

Methods for Assessing Child and Family Outcomes in Early Childhood Special Education Programs: Some Views from the Field

Stacey E. Mott, *Utah State University*, Rebecca R. Fewell, *University of Washington*, Michael Lewis, *Rutgers Medical School*, Samuel J. Meisels, *University of Michigan*, Jack P. Shonkoff, *University of Massachusetts Medical School*, and Rune J. Simeonsson, *University of North Carolina at Chapel Hill*

Although many concerns have been raised about methods of assessing outcomes in early childhood special education programs, professionals in the field are nevertheless faced with the need to select appropriate instruments for evaluating child and family outcomes as the result of intervention. A conference to address the current assessment needs of professionals was convened. This paper summarizes this conference, in which five prominent individuals in the field of early childhood special education gave specific recommendations for one child and one family outcome measure which would be applicable to a range of handicapped children between birth and age 5 being served in typical early intervention programs.

There has been much discussion in recent years about the methods and instruments appropriate for assessing outcomes of early childhood special education programs. Past evaluations of the effects of early intervention have focused heavily on the measurement of IQ and have

ignored other indicators of child progress or family functioning (Castro & Lewis, in press; Zigler & Balla, 1982). Shonkoff (1983) has questioned the practice of using norm-referenced instruments that were developed in a historical context inconsistent with current assessment goals and that excluded handicapped children from the standardization sample. Others have pointed to problems with assessment procedures that are not sensitive to program effects (Garwood, 1982). These concerns suggest the difficulty involved in selecting instruments to assess the outcomes of early childhood special education programs.

Logically, many have called for the development of better instruments for assessing the outcomes of early intervention programs for handicapped children or have given cautions about how the results of such assessments should be interpreted (Brooks-Gunn & Lewis, 1983; Ramey, Campbell, & Wasik, 1982). Although these discussions have been useful in sensitizing the field to the importance of good assessment procedures, they do not satisfy the current needs in the field. For example, many state coordinators of early childhood special education programs must have specific information about the most appropriate assessment instruments to use for evaluating programs that operate in a variety of service settings and that serve children with a range of handicapping conditions. The current practice of selecting instruments based on familiarity or ease of administration alone highlights this need.

Another impetus for scrutinizing assessment instruments comes from the research community. The Early Intervention Effectiveness Institute at Utah State University is faced with the challenge of conducting 16 longitudinal studies on the effects of early intervention with handicapped children. In this research, comparisons are to be made between programs that serve children with a range of handicapping conditions and that use different intervention models. The use of at least one common child measure and one family measure would facilitate these comparisons.

In December 1985, a conference was convened to address both the needs of early intervention programs and those of the research community. Five prominent individuals from the field of early childhood special education were asked to recommend one child outcome measure and one family outcome measure that would be appropriate for the evaluation of a typical early childhood special education program serving children from birth to 5 years with a range of handicapping conditions. This paper summarizes their recommendations.

Views From the Field

Rune J. Simeonsson, PhD, University of North Carolina at Chapel Hill

Approach to Measures Selection. Given the problems in this area, it is clear that identification of ideal child and family instruments for widespread use is difficult. However, given the relative advantages and limitations of various instruments currently available, there are at least two that can be recommended.

Child Outcome Measure. The Griffiths Developmental Scale (Griffiths, 1970) is recommended for several reasons. The Griffiths can be used from birth to 8 years of age and is a well-standardized instrument. It generates a profile on the basis of several subscales, a characteristic not found in many of the instruments relevant to this age range. It has good psychometric properties, and information about performance on the items could be used to plan intervention for individual children.

The Griffiths has 498 items. From birth to 2 years, there are five subscales: Locomotor, Personal Social, Hearing and Speech, Eye-Hand Coordination, and Performance. From years 3 through 8 another subscale is added, Practical Reasoning. Quotients can be obtained for each subscale and the quotients are averaged across subscales for the general quotient of development. Mental ages can also be obtained.

