Measurement Options for the Assessment of Head Start Quality Enhancements

Final Report
Volume II

September 30, 2005

Shefali Pai-Samant Joanna DeWolfe Sarah Caverly	*Xtria, LLC.*
Kimberly Boller	*Mathematica Policy Research, Inc.*
Sharon McGroder	*Consultant*
Jennifer Zettler Jessica Mills	*Xtria, LLC.*
Christine Ross	*Mathematica Policy Research, Inc.*
Cheryl Clark Mariel Quinones Jamie Gulin	*Xtria, LLC.*

Submitted to:

Office of Planning, Research and Evaluation
Administration for Children and Families
370 L'Enfant Promenade
Seventh Floor West
Washington, DC 20447

Project Officer:
Anne Bergan

Submitted by:

Mathematica Policy Research, Inc.
P.O. Box 2393
Princeton, NJ 08543-2393
Telephone: (609) 799-3535
Facsimile: (609) 799-0005

Project Director:
Christine Ross

Contract No.: 03Y00416401D/SIN 874-3
MPR Reference No.: 6030-370

CONTENTS

Page

TABLES

Table **Page**

MEASUREMENT OPTIONS FOR THE ASSESSMENT OF HEAD START QUALITY ENHANCEMENTS

This volume of the final report for the project, *Design Options for the Assessment of Head Start Quality Enhancements*, provides a compendium of measures that could be used to evaluate the effectiveness of Head Start enhancements. We focus primarily on child outcome measures, although we also present and discuss measures pertaining to intermediate outcomes related both to changes in the program (for example, program management, teacher-child interaction, teacher knowledge and behaviors, and global classroom quality) and to changes in the home, (for example, parenting practices and the emotional and cognitive stimulation available in the home environment), outcomes that some Head Start enhancements could target. We also review measures and variables pertaining to baseline characteristics and conditions that may serve as moderators of program impacts. All of the measures reviewed here have been used in studies of Head Start programs, children, and families.

CHILD OUTCOMES

This section presents and discusses measures of child outcomes that could be considered for inclusion in evaluations of Head Start enhancements. We address strengths and limitations of these child outcome measures and note gaps in measurement that may need to be filled to most effectively evaluate Head Start enhancements. Selection of measures for a specific enhancement study will be guided by the theory of change articulated by the enhancement developer.

In deciding which child outcome measures to review, we first identified recent large-scale studies and program evaluations involving low-income families with preschool-age children. We then selected child outcome measures most germane to aspects of school readiness identified in the Head Start Child Outcomes Framework (see Appendix A) and chose the measures applicable to preschool-age children. We reviewed child outcome measures from the following projects:

- Head Start Family and Child Experiences Survey (FACES)
- Head Start National Reporting System (NRS)
- Head Start Impact Study (HSIS)
- Early Head Start Pre-Kindergarten Follow-Up Study (TPK)
- Descriptive Study of Head Start Health Services (HS Health)
- Preschool Curriculum Evaluation Research Project (PCER)
- Early Reading First Evaluation (ERF)
- Classroom Literacy Interventions and Outcomes Study (CLIO)
- NICHD Study of Early Child Care and Youth Development (NICHD)
- Head Start Quality Research Center Consortium (QRC)[1]

Each measure is profiled in Appendix B with (1) a detailed description of the measure and the child development construct(s) assessed, (2) a listing of which of the above-mentioned large-scale studies included the measure, (3) information relating to the recency and composition of the norming or research sample, (4) psychometric information on demonstrated reliability and validity, (5) information on the qualifications and time needed to administer the measure, and (6) an overall rating of ease of administering the measure. Box 1 describes the approach and definitions we used in developing the domain tables. These profiles are organized and tabled by domain and domain element of the Child Outcomes Framework (COF):

- Language Development (Appendix Table B.1)
- Literacy (Appendix Table B.2)
- Mathematics (Appendix Table B.3)
- Science (Appendix Table B.4)
- Creative Arts (Appendix Table B.5)
- Social and Emotional Development (Appendix Table B.6)
- Approaches to Learning (Appendix Table B.7)
- Physical Health and Development (Appendix Table B.8)

[1] Only a few child outcome measures from the QRCs are included in the tables. The ones included are distinct from those used in the FACES study.

