

ATTACHMENT 6—CHILD ASSESSMENT MEASURES

A. The Child Assessment Measures

All of the other major subject areas and indeed almost all of the questions in this survey have been asked in prior survey rounds. There has been little respondent resistance to any of the items, as witnessed by the very high response rate, continued respondent cooperation, and high quality data collected. The child assessment questions were all asked in the 1986, 1988, 1990, 1992, and 1994–2012 survey rounds. Discussions with NORC field staff indicate that these materials are greeted in an enthusiastic manner by the respondents and their children. The respondents view the assessment component of the survey as a positive experience. Response rates have been very high on these assessments, generally over 90 percent, and analysis suggests that the quality of the data is very good. We include here the information that describes the assessments in detail, including estimates of their established reliability and validity.

The decisions about which child aptitude tests to ask of the sample members were carefully considered from a number of perspectives. These decisions were made jointly by staff from the National Institute of Child Health and Human Development (NICHD), the Ohio State University Center for Human Resource Research (CHRR), and nationally recognized panels of experts from medicine and the social sciences. Each of these experts has specialized areas of interest central to this study. A list of these individuals and their affiliations is shown on page 13 in table 3.

For a test to be considered for possible inclusion in this study, it needed to meet a number of specific criteria. The obvious one being that it is essential as either an input or outcome measure (or both) for analyses of interest to NICHD, which provides funding to the Bureau of Labor Statistics for the Child Survey. All of the aptitude tests chosen met the following criteria.

- (a) They are “tried and true” tests that have been extensively used by data collectors in a variety of social, economic, and cultural situations. They have been used in household settings similar to the interviewing procedures used with the NLSY79, and they have been administered by non-technical interviewing personnel to a full cross-section of American youth, middle class and economically disadvantaged nonblack/non-Hispanics as well as black and Hispanic youth.
- (b) All of the tests are recognized by the social science community as well-established. Reliability and validity statistics indicate that they are all highly reliable and valid. High reliability means that if the same individual is repeatedly given the same test, that person will repeatedly have similar scores. Validity means that the test measures what it purports to measure, and it is generally determined by comparing results on the given test with results for the same individual on a different test whose validity has already been well established. All of the tests are rated highly in *Tests in Print* (Vol. 3, Burroughs, 1983), a highly regarded testing manual that rates all of the major aptitude tests.
- (c) All of the tests chosen for use are inexpensive to administer, require very little equipment (important for tests to be administered in a home setting), and are relatively short.
- (d) The utility and appropriateness of all of the tests have been considered from both longitudinal and cross-sectional perspectives. First, the participants in the questionnaire development process have carefully ensured that tests are included that cover all the critical cognitive, personality, and physical health dimensions for children at all maturational stages. Second, every effort has been made to assure that the tests complement each other analytically from a longitudinal perspective. Finally, to the maximum extent feasible, every effort has been made to include basic cognitive and personality scales that can, in a cross-sectional context, be compared across age groups at one point in time.

- (e) None of the tests involve any physical or psychological risk to either the child or any other family member. In all instances, the mother, who is the sample respondent, will be intimately involved in the testing procedures and indeed, for a large proportion of the tests, the questions are addressed directly to the mother. If at any time there is any reticence regarding a procedure by either a child or the respondent, testing will immediately cease. We will in no instance jeopardize the quality of the material being collected for the Department of Labor or the basic integrity of the longitudinal design.

B. Summary of Child Aptitude Measures To Be Used

This section presents summary descriptions of the tests and measures to be used in the 2014 survey round. In addition to briefly describing the measure, we will include summary documentation about its validity and reliability, its utility for evaluating aptitudes of minority respondents, and the estimated time required to give the test.

1. Home Observation for Measurement of the Environment (HOME)

This set of scales measures the nature and quality of the child's developmental environment. Some of the items are maternal self-reports, while others are interviewer observations. There are three versions: for infant (0–2 years), preschool (3–5 years), and elementary-aged (6 years and older) children. The instrument yields a total score reflecting the quality of the home environment and mother-child relationship, two general indicators of emotional support and cognitive stimulation, and subscales measuring several processes of the home environment. The HOME Inventory's total score, factor scores, subscale scores, and item scores have been used by previous researchers. The scales have been previously used across a full range of ethnic and SES subgroups.