The standardization sample consisted of 2,263 British children stratified by age levels. The test-retest reliability for the infant scale is .87, while test-retest for the overall scale is .77. The internal consistency of the instrument is represented by intercorrelations among subscales ranging from .64 to .78. Studies that have compared the Griffiths general quotient with the Stanford Binet have identified correlations ranging from .79 to .81. In studies comparing the Griffiths and the Bayley Scales of Infant Development (Bayley, 1969), the Griffiths consistently yields values approximately 10 points higher on both the mental and the motor scales. Thus, Griffiths and Bayley scores are not directly comparable (Ramsay & Fitzhardinge, 1977).

Family Outcome Measure. An instrument worthy of consideration is the Parenting Stress Index (PSI) (Abidin, 1983). Since the PSI

is a relatively new instrument, the literature pertaining to its use is not extensive. Available findings, however, support its utility in early intervention research (Loyd & Abidin, 1985).

First of all, the instrument is straightforward and the content addresses many of the issues that are essential to raising a handicapped child. Although it was not devised specifically for assessing parents who are raising a handicapped child, the items have a great deal of face validity for this group. Another very important consideration is that the PSI not only captures family characteristics pertaining to stress—financial stress, the family's emotional and physical state—but also provides for the assessment of specific characteristics of the child.

The PSI consists of 101 items and yields three major domain scores—a child domain, a parent domain, and an optional life stress scale. Respondents rate each item on the scale by indicating “strongly agree” or “strongly disagree.”

Percentile ranks are used in interpretation of scale results. Values exceeding the 75th percentile reflect clinically significant levels of stress. The manual indicates that while an extremely low score has some significance, and can be interpreted, values of importance are those which exceed the 75th percentile on the subscales or the overall score.

In the standardization sample of 534 parents, more than 90% of the parents were white, while only 6% were black. In addition, the sample is probably more representative of the higher educational range as it was developed in a university town. The age range of the children was from 1 to 19 years. However, the mean is about 14 months. Abidin thus indicates that this instrument is most suited for children under 3 years of age.

The alpha reliabilities for the items in the scale are adequate. The intercorrelation matrix of the subscales shows acceptable values. The test-retest reliability of the instrument was .95 for the total stress score. Thus, there is evidence that the instrument has reasonable psychometric properties.

The instrument also appears to have a substantial degree of discriminant validity. The PSI has been found to identify parents in terms of significantly higher stress scores as a function of raising a handicapped child (Kazak & Marvin, 1984). Any instrument used to assess the adaptation of families with handicapped children should capture this characteristic.

Rebecca R. Fewell, PhD, University of Washington

Approach to Measures Selection. Several critical factors should be considered in instrument selection, including: (1) the purpose for testing; (2) the uses of the data; (3) the examiners; (4) the children to be assessed; (5) those who are to receive the results; and (6) the fiscal and material resources available for the testing.

The State of Washington test standards were used in my selection of a test. The test must: (1) measure skills in one or more of the five areas of cognition, communication, fine motor, gross motor, and social-emotional development; (2) be properly standardized; (3) provide scores expressed by a measure of central tendency and variance; and (4) have a test-retest or split-half reliability level of .80 or greater.

Child Outcome Measure. The Battelle Developmental Inventory (BDI) (Newborg, Stock, Wnek, Guidubaldi, & Svinicki, 1984) meets the requirements identified. The BDI is a nationally standardized, individually administered developmental battery for children from birth to 8 years of age. The battery includes a screening test and scales in five domains: Personal-social, adaptive, motor, communication, and cognition. Domains are further divided into subdomains.

The BDI screening test can be given in 20 minutes and the entire battery takes about 1-1/2 to 2 hours. The test kit includes individual booklets for each domain, the screening test, protocols, and the Examiner's Manual. However, it is necessary to secure testing materials. Items are administered directly to the child or observed spontaneously, or inquiries are made of parents. Unlike many early childhood tests, the BDI meets the psychometric requirements of the Standards for Educational and Psychological Tests. The following practical aspects and applications make it appealing to programs:

1. A recently standardized screening test and comprehensive assessment across five domains.
2. Developmental sequence appropriate for IEPs.
3. Can be administered by program staff.
4. Manual is thorough and easy to follow.
5. Provides adaptations for the handicapped.
6. Can be used to measure individuals or groups.
7. Scoring system of 0, 1, and 2, permits assessment of change in moderately and severely handicapped students.