BOX 1. PROCESS FOR COMPILING THE EIGHT DOMAIN TABLES

Studies Included

We reviewed preschool child outcome measures used in ten studies: (1) the Family and Child Experiences Survey (FACES); (2) the Head Start National Reporting System (NRS); (3) the Head Start Impact Study (HSIS); (4) the Early Head Start Pre-Kindergarten follow-up study (TPK); (5) the Descriptive Study of Head Start Health Services (HS Health); (6) Preschool Curriculum Evaluation Research (PCER); (7) Early Reading First (ERF); (8) Even Start Classroom Literacy Interventions and Outcomes Study (CLIO); (9) the NICHD Study of Early Child Care and Youth Development (NICHD); and (10). Head Start Quality Research Center Consortium (QRC)

Measure Inclusion/Exclusion Criteria

Only measures suitable for assessing outcomes of preschool-age children are profiled in Tables B.1 to B.9 (although some of these measures are also suitable for use with older childrens). Measures designed specifically for use with elementary school-age children in these studies were not profiled but are listed in Appendix Table B.11. When a measure was dropped from a multi-year study, it was not included in the domain tables, because dropping the measure indicates (1) the measure did not add anything unique to the explanatory power, (2) the measure exhibited ceiling or floor effects, or (3) the measure presented a training challenge. When available, we reviewed the measures used in all the data collection periods of a given study. We note changes in measures over the course of various data collection periods (specifically for FACES, NRS, and CLIO), including reduction of items, rewording of instructions and items, and the use of different forms of the measure.

Description of the Measure

The description of the measure was obtained from the study and the publisher. The data source (child, parent, teacher, and assessor) for each measure is listed. The number of items in the original measure and the study is also reported. The ages that the measure can be administered to are indicated in terms of the publisher-reported ages and study-sample ages. Language refers to the language used for the measure. If the Spanish measure is a literal translation of an English measure, we list the Spanish measure in the same row as the English measure. If the Spanish measure is a different measure, it is listed in alphabetical order in the domain tables.

Prior Use

This area indicates the study/studies where the measure was used.

Psychometrics (Norming, Reliability & Validity)

Publisher-reported psychometrics are reported when relevant and available. If the measure was adapted, both the publisher and study-specific psychometrics are reported. If a measure was developed for a specific study, only the study-specific psychometrics (if available) are reported.

Box 1 (continued)

Domain and Indicator Coding

Coding of the domains and indicators was based on the description of the measure as provided in the publisher's manual and/or on the publisher's website. If the study made modifications to the measure, the description and coding provided by the study is also included. When a measure was developed for a study or significantly adapted for the study, we coded the measure based on the description provided by the study. If the child assessment booklet was available for a given study (most often FACES, NRS, CLIO), the coding was also based on a review of the items. A single measure could be coded for multiple domains and multiple indicators.

Training and Complexity for Use by Researchers

This information comes from the publisher for non-adapted measures and both publisher and study if adapted or developed for the study. When a measure was used in more than one study, we have restricted the training information to include only the FACES training information although training could potentially be different for various studies. If a measure is a non-FACES measure, we report the training information from the study/studies where it was used. Coding information does not include time spent by study home offices for additional coding or time for deriving standard and weighted scores.

List of Sources (Appendix B.10 provides a complete list of sources by study)

For each study, the list of preschool child outcome measures used in the study was obtained from the study website. When available, study OMB packages, technical and progress reports, training manuals, child assessment booklets, and PowerPoint presentations were reviewed. Another resource was documents containing compilations of measures. In some cases, the study research team was contacted for clarifications about the measures. Publisher's web sites and publisher's manuals were also reviewed.

Appendix Tables B.1 through B.8 include only measures that have a scale or subscale addressing the COF domain of interest. Appendix Table B.9 lists additional child outcome measures classified by COF domains that have items addressing the domain, but not a complete scale or subscale representing domain elements. For example, the Child Observation Record (COR) has single items measuring individual domain elements which, taken together, could measure various domains but cannot provide a valid measure of specific domain elements Most of the measures listed in Table B.9 represent the social and emotional development domain.

Appendix A.2 lists the reviewed measures alphabetically and indicates how we categorized each according to the COF domains and domain elements. Appendix B.10 lists the sources we used to identify the measures and their properties by project. Appendix Table B.11 indicates which measures are also applicable to elementary school-age children.

In deciding which measures should be listed under each COF domain and domain element, we considered both publishers' and study researchers' descriptions of which child development constructs the given instrument measured. Because publishers and researchers did not explicitly use the COF framework or terminology in describing these measures, we also reviewed the content of each measure through the lens of the COF framework. We listed a measure under a COF domain element if an entire subscale or scale related to the given domain element. For the Social and Emotional Development domain, we had to make an exception because most measures do not have an entire scale or subscale representing a single domain element. Because the COF refers to child outcomes at the positive end of the developmental continuum, it was a challenge, and required subjective judgment to decide where to list measures that tapped the problem end of the continuum (for example, problem behavior scales).