The infant version consists of six categories: maternal emotional and verbal responsiveness, maternal acceptance of and involvement with her child, materials for play, organization of the environment, and variety of stimulation. The preschool version consists of eight categories: maternal warmth and acceptance; organization of the environment; variety of stimulation; modeling of maturity; learning, language, and academic environmental processes; and avoidance of harsh discipline. The elementary version also consists of eight categories: maternal responsiveness, emotional climate, organization of the environment, modeling of maturity, family participation in growth experiences, paternal involvement, opportunities for growth, and provision for active stimulation.

This widely used battery was created by Dr. Bettye Caldwell and Dr. Robert Bradley of the University of Arkansas at Little Rock. For each HOME subscale, three items for the NLSY79 child survey were selected based on previous factor analysis (Bradley and Caldwell, 1979) and Dr. Bradley's consultation.

The HOME is considered very reliable. Bradley and Caldwell (1981) report inter-rater reliabilities from six studies in the high .80s to low .90s. Bradley, Caldwell, and Elardo (1979) found that six-month test-retest subscale correlations ranged from .45 to .87. Studying children from 6 to 42 months of age, Yeates et al. (1983) found twelve-month test-retest reliabilities from .43 to .68, and two-year test-retest reliabilities of .38 to .56. Ramey et al. (1984) reported two-year test-retest reliabilities of .56 and .57. Van Doornick et al. (1981) reported high total score stability ($r = .86$) among siblings tested at least ten months apart. In the NLSY79, Mott and Quinlan (1991) found that reliability for the subscales of the HOME used in the NLSY79 was 0.59 for children under three, rising to 0.69 for children three to six, 0.73 for children six to nine, and 0.68 for children ten and over. Two-year test-retest reliabilities were 0.52 for children three to five and 0.59 for children six and over, similar to levels found in other studies.

Prior longitudinal research indicates that the HOME predicts later cognitive, social, and physical development. Yeates et al. (1983) longitudinally compared the predictive ability of the HOME relative to the predictive ability of maternal intelligence for child intellectual development at two, three, and four years of age, finding that maternal intelligence was initially more predictive, but by age four the quality of

the home environment was most predictive of cognitive development. The HOME is more predictive of subsequent cognitive development than is concurrently measured cognition (Elardo, Bradley, and Caldwell, 1975). When administered as early as two months of age, the HOME has correlated from .34 to .72 with intelligence tests subsequently administered as late as four-and-a-half years of age. The HOME at one and two years correlated (.33 to .65) with academic achievement in the first through fourth grades of school (Bee et al., 1982; Bradley and Caldwell, 1976, 1980, 1984; Elardo, Bradley, and Caldwell, 1975; Van Doornick et al., 1981).

Besides these strong predictive correlations with subsequent mental development, the HOME has also indicated a variety of developmental risks and delays such as clinical malnutrition, lead burden, failure-to-thrive, socio-cultural retardation, language delay, developmental delay, and poor academic achievement (Elardo and Bradley, 1981). The HOME is moderately related to SES and parental education ($r = .2$ to $.6$, Elardo and Bradley, 1981). A meta-analysis of the correlation between SES and intelligence found that measures of the home environment accounted for four to eleven times as much of the variation in academic achievement and intelligence (median $r = .55$) as did standard measures of SES. The homes of divorced working mothers provided less cognitive stimulation and emotional support according to the HOME Inventory than did the homes of married (working or nonworking) mothers. Six studies found relationships between temperamentally difficult and unsociable infants, and decreased cognitive stimulation and emotional support available in their homes. In general, this has been the most widely used of all the NLSY79 child assessments. As in previous rounds, the HOME will be administered for the full sample of children.