6. TOPICS IN EARLY CHILDHOOD SPECIAL EDUCATION 6:2

8. Test protocol provides summary profile across domains.
9. Commercially available at reasonable cost.

Family Outcome Measure. The following criteria were used to select a family outcome measure: (1) the instrument must measure the impact of a handicapped child on the family yet be broad in the issues and concerns addressed; (2) data must be available to support its technical qualities; (3) the instrument must be easy to administer, score, and interpret; (4) the measure must be available and reasonable in price; and (5) the measure must require minimum administration time.

The Questionnaire on Resources and Stress-Friedrich Edition (QRS-F) (Friedrich, Greenberg, & Crnic, 1983) closely complies with these criteria. The QRS-F is a 52-item questionnaire derived from Holroyd's (1974) Questionnaire on Resources and Stress (QRS). The original 285-item true/false scale was developed to measure the impact of a developmentally delayed, handicapped, or chronically ill child on other family members.

Friedrich et al. (1983) recognized the problems of length and technical quality in the QRS and developed a shorter and psychometrically stronger inventory, the QRS-F. The QRS-F has four factors: (1) parent and family problems, (2) pessimism, (3) child characteristics, and (4) physical incapacitations.

If this scale is selected for use in a program, the user might consider a 0-5 point scale rather than the true/false version of the QRS and QRS-F. This addition would not change the quality of the scale but it would provide the staff with finer discriminations and more useful information. In summary, the QRS-F has the following characteristics that suggest it is appropriate for use in programs for young handicapped children and their families:

1. Designed specifically for families with handicapped children.
2. Data are available to support its technical aspects.
3. Questionnaire is easy to give, available in the literature, and cost is minimal.
4. Has been effective in measuring parent change resulting from intervention over time.

Michael Lewis, PhD, Rutgers Medical School

Approach to Measures Selection. I will suggest some specific measures that will differ from the types of measures that might normally

be used in a program evaluation. I do so with the understanding that they are unlikely to be the types of measures that many would readily adopt in a large program evaluation.

Child Outcome Measure. One question we have to ask is what to measure as relevant to task. I am really concerned about this question because many programs select a measure of general competence. Now, in the period from birth to 3 years, we know that general competence measures are poor. One reason for this is that the test-retest reliability of measures for young infants is poor. That is, infants are quite variable as a function of state, of situation, of the nature of the examiner, or even the temperature of the room. Another reason is that the item pool is poor. We know that the items in this age range are heavily weighted toward motor milestones. We know that with the handicapped child, the use of a heavily weighted motor set of items is not going to measure the general competence we seek. Third, it has also been argued that the nature of the general competence that we wish to measure in infancy has nothing to do with its measure in early childhood or adulthood.

While I want to avoid general competence, I do want to talk about cognitive skills. What kinds of cognitive skills do we wish to consider in the infancy period? It would seem to me that one of the consequences of intervention, if successful, should be a general increase in the cognitive capacity to attend to and take in information from the environment. Attention predicts well to subsequent development (Lewis & Brooks-Gunn, 1981), and is relatively easy to measure (Lewis, 1982).

The assessment of attentional skills would require that you place a child in a situation in which you can observe the child directly, and take fairly exact measurements. For example, if you are looking at attention in the visual mode, you have to observe the child's orienting behavior toward a target. The score that you obtain from such observations can be expressed in a fashion that is easy to use.

There is a good deal of convergent evidence supporting a general cognitive measure such as attention. If you want to find out how children are developing, you should measure them directly and use careful measurement procedures. In a large-scale evaluation program it might be difficult to collect such direct measures, and it certainly would be costly. However, in my experience it has been possible to collect information on attentional skills in a large project.