Researchers and program developers will have different measurement needs for any particular study of a program enhancement. For some, a measure that taps a broad area, with individual items representing many different domain elements, might be the best choice. For others, a measure that focuses on a particular domain element will be appropriate. The list of measures includes broad and narrowly-focused measures. In addition, our decision rules for classifying measures result in listing some measures under more than one domain and domain element. Despite the repetition, this yields, for each domain element, a comprehensive menu of choices for researchers and program developers who seek outcome measures that relate to outcomes that a particular enhancement targets.

Many similar or slightly revised versions of measures reviewed here appeared in more than one of the studies listed above. This reflects the tendency in designing large-scale studies to avoid reinventing the wheel and, instead, to rely on measures that have been successfully used in previous studies—especially if they represent the field's most recent attempts to improve upon existing measures for use in a large-scale study or with a particular population. For this reason, measures used in FACES are well represented in this paper, both because we reviewed the FACES measures explicitly and because many of the other studies we reviewed also drew upon the FACES work in designing their measurement strategies.

The following section provides an overview of measures available to assess aspects of preschool children's school readiness, organized by domain and domain element of the COF. We note the number of measures available to assess each school readiness construct (domain element) and highlight ones that meet most of the selection criteria. In referencing the number of measures available under each domain element, measures that have been sufficiently modified from their original version (that is, if items have been revised, added, and/or dropped) are listed as separate measures, with the corresponding references to the authors or research team responsible for modifying/creating the measure shown in the table. English and Spanish versions of the same measure (usually a direct translation of the measure) were counted as one measure, whereas English and Spanish versions of measures that tap the same domain element but are not simply direct translations (for example, the PPVT-III and the TVIP) are counted as separate measures. Versions of the same measure that have been used with mothers and with fathers are not counted as separate measures.

Versions of the same measure that have been used with parents and with teachers are counted as separate measures because they provide data about the child's behavior in two different settings (home and Head Start). A stand-alone test that is part of a larger battery (such as the WJ-III Letter-Word Identification test) is counted separately from another test from the same battery, whereas a scale that can be decomposed into subscales (such as the CBCL) is counted as one measure.

Language Development Constructs and Measures

Many measures of preschool children's language development meet most of our selection criteria and would be suitable for use in evaluations of Head Start enhancements. Table B.1 summarizes the measures reviewed in the two domain elements of Language Development: (1) Listening and Understanding and (2) Speaking and Communicating.

Listening and Understanding. "Listening and Understanding" refers to the child's increasing ability to understand and infer meaning from spoken language. It includes the ability to attend to and follow simple directions, the understanding of an increasingly complex and varied vocabulary and, for non-English-speaking children, demonstrated progress in understanding English.

We identified 10 distinct measures relating to preschool-age children's "Listening and Understanding." Measures containing psychometrically sound subscales include the Preschool Language Scale-IV/Auditory Comprehension Subscale (PLS-IV AC) and the Pre-LAS 2000/Oral Language Component. The PLS-IV AC Subscale has not been used with Head Start populations, and the Pre-LAS has been used with Head Start populations only as an English-language screener. Measures that focus solely on measuring children's listening and understanding include the PPVT-III and the Woodcock Johnson-III Tests of Achievement/Oral Comprehension Subscale, currently being used in the national Head Start Impact Study (though, unfortunately, psychometric information from this study is not yet available). Of the measures reviewed, the PPVT-III is one of the most in-depth measures of children's listening and understanding, with strong psychometric properties compared to the other measures listed under this domain element. The PPVT-III has been used with Head Start populations; for FACES, trained paraprofessionals administered and scored a shortened version of the PPVT-III in about 10 minutes. The PPVT-III can also be used with elementary school-age (and even older) children, should any Head Start enhancement evaluations decide to measure children's listening and communicating beyond the Head Start year.

Speaking and Communicating. "Speaking and Communicating" refers to the child's increasing ability to produce sounds, pronounce words clearly, speak in sentences of increasing complexity, and convey information through conversation. It includes the use of an increasingly complex and varied vocabulary and, for non-English-speaking children, demonstrated progress in speaking English.

We reviewed nine distinct measures related to preschool-age children's "Speaking and Communicating." Some measures contain subscales (for example, the Expressive

Communication Subscale of the Preschool Language Scale-3) with psychometric properties indicating the valid and reliable measurement of "Speaking and Communicating." The Expressive One-Word Picture Vocabulary Test-III (EWOPVT-III), an in-depth measure that focuses solely on expressive vocabulary, has strong psychometric properties compared to other measures listed under this domain element, but it takes 15 to 20 minutes to administer and score. A shorter measure, the Picture Naming Individual Growth and Development Indicator (IGDI), assesses and scores a child's vocabulary in under three minutes.

Literacy Constructs and Measures

Many measures of preschool children's literacy and prereading knowledge and skills meet most of our selection criteria and would thus be suitable for use in evaluations of Head Start enhancements. Table B.2 summarizes the measures reviewed in the five domain elements under this domain: (1) Phonological Awareness, (2) Book Knowledge and Appreciation, (3) Print Awareness and Concepts, (4) Early Writing, and (5) Alphabet Knowledge.

