusive	1
	Sleep behavior disorder	3		
Minde & Tidmarsh (1997)	No diagnosis/not reported	21	No diagnosis	47
	Regulatory disorder	37	Under involved	35
	Affect disorder	14	Over involved	11
	Other (not specified)	12	Anxious/tense	5
	Adjustment disorder	11	Angry/hostile	2
	Disorders of relating and communicating	5	Mixed	0
			Abusive	0
Perez et al. (2003)	No diagnosis	0	No diagnosis	0
	Adjustment disorder	35	Mixed	32
	Affect disorders	18	Under involved	23
	Traumatic stress disorder	18	Anxious/tense	23
	Regulatory disorder	17	Abusive	18
	Disorders of relating and communicating	7	Over involved	4
	Eating behavior disorder	3	Angry/hostile	0
	Sleep behavior disorder	2		
Reams (1999)	No diagnosis/not reported	39	Not reported	–
	Adjustment disorder	23		
	Regulatory disorder	22		
	Affect disorder	7		
	Other (not specified)	5		
	Traumatic stress disorder	2		
	Disorders of relating and communicating	2		
Skovgaard et al. (2005)	No diagnosis	55	No Diagnosis	89
	Regulatory disorder	11	Under involved	11
	Disorders of relating and communicating	11		
	Affect disorder	11		
	Sleep behavior disorder	6		
	Adjustment disorder	6		
Thomas & Clark (1998)	No diagnosis	6	Not reported	–
	Affect disorders	41		
	Regulatory disorder	30		
	Traumatic stress disorder	23		
Thomas & Guskin (2001)	No diagnosis	0	Not reported	–
	Disorders of affect	42		
	Regulatory disorder	35		
	Traumatic stress disorder	23		

NOTE: Subcategories of Axis I diagnoses were combined into single primary diagnosis classifications.

Table 5
DC: 0-3 Diagnoses on Axes III, IV, and V

Study	Axis III Medical/Developmental		Axis IV Psychosocial Stressors		Axis V FED[a] Level	
	Diagnosis	Percent	Diagnosis	Percent	Diagnosis	Percent
Cordeiro (1997)	No diagnosis	23	Not reported	–	Not reported	–
	Developmentally delayed	77				
Cordeiro (2003)	Not reported	–	Not reported	–	Not reported	–
DeGangi et al. (2000, Sample 2)	Not reported	–	Not reported	–	Not reported	–
DeGangi et al. (2000, Sample 3)	Not reported	–	Not reported	–	Not reported	–
Dunitz (1996)	Not reported	–	Not reported	–	Not reported	–
Frankel et al. (2004)	Speech language delays/problems	50	Not reported	–	Not reported	–
	Cognitive delays	50				
Guedeney et al. (2003)	No diagnosis	32	*Description of most frequent stressors*		No problems	62
	Diagnosis (not specified)	68	Parental psychopathology	35	Severe impairment	31
			Home conflicts	26	Moderate impairment	7
			Negligence	15		
			Impact of stressors			
			Moderate impact	39		
			Severe impact	26		
			No impact	18		
			Mild impact	17		
Keren et al. (2001, Sample 1)	No diagnosis	79	No diagnosis	48	Not reported	–
	Global developmental delay	12	Parental psychopathology	27		
	Failure to thrive	3	Parental conflict	19		
	Developmental language disorder	3	Life events	4		
	Cerebral palsy	2	Parental divorce	2		
	Mental retardation	1	Parental medical illness	1		
Kerren et al. (2001, Sample 2)	No diagnosis	43	No diagnosis	57	Not reported	–
	Global developmental delay	23	Parental psychopathology	27		
	Failure to thrive	13	Parental conflict	13		
	Developmental language disorder	13	Life events	3		
	Cerebral palsy	7				
Keren et al. (2003)	No diagnosis	79	No diagnosis	60	Not reported	–
	Developmental delays	13	Parental psychiatric problem	22		
	Failure to thrive	3	Marital conflict	8		
	Medical diagnosis (chronic illness)	3	Other	3		
	Long-term consequences of premature birth	1	Loss	2		
			Divorce	2		
			Acute trauma	1		
			Birth of sibling	1		
			Abuse	1		
			Adoption	<1		

Table 5, continued

Study	Axis III Medical/Developmental		Axis IV Psychosocial Stressors		Axis V FED[a] Level	
	Diagnosis	Percent	Diagnosis	Percent	Diagnosis	Percent
Maldonado-Duran et al. (2003)	No diagnosis	15	Financial problems for family or poverty	49	Not reported	–
	Other health condition (urinary, GI, cardiac, coetaneous, etc. and chronic illnesses)	26	Marked conflict within the family	40		
	Failure to thrive	18	Violence in the environment	26		
	Other health condition (urinary, chronic or repetitive otitis media)	10	History of psychological trauma to significant other in infant's life	25		
	Neurological disorder (other than epilepsy)	9	Separation from parent due to work	22		
	Asthma	8	History of neglect	21		
	Iron deficiency, obesity	7	Move	20		
	Seizure disorder	4	History of emotional maltreatment	19		
	Major allergies	2	Foster home placement	17		
	Visible malformation	1	Loss of contact with parent	16		
			Birth of sibling	7		
			Divorce of parents	7		
			History of physical abuse	6		
			History of sexual abuse	2		
Minde & Tidmarsh (1997)	Specific medical problem	73	Marital problems	44	Not reported	–
	Developmental problems	18	Parental psychopathology	33		
	Mental retardation	9	Drug use	5		
			Economic problems	4		
			Medical problems	5		
			Others	9		
			Effects on child			
			Mild	30		
			Moderate	56		
			Severe	14		
Perez (2003)	Not reported	–	Not reported	–	Not reported	–
Reams (1999)	Not reported	–	Not reported	–	Not reported	–
Skovgaard et al. (2005)	Not reported	–	Not reported	–	Not reported	–
Thomas & Clarke (1998)	Not reported	–	Not reported	–	Not reported	–
Thomas & Guskin (2001)	Not reported	–	Not reported	–	Not reported	–

[a]FED = Functional Emotional Development Level

Table 6

Studies Including Reliability, Validity, and Accuracy Data for DC: 0-3 Classificatory Assignments

Study	Type of Reliability		Type of Validity		Accuracy	
	Interrater	Test/Retest	Convergent	Discriminant	Specificity	Sensitivity
Corderio (1997)	No	No	No	No	No	No
Corderio et al. (2003)	No	Yes	No	No	No	No
DeGangi et al. (2000)	Yes	No	No	Yes	No	No
Dunitz et al. (1996)	No	No	Yes	No	No	No
Frankel et al. (2004)	Yes	No	Yes	No	No	No
Guedeney et al. (2003)	Yes	No	Yes	No	No	No
Keren et al. (2001)	Yes	No	No	No	No	No
Keren et al. (2003)	No	No	No	No	No	No
Maldonado-Duran et al. (2003)	No	No	No	No	No	No
Minde & Tidmarsh (1997)	No	No	Yes	No	No	No
Perez et al. (2003)	No	No	Yes	No	No	No
Reams (1999)	No	No	Yes	No	No	No
Skovgaard et al. (2005)	Yes	Yes	Yes	No	No	No
Thomas & Clark (1998)	No	No	Yes	No	No	No
Thomas & Guskin (2001)	No	No	Yes	No	No	No
Percent of Studies	33	13	60	7	0	0