Family Outcome Measure. It comes as no surprise that we have learned that it is necessary to know both about the child and the

environment in order to understand development (Lewis, Feiring, McGuffog, & Jaskir, 1984). Thus, we cannot afford *not* to look at families directly. For example, we proposed to study families of handicapped children in interaction at the dinner table. This proposal was based on our previous work with normal children and their families (Lewis & Feiring, 1982). At present, we are working on a longitudinal study of 150 children and their families. We have followed them from birth to 12 years of age and plan to follow them through adolescence. We saw the children and their families at dinner when the children were 3 years of age. In our recent work, we looked at the patterns of mothers, fathers, and children talking to each other at the dinner table and we observed important relationships between mother-father language patterns and the child's language development. Through direct measurement, we have found that family interactions are related to children's development (Lewis & Feiring, in press). Again, although direct measurement is difficult, it is likely to result in a better understanding of the child's development than measures that rely on parental report. And although observational measures are difficult in a large-scale study, we should be prepared to use such techniques.

Samuel J. Meisels, EdD, University of Michigan

Approach to Measures Selection. Two methodological constraints come to mind in selecting measures for an early intervention program. The first constraint has to do with the inappropriateness of the traditional treatment/control group design for assessing program impact on young handicapped children. If there is no simple and immediate means for evaluating the impact of an intervention in relation to a "no-intervention" group, this fact should influence the selection of assessment measures. The second constraint has to do with the absence of a major resource pool of instruments that are developmentally appropriate and have been psychometrically standardized for populations of disabled children.

Also influencing my recommendations is the problem of what to do with the data once they have been collected. If we wish to measure change in both the child and the family but we cannot conduct a classic treatment/control group design, then we not only have to consider which measures to use, but how these measures can be used to demonstrate program effectiveness.

Child Outcome Measure. The instruments principally used to assess the impact of early intervention programs fall into two categories, typically described as norm-referenced instruments or criterion-referenced instruments, respectively.

The most appropriate instruments for assessing change in early intervention programs are criterion-referenced instruments, for example, the Battelle Developmental Inventory (BDI) (Newborg et al., 1984). One consideration in the use of criterion-referenced tests is the decision about what to do with the data from these instruments because they, by definition, yield individual profiles on children that then must be aggregated in order to assess program effectiveness. One approach to managing all of this information is the use of a method known as goal attainment scaling, or GAS (Carr, 1979; Simeonsson, Huntington, & Short, 1982). GAS enables one to assess the degree to which the goals established for each child have been attained. One uses a set of objectives, assigns weights to the objectives, develops a set of expected outcomes for each objective, scores the outcomes, and calculates a summary score of the outcomes across the objectives. Both individual progress and program effectiveness can be assessed. Goal attainment scaling provides a metric, or a way of making sense of individual change, in a wide variety of areas. That is, the change, or overall improvement scores, is not dependent on specific goals or on the theoretical or methodological approaches selected. To the extent that a test or measurement instrument has been used extensively with various handicapped populations (e.g., the BDI), GAS can be used to transform individual scores into meaningful group data (Simeonsson et al., 1982).

One problem in using Goal Attainment Scaling is the validity of setting goals and assigning favorable outcomes. What is needed to overcome this problem is a series of large N studies of handicapped populations that are involved with criterion-referenced measurements and activities. Such studies would allow the establishment of standards across groups of children, within types of handicapping conditions, and for specific domains of skills.

Family Outcome Measure. Families are extremely complex entities and can no more be assessed by means of a single unidimensional instrument than can children. To select a family measure, it would be essential to know what can be expected to change as a result of the intervention program. Moreover, it would be important to select a scale with several subtests rather than a single instrument. While the

Parenting Stress Index (PSI) (Abidin, 1983) is multidimensional, its hypotheses are too broad for many researchers and its format is too formidable for many parents. Another source of family measures comes from the work of Olson and McCubbin at Minnesota (Olson et al., 1982). They have developed a range of family scales and inventories from which a selection can be made to match individual hypotheses and specific population characteristics. In particular, two of the scales, the FACES II Scale and the Family Satisfaction Scale, are particularly useful with families of disabled children. The factors assessed by these scales are family cohesion and family adaptability. These two brief questionnaires focus on emotional bonding, independence, family boundaries, coalitions, and use of time. However, although these scales are extremely well developed and researched, they provide only part of the family picture. It is necessary to return to one's hypotheses and program goals when selecting family outcome measures so that neither too little nor too much is expected of these very limited instruments.