Phonological Awareness. "Phonological Awareness" refers to the child's increasing ability to discriminate and identify sounds in spoken language and an understanding of the association between sounds and written words.

Our review identified five measures relating to preschool-age children's phonological awareness. The Phonemic Awareness Subscale of the Test of Language Development: Primary to 3rd Grade Edition (TOLD: P-3) is an in-depth, valid, and reliable measure of children's phonological awareness through third grade, though it is not available in Spanish. The Preschool Comprehensive Test of Phonological and Print Processing (Pre-CTOPPP) contains an "Elision Task" (which assesses the child's phonemic awareness), and a "Blending Task" (which assesses the child's ability to combine word parts). The Elision Task is also available in Spanish, and each task takes eight to 10 minutes to administer and score. The Pre-CTOPPP has good psychometric properties.

Book Knowledge and Appreciation. "Book Knowledge and Appreciation" refers to the child's interest in books and reading-related activities, such as listening to and retelling stories and pretending to read.

Our review identified two measures relating to preschool-age children's book knowledge and appreciation. Perhaps the most in-depth measure of book knowledge and appreciation as conceptualized in the COF is the Story and Print Concepts measure, which is also available in Spanish. The available information from FACES on its psychometric properties suggests less than optimal reliability and mixed evidence of its validity with a Head Start population.

Print Awareness and Concepts. "Print Awareness and Concepts" refers to the child's increasing awareness of print as a form of communication. It includes the recognition of words as a unit of print, an increased ability to associate spoken with written words, and

increased awareness of the mechanics of reading (for example, from left to right, top to bottom).

Our review identified six measures relating to preschool-age children's print awareness and concepts. The Pre-CTOPPP contains a "print awareness task" that contains a few items that tap this domain element. The Conventions Subtest of the Test of Early Reading Ability—3rd edition (TERA-3) is another in-depth assessment. It has good reliability and can be used with elementary school-age children, although it is not available in Spanish. The full scale of the TERA can be used both as a screener and as a child outcome measure. The Letter-Word Identification Test of the Woodcock-Johnson III also has good psychometric properties (though the psychometric properties of its Spanish counterpart, Batería Woodcock Muñoz Pruebas de Aprovechamiento, Identificacion de Letras y Palabras, are unclear). It has been used with diverse populations, can be administered in eight minutes, and can be used with elementary school-age children. The Story and Print Concepts measure is also available in Spanish but has questionable psychometric properties.

Early Writing. "Early Writing" refers to the child's interest and engagement in writing activities. It includes the use of an increasing variety of writing tools and materials, and progress from using scribbles and pictures to convey ideas, to using letterlike symbols and, eventually, to writing one's own name.

Our review identified three measures relating to preschool-age children's early writing. The most in-depth assessment of early writing knowledge and ability we found is the Dictation Test of the Woodcock-Johnson III, which has good psychometric properties and—along with its Spanish counterpart, Batería Woodcock Muñoz Pruebas de Aprovechamiento, Dictacion—can be administered and scored in five minutes. It also can be used with elementary school-age children.

Alphabet Knowledge. "Alphabet Knowledge" refers to the child's increasing knowledge of letters and their uses. It includes the ability to name letters of the alphabet, as well as progress in associating the names of the letters with their corresponding shapes and sounds.

We identified eight measures relating to preschool-age children's alphabet knowledge. The "print awareness task" of the Pre-CTOPPP taps the child's ability to identify letters. The Alphabet Subtest of the TERA—3 has good psychometric properties and takes about 10 minutes to administer and score. Likewise, the Letter-Naming Task (and its Spanish counterpart, Nombrando Las Letras) is a psychometrically sound measure and takes only about five minutes to administer and score. The English and Spanish versions of the Letter-Word Identification Test of the Woodcock-Johnson III measure print awareness and can also be used with children of all ages to measure alphabet knowledge.

Mathematics Constructs and Measures

Our review identified six measures relating to preschool children's mathematics, many of which meet most of our selection criteria and would thus be suitable for use in evaluations of Head Start enhancements. Table B.3 summarizes the measures reviewed in

the three domain elements under this domain: (1) Number and Operations, (2) Geometry and Spatial Sense, and (3) Patterns and Measurement.

Number and Operations. "Number and Operations" refers to the child's interest in, and awareness of, numbers and counting as a way to determine quantity. It includes the ability to count, knowledge of the one-to-one association of numbers to objects when counting, and an increasing ability to count and compare quantities using such terms as "more" and "less."