References

- Bee, H.; Bernard, K.; Eyres, S.; Grey, C.; Hammond, M.; Spietz, A.; Snyder, C.; and Clark, B. "Prediction of IQ and Language Skill from Perinatal Status, Child Performance, Family Characteristics, and Mother-Infant Interaction." *Child Development* 53 (1982): 1134-1156.
- Bradley, Robert H. and Caldwell, Bettye M. "Home Observation for Measurement of the Environment: A Revision of the Pre-School Scale." *American Journal of Mental Deficiency* 84 (1979): 235-244.
- Bradley, Robert H. and Caldwell, Bettye M. "The Relation of Infants' Home Environments to Mental Test Performance at Fifty-Four Months: A Follow-Up Study." *Child Development* 47 (1976): 1172-1174.
- Bradley, Robert H. and Caldwell, Bettye M. "The HOME Inventory: A Validation of the Pre-School Scale for Black Children." *Child Development* 52 (1981): 708- 710.
- Bradley, Robert H. and Caldwell, Bettye M. "The Relation of Home Environment, Cognitive Competence, and IQ among Males and Females." *Child Development* 51 (1980): 1140-1148.
- Bradley, Robert H. and Caldwell, Bettye M. "The Relation of Infants' Home Environment to Achievement Test Performance in First Grade: A Follow-Up Study." *Child Development* 55 (1984): 803-809.
- Bradley, Robert H.; Caldwell, Bettye M.; and Elardo, R. "Home Environment and Cognitive Development in the First Two Years: A Cross-Lagged Panel Analysis." *Developmental Psychology* 15 (1979): 246-250.
- Caldwell, Bettye M. and Bradley, Robert H. Home Observation for Measurement of the Environment. Little Rock, AR, 1984.
- Elardo, Richard D. and Bradley, Robert H. "The Home Observation for Measurement of the Environment (HOME) Scale: A Review of Research." *Developmental Review* 1 (1981): 113-145.
- Elardo, Richard D.; Bradley, Robert H.; and Caldwell, Bettye M. "The Relation of an Infant's Home Environment to Mental Test Performance from Six to Thirty-Six Months." *Child Development* 46 (1975): 71-76.

- Mott, Frank L. and Quinlan, Steven V. *Children of the NLSY79: 1988 Tabulations and Summary Discussion*. Columbus, OH: Center for Human Resource Research, March 1991.
- Ramey, C.; Yeates, K.; and Short, E. "The Plasticity of Intellectual Development: Insights from Preventive Intervention." *Child Development* 55 (1984): 1913-1925.
- Van Doornick, W.; Caldwell, Bettye M.; Wright, C.; and Frankenburg, W. "The Relationship between Twelve-Month Home Stimulation and School Achievement." *Child Development* 52 (1981): 1080-1083.
- Yeates, K.; MacPhee, D.; Campbell, F.; and Ramey, C. "Maternal IQ and Home Environment as Determinants of Early Childhood Intellectual Competence: A Developmental Analysis." *Developmental Psychology* 19 (1983): 731-739.

2. Wechsler Intelligence Scale for Children–Revised: Digit Span Subscale

There are two parts to this measure of short-term memory for children age 7 and older. First, the child listens to and repeats a sequence of numbers said by the interviewer. In the second part, the child listens to a sequence of numbers and repeats it in reverse order. In both parts, the length of the sequence of numbers increases as the child responds correctly.

This subscale is from the revised Wechsler Intelligence Scale for Children (WISC-R) published by the Psychological Corporation. The WISC-R is one of the most respected measures of child intelligence. The Digit Span score is a good measure of short-term memory and attentiveness for children 7 and older. It correlates ($r = .45$) with the PIAT's reading recognition. Its parallel form reliability is about 0.53. It is administered to the NLSY79 children age 7–11.

References

- Jackson, N. and Myers, M. "Letter Naming Time, Digit Span, and Precocious Reading Achievement." *Intelligence* 6 (1982): 311-329.
- Wechsler, D. *Wechsler Intelligence Scales for Children–Revised*. New York: The Psychological Corporation, 1974.