**Jack P. Shonkoff, MD, University of Massachusetts
Medical School**

Approach to Measures Selection. Appropriate instrument selection demands consideration of several issues that are often neglected. The first concerns the ordering of tasks. We often approach program evaluation by looking first at available instruments and then deciding to use those that seem best. I suggest that we spend more time thinking about the outcomes we wish to measure before we even begin to consider the selection of instruments. Toward this end, my colleagues and I have spent a good part of the past year talking to service providers about what they feel are their most important intervention objectives. We also convened two parent-advisory groups to explore parallel issues. This exercise proved fruitful. Only after we had struggled with questions about what variables ought to be measured did we feel prepared to consider the selection of specific instruments.

The second issue that must be addressed is the difference between short-term and long-term outcomes. Although the important long-term effects of services for children are, by necessity, child-oriented, short-term influences on families may be critical, and in fact, might be the mediating variables that have the greatest impact on long-term outcomes for children. Moreover, we should not minimize the independent value

of family impacts themselves as legitimate goals of early intervention efforts.

One additional consideration is that we should select measures that focus not only on deficits, but also on adaptive behaviors. Only after the above decisions are made should we proceed to look for those instruments with acceptable psychometric properties that meet such practical considerations as cost and the logistics of administration.

Child Outcome Measure. If only one measure were to be selected, I would focus on social competence and therefore suggest the newly revised Vineland Adaptive Behavior Scales, which have been designed to assess, "personal and social sufficiency in children and adults" (Sparrow, Balla, & Cicchetti, 1985). The Vineland has been recommended for use in research and program evaluation, as a diagnostic instrument, and for program planning. Versions of the scale can be used to develop Individualized Educational Plans (IEPs) or for screening and diagnostic purposes. Three formats are available. The most appropriate for evaluating the impacts of early intervention services is the Survey Edition. This is a 297-item interview/questionnaire that takes anywhere from 20 to 60 minutes to administer. The Scales assess four domains, with 11 subdomains. The four domains are *communication*, *daily living skills*, *socialization*, and *motor skills*. An optional maladaptive behavior scale can be administered beginning at age 4 or 5. An elaborate standardization process is clearly described in the manual. The Scales can be used from birth to 18 years, 11 months, thereby facilitating long-term follow-up with the same instrument. Supplementary norms are available for populations of individuals with mental retardation, physical disabilities, and emotional disturbance, as well as children with hearing or visual impairments.

Each of the subdomains generates an adaptive level and an age equivalent score. Psychometric data for this instrument demonstrate adequate split-half, interrater, and test-retest reliability as well as construct, content, and criterion-related validity.

Family Outcome Measure. If only one family measure were to be selected, I would like to suggest the Impact on Family Scale developed by Stein and Riessman (1978). This instrument assesses change in family behaviors that may be attributed to a child with a disability. The original scale was developed and normed for a population of children with chronic illness, which included some youngsters with developmental

disabilities such as spina bifida. Because of its orientation toward a chronically ill population, modifications in the wording of some of the items are necessary to address issues related to disability.

The Impact on Family Scale measures four factors. The first factor is financial burden, which for many families may be a major stressor. The second is a familial-social support factor that addresses social interactions both within and outside the family. The third factor is a personal strain factor that addresses subjective feelings of family stress. The fourth factor is a positive one, mastery, which addresses change in a parent's sense of mastery over the management of his or her disabled child.

I would like to briefly mention two additional measures that address both family and child issues. The Nursing Child Assessment Teaching Scale (Barnard, 1978) can be used to evaluate important parent and child behaviors in an interactional context. The Parenting Stress Index (PSI) (Abidin, 1983) also addresses the interactional nature of the parent-child system. It allows the program evaluator not only to assess change in the parent's perception of caregiving stress, but also to discover the differential contributions of the child, the parent, and the family environment to stress. In that sense, the PSI can be regarded simultaneously as a family measure, a child measure, and an interactional measure.

In closing, I would like to reaffirm my commitment to the importance of multiple measures and my belief that early intervention impact studies must address more explicitly family-oriented dependent variables as mediators of child-oriented effects as well as important outcomes in their own right.