All six measures tap some aspect of children's knowledge of numbers and operations. Some measures comprise entire scales focusing on the child's ability to count (Color Name and Counting) or the ability to count *objects* (the Counting Block Test). A more in-depth scale—the Applied Problems Test of the Woodcock-Johnson-Revised and the Woodcock-Johnson III—assesses the child's ability to solve age-appropriate math problems, which requires an understanding of counting and simple operations like addition and subtraction. Perhaps the most in-depth scale with broad coverage of this domain element as conceptualized by the COF is the Early Math Skills assessment (and its Spanish counterpart, Conocimiento Básicos de Matemáticas), which taps children's knowledge and abilities to recognize numbers, count, identify and name shapes, and interpret simple graphs.

Geometry and Spatial Sense. "Geometry and Spatial Sense" refers to the child's increasing awareness of size, shape, and position. It includes the ability to recognize and name shapes and match and sort objects, as well as an increasing understanding of words such as up, down, under, over, in front, and behind. While some scales had a few items related to geometry or spatial sense, no measures were identified that addressed these topics as a scale or subscale.

Patterns and Measurement. "Patterns and Measurement" refers to the child's increasing awareness of patterns and the concept of measurement. It includes the ability to identify, duplicate, and extend patterns, as well as to demonstrate progress in using tools (such as rulers) to measure objects. No measures were identified related to this area.

Science Constructs and Measures

This domain encompasses two areas: (1) Scientific Skills and Methods and (2) Scientific Knowledge.

Scientific Skills and Methods. "Scientific Skills and Methods" refers to children's increasing knowledge and skills relating to observing, describing, and making predictions about the world around them. It includes the ability to observe, collect, and record information; compare and contrast objects or phenomena; make and test predictions; and draw conclusions.

Scientific Knowledge. "Scientific Knowledge" refers to a child's growing knowledge of the natural world and living things(for example, their bodies, the environment), as well as growing awareness of time, temperature, and cause and effect relationships.

As indicated in Table B.4, no measures were found that adequately assess either area in this domain.

Creative Arts Constructs and Measures

There are four domain elements under this domain: (1) Music, (2) Art, (3) Movement, and (4) Dramatic Play.

Music. The "Music" domain element is meant to tap the child's interest, enjoyment, and developing skills in music-related activities, such as listening to music, singing, and playing instruments. None of the measures reviewed tap this domain element.

Art. The "Art" domain element is meant to tap the child's interest in, enjoyment of, and developing skills in art-related activities, such as creating drawings, paintings, and other artwork. None of the measures reviewed tap this domain element.

Movement. The "Movement" domain element is meant to tap the child's creative expression through movement to different rhythms, beats, and tempos of music. None of the measures reviewed tap this domain element.

Dramatic Play. The "Dramatic Play" domain element is meant to tap the child's interest in engaging in make-believe, with play becoming increasingly complex. The Howes Peer Play Observation scale provides the possibility for coding children's dramatic play.

Social and Emotional Development Constructs and Measures

Table B.6 describes measures that may be used to evaluate preschool-age children's social and emotional development in five domain elements: (1) Self-Concept, (2) Self-Control, (3) Cooperation, (4) Social Relationships, and (5) Knowledge of Families and Communities. In general, many options exist for measuring aspects of preschool-age children's social and emotional development. In fact, deciding which measure or measures of aspects of children's social and emotional development to include in an evaluation of a Head Start enhancement may be daunting. This is because of the range of measures available that differ in the balance struck between tapping breadth or depth of the construct, the number of items used, psychometric properties of the measure, and whether the positive or only the problem end of the developmental spectrum is assessed. Measures of Social and Emotional Development have typically been measured through parent and teacher reports in studies of head start populations and for older children, child reports, rather than through observational assessments, although we include a few observational assessments on our list.

Self-Concept. "Self-Concept" refers to children's awareness of their specific abilities, characteristics, and preferences. It includes the child's growing confidence and growing capacity for independence.

Our review identified five measures that relate to preschool-age children's self-concept but also cover many other aspects of children's social and emotional development. The most common self-concept construct measured in the reviewed studies was child

confidence. Our review found no psychometrically-sound multi-item scale or subscale that exclusively taps one or more aspects of children's self-concept.

Self-Control. "Self-Control" refers to children's growing ability to express their feelings, needs, and opinions in conflict situations without causing harm to themselves or others. It includes the growing understanding of how their actions affect others and the ability to follow rules.