3. Peabody Picture Vocabulary Test–Revised (PPVT-R)

The PPVT-R measures receptive vocabulary knowledge of orally presented words by directing the child to nonverbally select the picture (out of four) that best describes the word's meaning. The PPVT-R is among the most highly recognized and established indicators of verbal intelligence and scholastic aptitude across childhood (from age 3 and older). Since 1978, it is the fourth most frequently cited test in Mitchell's *Tests in Print* (1983). It is currently administered to NLSY79 children age 4–5 and 10–11.

Numerous studies have replicated the reliability estimates from the PPVT-R's standardization sample (4,200 children ages 2 years and 6 months to 18 years and 11 months): a median split-half reliability of .80 (ranging from .67 to .88), a median parallel form reliability of .70 (ranging from .50 to .87), and a median 9- to 31-day test-retest reliability of .78 (.52 to .90; Dunn and Dunn, 1981). Goldstein et al. (1970) reported a 21-month test-retest reliability of .61 among 160 disadvantaged three- to seven-year-olds, and Costello and Ali (1971) found a two-week retest reliability of .77 among thirty-six black preschoolers.

The PPVT-R demonstrates a high construct validity with a variety of intelligence tests. Its median correlation with other vocabulary tests was .71 (based on 55 criterion validity coefficients, ranging from .20 to .89); with other individual intelligence tests it was from .38 to .72 (based on 291 correlations ranging from -.16 to .92). Its correlation was higher with the Binet and Wechsler tests than with less

reputable tests; and correlations were higher with verbal intelligence (.66 to .71) than with performance (.46 to .65; Dunn and Dunn, 1981).

Because it demonstrates high predictive validity with a variety of achievement measures, combined with other information, the PPVT is an extremely important predictor of early and middle school outcomes. Median correlation with math achievement was .50 (based on 16 correlations ranging from .27 to .77 with the Wide Range Achievement Test (WRAT), California Achievement Test (CAT) and PIAT); with language achievement it was .44 (16 correlations, from .02 to .66 with the WRAT, PIAT, CAT, with Metropolitan Achievement Test (MAT)); with reading comprehension it was .63 (seven from .42 to .70 with the CAT and PIAT); and with reading recognition it was .38 (WRAT) and .52 (PIAT) (14 ranging from .01 to .71; Dunn and Dunn, 1981).

Zigler, Abelson, and Seitz (1973) found an inverse relationship ($r = -.53$) between the magnitude of the increased IQ on retest and the initial IQ estimate. This indicates that a disadvantaged preschooler's measured intelligence is influenced by anxiety and sociability during assessment, that emotional patterns are distinct from cognitive deficits and are measured by the temperament and interviewer relationships between sociability and cognitive performance.

References

- Costello, J. and Ali, F. "Reliability and Validity of PPVT Scores of Disadvantaged Preschool Children." *Psychological Reports* 28 (1971): 755-760.
- Dunn, L. and Dunn, L. *PPVT-R Manual*. Circle Pines, MN, 1981.
- Goldstein, L.; Collier, A.; Dill, J.; and Tilis, H. "The Effect of a Special Curriculum for Disadvantaged Children on Test-Retest Reliabilities of Three Standardized Instruments." *Journal of Educational Measurement* 1 (1970): 171-174.
- Lamb, M.; Garn, S.; and Keating, M. "Correlations between Sociability and Cognitive Performance among Eight-Month-Olds." *Child Development* 52 (1981): 711-713.
- Plant, W.T. and Southern, M.L. "The Intellectual and Achievement Effects of Pre-School Stimulation of Poverty Mexican-American Children." *Genetic Psychology Monographs* 86 (1972): 141-173.
- Radin, N. "Maternal Warmth, Achievement Motivation, and Cognitive Functioning in Lower-Class Pre-School Children." *Child Development* 42 (1971): 1560-1565.
- Robertson, G. and Eisenberg, J. *PPVT-R Technical Supplement*. Circle Pines, MN: American Guidance Service, 1981.
- Zigler, W.E.; Abelson, W.; and Seitz, V. "Motivational Factors in the Performance of Economically Disadvantaged Children on the PPVT." *Child Development* 44 (1973): 294-303.