Discussion

The purpose of this assessment conference was to identify the best available instruments that can be used across a variety of programs to evaluate child and family outcomes of early intervention for handicapped children. Although it is clear that no single instrument should be the only tool for assessing either child or family functioning, applied researchers, state coordinators, and program directors often need to collect common data for children enrolled in different programs. Such measures would yield valuable information as a part of a more comprehensive assessment battery that would be unique to each program.

Several specific child outcome measures were selected, including

the Griffiths, the BDI, and the Vineland. Although the suggestion that these instruments represent the best currently available tools is valuable, it is also important to recognize that these instruments must be viewed cautiously. The BDI is a new instrument which has not yet been subjected to extensive research. While the psychometric data presented in the manual are commendable, several characteristics of this test may influence test outcomes. For example, the lack of standardized test materials and the use of three potential sources of information for scoring may present serious problems. The Griffiths, a British scale, has not been used extensively in the United States. The interpretation of the norms may be problematic due to cultural differences. The revised Vineland is also a new instrument which, like the BDI, has adequate psychometric properties. However, reliance on caregiver reports may be unacceptable in programs that serve children whose parents are developmentally disabled or may for some other reason be unreliable reporters of their child's progress.

The family outcome measures recommended require similar cautions. The area of family assessment is clearly less well developed than the area of child assessment, and thus the measures recommended were all relatively new and untried. The specific family outcomes that are expected to result from intervention are also quite open to debate. This further confounds the selection of an appropriate family measure. For example, the PSI seems to be attractive, as it assesses a variety of outcomes. Other scales, such as the QRS-F, may be adequate for some programs, but may be too narrow to tap the outcomes of others. The need to include complementary outcome measures thus seems particularly relevant to the area of family assessment. In addition, because many of the measures selected are new and untried, users should identify opportunities to gather additional information about them.

Concerns about the type of data to be collected were raised. An extensive analysis of the relative merits of norm-referenced assessment, criterion-referenced assessment, or direct observation was beyond the scope of this conference. Such an analysis would probably reveal that each method has both advantages and disadvantages when used for program evaluation purposes. Thus, rather than debate the type of data that are most appropriate for conducting a program evaluation, it would be more fruitful to begin by clarifying the expected outcomes of an intervention program, the purposes of the program evaluation, and the resources available for it. The selection of an assessment procedure and specific instruments that meet these program requirements can then follow.

Authors' Note

The work reported in this article was carried out in part with funds from the U.S. Department of Education (Contract No. 300-85-0173) to the Early Intervention Effectiveness Institute at Utah State University.

References

- Abidin, R. R. (1983). *Parenting Stress Index*. Charlottesville, VA: Pediatric Psychology Press.
- Barnard K. (1978). *Nursing Child Assessment Teaching Scales*. Seattle, WA: University of Washington.
- Bayley, N. (1969). *Bayley Scales of Infant Development*. New York: Psychological Corporation.
- Brooks-Gunn, J., & Lewis, M. (1983). Screening and diagnosing handicapped infants. *Topics in Early Childhood Special Education*, 3(1), 14-28.
- Carr, R. A. (1979). Goal Attainment Scaling as a useful tool for evaluating progress in special education. *Exceptional Children*, 46, 88-95.
- Casto, G., & Lewis, A. (in press). Selecting outcome measures in early intervention. *Journal for the Division of Early Childhood*.
- Friedrich, W., Greenberg, M., & Crnic, K. (1983). A short form of the questionnaire on resources and stress. *American Journal of Mental Deficiency*, 88(1), 41-48.
- Garwood, G. (1982). (Mis)use of developmental scales in program evaluation. *Topics in Early Childhood Special Education*, 1(4), 61-69.
- Griffiths, R., (1970). *The abilities of young children*. London: Association for Research in Infant and Child Development.
- Holroyd, J. (1974). The questionnaire on resources and stress: An instrument to measure family response to a handicapped family member. *Journal of Community Psychology*, 2, 92-94.
- Kazak, A. E., & Marvin, R. S. (1984). Differences, difficulties, and adaptation: Stress and social networks in families with a handicapped child. *Family Relations*, 33, 67-77.
- Lewis, M. (1982). Attention as a measure of cognitive integrity. In M. Lewis & L. Taft (Eds.), *Developmental disabilities: Theory, assessment, and intervention* (pp. 185-212). New York: F.P. Medical and Scientific Books.
- Lewis, M., & Brooks-Gunn, J. (1981). Visual attention at 3 months as a predictor of cognitive functioning at 2 years of age. *Intelligence*, 5, 131-140.
- Lewis, M., & Feiring, C. (1982). Some American families at dinner. In L. Laosa & I. Sigel (Eds.), *The family as learning environments for children* (Vol. 1, pp. 115-145). New York: Plenum.
- Lewis, M., & Feiring, C. (in press). Direct and indirect effects during family interaction at dinner. In S. Feinman (Ed.), *Social influences and behavior*. New York: Plenum.
- Lewis, M., Feiring, C., McGuffog, C., & Jaskir, J. (1984). Predicting psychopathology in 6 year olds from early social relations. *Child Development*, 55, 123-136.
- Loyd, B. H., & Abidin, R. R. (1985). Revision of the Parenting Stress Index. *Journal of Pediatric Psychology*, 10, 169-177.
- Newborg, J., Stock, J., Wnek, L., Guidubaldi, J., & Svinicki, J. (1984). *Battelle Developmental Inventory*. Allen, TX: DLM Teaching Resources.