Our review identified 13 measures that relate to preschool-age children's self-control (also called "self-regulation" and "emotion regulation"). Some measures tap self-control only indirectly. For example, some measures tap the presence of aggressive and/or impulsive behavior (for example, the Problem Behavior Subscale of the SSRS, or the Externalizing Problems Subscale of the CBCL). High scores on these measures may indicate a lack of self-control. However, low scores on these measures cannot necessarily be interpreted as evidence of self-control, because these measures do not cover the positive end of the impulsive/self-control continuum. Other examples of indirect measures of self-control include those that tap prosocial behavior or social relationships more broadly—for example, the peer status ratings from the "Friends or Foes?" measure—and the Howes Peer Play Observation Scale.

Other measures tap self-control more directly, but they vary in how narrowly or broadly this is defined (that is, which behaviors constitute "self-control") and whether the measure contains only a few items or a multi-item scale or subscales. For example, the Self-Control Subscale of the Social Skills Rating System is a valid and reliable multi-item measure of children's self-control at home (the parent report) or in the classroom (teacher report). The Delay of Gratification Task and the Parent-Child Interaction Task require observational ratings of children's emotion regulation during structured and intentionally challenging or frustrating tasks.

Cooperation. "Cooperation" refers to a child's ability to sustain social interactions through helping, sharing, discussion, compromise, and taking turns, without being overly submissive or overly directive.

Our review identified 13 measures that contained items relating to preschool-age children's ability to cooperate. Most of these measures are the same as those listed under "Self-Control" on Table B.6. As with self-control, some of these measures tap cooperation only indirectly, while others tap it more directly, completely, and extensively. Measures of the presence of aggressive and/or impulsive behavior that reflect the absence of self-control (see above) can also be thought of as reflecting the absence of cooperation. As noted above, however, low scores on these measures do not necessarily indicate the presence of cooperation. In addition, as with self-control, some measures listed under cooperation on Table B.6 tap children's prosocial behavior or positive relationships more broadly (for example, the Friendship Interaction Coding Scale and the Howes Peer Play Observations Scale) and, thus, may not be good indicators of cooperation per se.

One measure that taps cooperation in greater depth and breadth is the Cooperation Subscale of the SSRS (SSRS-CS). The SSRS-CS has strong psychometric properties, though

it takes longer to administer (about 20 minutes) than the California Preschool Social Competency Scale (about 10 minutes).

Social Relationships. "Social Relationships" refers to a child's growing interest in and ability to develop friendships and positive relationships with adults. It includes being able to accept guidance and directions from familiar adults (such as teachers) and being able to express empathy and respond sympathetically to peers in need.

Our review identified 16 measures that contained items relating to preschool-age children's social relationships. Much overlap exists in the measures listed on Table B.6 under "Social Relationships" and under "Self-Control" and "Cooperation." This is not surprising, since positive social relationships and interactions require (among other things) self-control and cooperation. Thus, the domain element "Social Relationships" is, by its very nature, a broader construct that includes such things as the child's willingness to talk with and accept guidance and directions from teachers, ability to develop friendships, and ability to express empathy and care for others.

In fact, many measures listed on Table B.6 under "Self-Control" and "Cooperation" are probably better thought of as measures of social relationships more broadly. Many of the measures require observational coding of interactions with peers (Friends or Foes?, Friendship Interaction Coding Scale, Howes Peer Play Interaction Coding Scale), or parents (Parent-Child Interaction Task), or otherwise tap the dyadic nature of the child's social relationships. Some measures contain only a few items on "social competence" (for example, the Friendship Interaction Coding Scale). A few measures contain validated subscales of narrow child behavior constructs conducive to positive social relationships, such as empathy (the Empathy Subscale of the SSRS-Teacher Report), cooperation (see above), and self-control (see above). Still other measures constitute validated scales or subscales of children's social competence more broadly conceptualized (for example, the 10-item social competence subscale of the Social Competence and Behavior Evaluation or the Social Skills Scale of the SSRS).

Perhaps especially when selecting measures of Social and Emotional Development, evaluators must carefully articulate the theory of change underlying a given Head Start enhancement and select the measure or measures that best fit with this theory. Is change expected on a narrow socioemotional outcome, such as empathy or problems in friendships (perhaps because these are key components of the enhancement intervention), or is change expected on social competence more broadly (perhaps as an indirect impact of an enhancement focusing on language development)?

Knowledge of Families and Communities. The COF refers to "Knowledge of Families and Communities" as an increasing understanding of similarities and a respect for differences among people regarding gender, race, culture, language, and special needs. Our review did not identify any measures that tap this domain element of the COF.

(for example, by asking a child how a character in a vignette could solve the problem and accomplish their goal).

Physical Health and Development Constructs and Measures

Table B.8 describes measures that may be used to evaluate preschool-age children's Physical Health and Development: (1) Fine Motor Skills, (2) Gross Motor Skills, and (3) Health Status and Practices.

Fine Motor Skills. "Fine Motor Skills" refers to a child's proficiency in using tools requiring manual dexterity and control (such as scissors, a stapler, a hammer) and in tasks requiring hand-eye coordination (such as building with blocks, putting together puzzles, stringing beads, writing, and drawing).