4. Peabody Individual Achievement Test (PIAT)

The PIAT is a wide-range measure of academic achievement for children aged five and over that is widely known and used in research. It is perhaps the most widely used brief assessment of academic achievement having demonstrably high test-retest reliability and concurrent validity. From its national standardization sample, the PIAT's median one-month test-retest reliability is .74 for Math, .89 for Reading Cognition, and .64 for Reading Comprehension (Dunn and Markwardt, 1970). The concurrent validity coefficients for the three PIAT subscales were reported in the discussion of the PPVT-R. Of all psychological tests, the PIAT had the forty-second largest number of citations since 1978 in Mitchell's *Tests in Print* (1983). In addition to the raw and standard scores, percentiles, age equivalents, and grade equivalents are also available. Some of the PIAT's many correlations with measures of memory, the home environment, and intelligence have been previously mentioned. The PIAT subscales are administered to all eligible children ages 5-14. In this context, "eligible" means that a sufficient score

was achieved on the previous assessment to proceed to the next assessment. In particular, some children may not advance to the reading comprehension subscale.

(a) PIAT Math Subscale

This subscale measures ability in mathematics as taught in mainstream education. It consists of 84 multiple-choice items (each with four options) that increase in difficulty. It begins with such early skills as recognizing numerals and progresses to measuring advanced concepts in geometry and trigonometry. The child looks at the problem, then points to the answer.

(b) Peabody Reading Recognition Subscale

This subscale measures word recognition and pronunciation ability to indicate reading achievement. Children read a word silently, then say it aloud. Recognition contains 84 items (each with four options) that increase in difficulty from preschool to high school levels. Skills assessed include matching letters, naming names, and reading single words aloud.

(c) Peabody Reading Comprehension Subscale

This subscale measures the ability to derive meaning from sentences that are read silently. For each of the 66 items of increasing difficulty, the child silently reads the sentence once and then selects one of four pictures that best portrays the meaning of the sentence.

While the recognition and comprehension subscales cannot measure all dimensions of reading comprehension (our “ideal” objective requiring at least a half-hour of assessment time), they do measure two critical components of reading comprehension, word recognition and comprehension of the meaning of sentences. In Hammill and McNutt (1981), meta-analysis (8,239 coefficients from 322 studies) of reading correlates reports median concurrent correlations of .72 between recognition and composite reading, .72 between comprehension and composite reading, .85 between composite reading and general academic achievement, .74 between recognition and comprehension, and .62 between composite reading and math.

References

- Dunn, L. and Markwardt, F. *PIAT Manual*. Circle Pines, MN: American Guidance Service, 1970.
- Hammill, D. and McNutt, G. *The Correlates of Reading*. Austin, TX: PRO-ED, 1981.
- Naglieri, J. and Harrison, P. “McCarthy Scales, McCarthy Screening Test, and Kaufman’s McCarthy Short Form Correlations of the PIAT.” *Psychology in the Schools* 19: 149-155.
- Robertson, G. and Eisenberg, J. *PPVT-R Technical Supplement*. Circle Pines, MN: American Guidance Service, 1981.

5. Temperament Scales

Because the child’s temperament is partially a parental perception (Bates, 1980), the behavioral style of children is measured by a set of maternal-report items (for children younger than 7 years) and interviewer ratings (for children directly assessed). Because the review by Hubert et al. (1982) found no single instrument to be satisfactory, our scale is based on items from Rothbart’s IBQ, Campos and Kagan’s compliance scale, and other items from Dr. Joseph Campos. The maternal scale “How My Infant Usually Acts” addresses the activity, predictability, fearfulness, positive affect, and irritability of the 0–11-month-old child. “How My Toddler Usually Acts” addresses the fearfulness, positive affect, and irritability of the 12–23-month-old child. “How My Child Usually Acts” measures the compliance, affect, attachment, and sleep problems of children aged 2 years to 6 years and 11 months. The interviewer rates the child’s shyness at first meeting and at the end of the session; during the assessment, the interviewer records the

child's cooperation, interest and motivation, energy, persistence, and attitude about and rapport with the interviewer during the assessment.