- Olson, D.H., McCubbin, H. I., Barnes, H., Larsen, A., Muxen, M., & Wilson, M. (1982). *Family inventories*. St. Paul: University of Minnesota.
- Ramey, C., Campbell, F., & Wasik, B. (1982). Use of standardized tests to evaluate early childhood special education programs. *Topics in Early Childhood Special Education*, 1(4), 51-60.
- Ramsay, M., & Fitzhardinge, P. M. (1977). A comparative study of two developmental scales: The Bayley and the Griffiths. *Early Human Development*, 1/2, 151-157.
- Shonkoff, J. (1983). The limitations of normative assessments of high risk infants. *Topics in Early Childhood Special Education*, 3(1), 29-43.
- Simeonsson, R. J., Huntington, J. S., & Short, R. J. (1982). Individual differences and goals: An approach to the evaluation of child progress. *Topics in Early Childhood Special Education*, 1, 71-80.
- Sparrow, S. S., Balla, D. A., & Cicchetti, D. V. (1985). *Vineland Adaptive Behavior Scales*. Circle Pines, MN: American Guidance Service.
- Stein, R., & Riessman, C. K. (1978). *Impact on Family Scale*. New York: Albert Einstein College of Medicine.
- Zigler, E., & Balla, D. (1982). Selecting outcome variables in evaluations of early childhood special education programs. *Topics in Early Childhood Special Education*, 1(4), 11-22.

Sources

- | | |
|---|--|
| Battelle Developmental Inventory | Parenting Stress Index |
| DLM Teaching Resources | Pediatric Psychology Press |
| One DLM Park | 2915 Idlewood Dr. |
| PO Box 4000 | Charlottesville, VA 22901 |
| Allen, TX 75002 | |
| FACES II and Family Satisfaction Scale | Questionnaire on Resources and Stress (QRS) |
| Family Stress and Coping Project | Dr. Jeanne Holroyd |
| Department of Family Social Science | Neuropsychiatric Institute, UCLA |
| University of Minnesota | 760 Westwood Plaza |
| Saint Paul, MN 58108 | Los Angeles, CA 90024 |
| Griffiths Mental Development Scale Test Agency | Questionnaire on Resources and Stress (QRS-F) |
| Cournswood House | In Friedrich, W., Greenberg, M., & Crnic, K. (1983). A short form of the questionnaire on resources and stress. <i>American Journal of Mental Deficiency</i> , 88(1), 41-48. |
| North Wycombe | |
| Bucks, England | |
| Impact on Family Scale | |
| Dr. Ruth Stein | |
| Department of Pediatrics | |
| Albert Einstein College of Medicine of Yeshiva University | Vineland Adaptive Behavior Scales |
| 1300 Morris Park Ave. | American Guidance Service |
| Bronx, NY 10461 | Publishers' Building |
| | Circle Pines, MN 55014 |