Our review identified four measures that contain items relating to children's fine motor skills. The WJ-R Dictation Test and its renamed WJ-III counterpart (WJ-III Spelling) contain six items requiring the child to draw lines and copy letters. A more in-depth measure, the McCarthy Draw-a-Design Task, is an assessment designed explicitly to test a child's perceptual-motor performance (fine motor) skills. It requires the child to draw increasingly complex lines and geometric figures. The WJ-R Dictation Test and the Draw-a-Design Task are available in English and Spanish.

Gross Motor Skills. "Gross Motor Skills" refers to a child's proficiency at tasks requiring coordination of large muscle groups. It includes the child's growing ability to run, jump, march, hop, gallop, throw, catch, kick, bounce balls, slide, and swing. Our review identified no measures relating to children's gross motor skills.

Health Status and Practices. "Health Status and Practices" refers to a child's progress in physical growth, as well as whether the child engages in healthy behaviors (such as personal hygiene, good nutrition, and healthy activity levels) and observes health and safety rules (such as wearing safety belts and bike helmets and practicing fire safety).

Our review identified five measures that assessed children's health status and practices. The Descriptive Study of Head Start Health Services reviewed the Head Start Bureau's Child Health Record, which includes information on hospitalizations, illnesses, health problems, growth screenings, dates of physical examinations, immunizations, dental health, and nutrition information. Its parent interview also collected information on 11 health activities conducted at Head Start, 11 health topics discussed at home, as well as 11 health practice changes observed in the child. The NICHD Study of Early Child Care assesses children's height and weight, and it also asked parents about any hospitalizations, diagnosed health conditions, and the severity and impact of any illnesses the child has experienced.

Summary and Discussion

Our review of the key studies listed above yielded many measures that tapped one or more of the eight domains of the Child Outcomes Framework. Some measures tap a wide array of domains and domain elements and may be a good choice in an evaluation seeking a

more global measure of children's school readiness in many areas. Other measures tap a more narrow set of constructs but do so in greater depth. For example, the SSRS contains psychometrically valid subscales measuring a child's self-control, cooperation, and approaches to learning, and the PPVT-III was designed to measure children's receptive vocabulary ("Listening and Understanding" in the COF framework). These narrow, psychometrically sound measures may be a good choice in an evaluation targeting, or otherwise seeking, more fine-grained measures of a more limited set of school readiness outcomes.

A number of measures we reviewed contained items relating to one or more areas of development in the COF, but the items were not meant to tap the domain elements as the COF conceptualized and defined them. For example, the CBCL is a well-established, psychometrically sound measure of behavior problems whose scoring allows for the identification of children at risk of developing clinically significant problems. It was not designed to measure social and emotional development more broadly, though some of its items reflect children's self-control (or lack thereof), cooperation (or lack thereof), and ability to develop and maintain social relationships (or lack thereof). The CBCL would be suitable for use in evaluating a Head Start enhancement focusing explicitly on the reduction of behavior problems, perhaps with the goal of preventing the need for clinical treatment.

Measures in the Social and Emotional Development domain were especially difficult to align with the COF. For example, the best researched and most well-established measures in the social and emotional domain relate to children's behavior problems—measures whose problem focus makes it difficult to align with the positively worded, positively focused constructs on the COF. And while measures of children's positive development in the social and emotional domain exist, these measures are typically developed by individual researchers for their own study's purposes (Zaslow et al. 2004), leading to differences across studies in how various aspects of social and emotional development are conceptualized and operationalized. This dilemma is not new; in his review of the child care quality research, Lamb noted the lack of correspondence in how aspects of child development were sometimes conceptualized and operationalized across studies (Lamb 1998; cited in Zaslow 2004). As researchers and policy makers increasingly ask how children's "positive development" can be enhanced (ChildTrends 2000; Zaff and Hair 2003), and as large-scale studies increasingly include measures of children's positive social and emotional development, the field may converge on and/or develop more standardized measures of important aspects of development in this domain—perhaps with an eye toward constructs identified in the COF and school readiness literature (for example, Kagan et al. 1995) as important social and emotional outcomes for young children to achieve.