Temperament is important to child development, personality development, the child's impact on family members, and the development of behavioral problems (Bates, 1980). These scales include dimensions such as sociability, mood, adaptability, and compliance, all factors that are components of Thomas's easy/difficult temperament construct and are precursors to personality development and social adjustment (areas measured by the Behavior Problems Index that is discussed below), social relations, and performance on tests such as the Motor and Social Development Scale and PPVT-R (Lamb, 1982).

As with adult personality measures, reviews of temperament (Bates, 1980; Campos et al., 1983; Hubert et al., 1982) contend that the perceiver plays a significant role, that mild to moderate inter-rater agreement is the rule; median parent-observer correlations of .2 to .4 in infancy increase to .3 to .6 by age two, median between-parent correlations are .4 to .6. They also state that moderate internal consistency (.2 to .8) and retest reliability (-.1 to .9) are present and that fair validity coefficients (.3 to .6) are found with a wide variety of criteria. Hubert et al. (1982) state that the most consistent and substantial relationship is found between temperamental difficulty and infant distress/fussiness with people. Published correlates include levels of neurotransmitters associated with stress, spectrographic analysis of cries, respiratory distress, and post mature birth syndromes; maternal anxiety, sociability, responsivity, and stress; family moves, employment changes, paternal childcare, birth of siblings; sensitivity to change and adversity, social communication, subsequent behavior disorders (such as delinquency and emotional disturbance), and cognitive and motor development. This assessment will be given to children age 0–6.

References

- Baydar, Nazli. "Reliability and Validity of Temperament Scales of the NLSY Child Assessments." *Journal of Applied Developmental Psychology* 16,3 (July-September 1995): 339-370.
- Bates, J. "The Concept of Difficult Temperament." *Merrill-Palmer Quarterly* 26 (1980): 299-319.
- Campos, Joseph J.; Barrett, Karen C.; Lamb, Michael E.; Goldsmith, H.; and Stenberg, Craig. "Socioemotional Development." In P. Mussen (ed.), *Handbook of Child Psychology*, Vol. 2, 4th Ed. New York: Wiley, 1983.
- Hubert, N.C.; Wachs, T.D.; Peters-Martin, P.; and Gandour, M.J. "The Study of Early Temperament: Measurement and Conceptual Issues." *Child Development* 53 (1982): 571-600.
- Lamb, Michael E.; Garn, S.; and Keating, M. "Correlations between Sociability and Motor Performance Scores in Eight-Month-Olds." *Infant Behavior and Development* 5 (1982): 97-101.
- McDevitt, S.C. and Carey, W.B. "Stability of Ratings vs. Perceptions of Temperament from Early Infancy to 1-3 Years." *American Journal of Orthopsychiatry* 51 (1981): 342-345.
- Menaghan, Elizabeth G. and Parcel, Toby L. "Measuring Temperament in a Large Cross Sectional Survey: Reliability and Validity for Children of the NLS Youth." Working Paper, Department of Sociology, The Ohio State University, 1988.
- Peters-Martin, P. and Wachs, T.D. "A Longitudinal Study of Temperament and Its Correlates in the First 12 Months." *Infant Behavior and Development* 7 (1984): 285-298.

6. Perceived Competence Scale for Children/Self-Perception Profile

This self-report magnitude estimation scale measures the child's perceived competence in the academic skill domain and the child's sense of general self-worth. Children age 12–14 answer the questions in these two domains by first selecting one of two alternatives describing how they usually act, and then indicating how true that alternative is of themselves.

There are many studies that have documented the importance of this scale as a predictor of important child outcomes and behaviors. For example, it has been shown to correlate highly with teacher ratings of children and with a child's achievement motivation. It has high internal reliability ($r = .73$ to $r = .86$) and high (nine month) test-retest reliability ($r = .8$). Mott and Quinlan (1991) found reliability scores of 0.68 in the NLSY79. Prior uses of the schedule suggest no apparent cultural bias.

References

Harter, Susan. "The Perceived Competence Scale for Children." *Child Development* 53 (1982): 87-97.

Harter, Susan. "Developmental Perspectives on the Self-Esteem." In M. Hetherington (ed.), *Social Development: Carmichael's Manual of Child Psychology*. New York: Wiley, 1984.

Harter, Susan. "Effectance Motivation Reconsidered: Toward a Developmental Model." *Human Development* 1 (1978): 34-64.

Mott, Frank L. and Quinlan, Steven V. *Children of the NLSY79: 1988 Tabulations and Summary Discussion*. Columbus, OH: Center for Human Resource Research, March 1991.

7. Behavior Problems Index

This widely-used scale was created by Dr. Nicholas Zill and James Peterson of Child Trends to measure the frequency, range, and type of childhood behavior problems. These items are from Achenbach's (1978) Child Behavior Checklist developed at the NIMH Laboratory of Developmental Psychology, one of the most thoroughly researched and widely used parental report measures of behavior problems in childhood. Some of these items have been used in earlier national surveys (Cycle II of the Health Examination Surveys: 1963–1965 and the Foundation for Child Development's National Survey of Children in 1976). The present set of 28 items were selected for inclusion in the 1981 Child Health Supplement to the National Health Interview Survey based on their ability to distinguish children referred for psychological treatment from typical children in a large sample (1,300 children in each group). In addition to their discriminant validity, the items were selected to measure six behavioral syndromes (antisocial, anxious/depressed, hyperactive, stubborn/parental conflicting, social withdrawing/peer conflicts, and immature dependency) suitable for boys and girls age 4 and over.

The 1981 Child Health Supplement data analysis found the internal consistency reliability of the behavior problems index to be .89 (.91 for children 12 and over). The internal consistency reliabilities on the subscales ranged from .54 to .76. Mott and Quinlan (1991) found internal consistency reliability measures in the same range, suggesting the Child Survey from the NLSY79 was comparable. The two-week test-retest reliability of the hyperactivity subscale was .68, suggesting that the retest reliability of the total scale is in the low .90s. Having been employed in prior national surveys, children from the entire range of social, economic, and ethnic backgrounds may be accurately assessed.

References

Achenbach, T.S. "The Child Behavior Profile: Boys Aged 6-11." *Journal of Consulting and Clinical Psychology* 46 (1978): 478-488

Achenbach, T.S. and Edelbrock, C.S. "Behavioral Problems and Competencies Reported by Parents of Normal and Disturbed Children Aged Four through Sixteen." *Monographs of the Society for Research in Child Development*, Serial No. 188, Vol. 46, Whole No. 1, 82 pages, 1981.

Mott, Frank L. and Quinlan, Steven V. *Children of the NLSY79: 1988 Tabulations and Summary Discussion*. Columbus, OH: Center for Human Resource Research, March 1991.

National Center for Health Statistics. "Current Estimates from the National Health Interview Survey: United States, 1981." *Vital and Health Statistics*, Series 10, No. 141, DHHS Publ. No. (PHS) 83-1569. Public Health Service. Washington, D.C.: U.S. Government Printing Office, October 1982.

Peterson, J.L. and Zill, Nicholas. "Marital Disruption, Parent-Child Relationships, and Behavioral Problems in Children." *Journal of Marriage and the Family* 48,2 (May 1988).

Rutter, M.; Tizard, J.; and Whitmore, K. *Education, Health, and Behaviour*. London: Longman, 1970.

Schreiner, I. "Analysis of the Reinterview Data from the 1981 Child Health Supplement to the National Health Interview Survey." Statistical Methods Division Memorandum. Washington, D.C.: U.S. Bureau of the Census, February 17, 1983.

The mother and child questionnaire supplements included in Attachment 11 describe the assessment battery more fully. With the exception of the copyrighted testing materials (for example, the standard PPVT-R and PIAT examining materials and scoring sheets), these attachments fully describe the child survey.