While many measures reviewed here are good candidates to consider including in Head Start enhancement evaluations, there are some limitations. Most notably, while some domains of the COF are well covered by the measures reviewed here, other domains or domain elements are covered less well, or not at all. Our review identified no measures tapping the "Geometry and Spatial Sense" and "Patterns and Measurement" domain elements in Mathematics; ; "Scientific Skills and Methods" and "Scientific Knowledge" domain elements in Science; "Music", "Art", and "Movement" domain elements in Creative

Arts; ;"Knowledge of Families and Communities" domain element in Social and Emotional Development; and "Gross Motor Skills" in Physical Health and Development. On the other hand, the Language Development and Literacy domains have many, and perhaps enough, measurement options. This is not surprising; the studies we examined for this paper reflect the current focus on children's prereading skills and knowledge, both in educational interventions and more generally in performance measures and evaluations of Head Start. As a result, measures of children's literacy and language development are well developed. In contrast, there has been less focus in the early childhood education field on young children's math and science knowledge and skills; consequently, measures development in these areas is lagging behind measures development in language and literacy. Adequately addressing these limitations will require examining additional measures from outside the set of studies considered here. Although standards for early childhood mathematics education exist (Clements et al. 2004), the focus has been on using assessment to inform education practice, not for developing norm-referenced tests in broad skill areas. Recently, the National Research Council's workshop planning committee on Mathematical and Scientific Development in Early Childhood held a one-day session to examine the state of the research base on early childhood (3- to 5-year-olds) in mathematics and science education. The planning committee assembled two expert panels to stimulate discussion at the workshop on: (1) Mathematical and Scientific Cognitive Development in Early Childhood; and (2) Going from Knowledge to Practice. This effort will also inform measures development in early math and science.

Moreover, whereas some of the measures reviewed (for example, the PPVT-III) have good psychometric properties for use with Head Start populations, others (for example, the Story and Print Concepts Assessment) have questionable psychometric properties in Head Start samples, and still others (for example, the California Preschool Social Competency Scale) would need to be validated with a Head Start population—(perhaps in a Stage 1 enhancement evaluation) before being included in a Stage 2 or 3 Head Start enhancement evaluation. That many of the measures reviewed here perform well with Head Start populations is not accidental; this review concentrated on measures used in studies of Head Start populations. In fact, it is also not accidental that many of the studies reviewed (and, thus, the measures profiled) originated from the FACES work, which represents the most recent and comprehensive effort to date to modify or develop measures of children's development suitable for use with Head Start populations. This strength may also be a limitation, however, because new studies may sometimes adopt FACES measures to permit comparisons of their findings, when child outcome measures more suitable to the particular study's goals may be more appropriate. This is a reasonable strategy, given the otherwise limited number of measures of child development suitable for use in large-scale evaluations of Head Start populations. However, evaluations of Head Start enhancements may need to consider modifying or developing measures to better meet the needs of an experimental evaluation of particular Head Start enhancements that target a particular child outcome domain.

The following section discusses possible changes that could occur in the classroom or in a child's home environment as a result of a Head Start enhancement. It also suggests ways to measure these classroom- and family-level intermediate outcomes.

Start staff. Program management interventions focused on improving directors' or coordinators' skills fall into this area. Head Start staff—such as teachers, education coordinators, health coordinators, and program directors—have varying levels of early childhood experience, knowledge, and skills. Therefore, staff interviews or self-administered questionnaires would capture information across several domains, among them, beliefs about developmentally appropriate practices, early childhood development knowledge, early childhood instruction knowledge, child assessment knowledge, and classroom or program management skills.

Measures of staff knowledge and skills are limited in comparison to measures of classroom quality. Reliable measures of staff knowledge and skills may not yet exist. Particularly in the case of new quality initiatives, measures of teacher knowledge and behavior may not yet have reliable measures. For example, the evaluations of Early Reading First and the Even Start CLIO evaluation struggled to identify measures of language and literacy aspects of the classroom environment and teacher behavior. Information about teacher knowledge and attitudes, however, can be collected through staff interviews or self-administered questionnaires.

Existing measures of staff knowledge and beliefs include teacher interviews developed for the FACES and PCER studies. Beliefs about developmentally appropriate practices are captured through items that focus on how children should be taught and managed—for instance, children's level of autonomy, teachers' philosophies of incentives/punishment, children's level of classroom involvement, and teacher self-efficacy. In addition, it is important to collect information on classroom or program management skills and any special training in these areas.

The QRC project that focused on using assessment as a program improvement mechanism includes a staff survey called the "Early Childhood Work Environment Survey" (Jorde-Bloom 1996). The FACES 2003 study now uses the support subscale of the Policy and Program Management Inventory (Lambert 2002; Lambert et al. 1999a, 1999b), which is designed to measure teacher satisfaction with the management climate of their Head Start center. Intermediate outcome measures in this area focus on staff communication, feelings that staff have about whether managers are responsive and supportive of their ideas, and generally how satisfied they are with their job and their managers.

Teacher Behavior. Many of the proposed enhancement initiatives are focused on changing what teachers do with children during the Head Start day. Intermediate outcome measures in this area include measures of teacher-child interaction, inventories of the range of classroom activities, and reviews of lesson plans. The measures vary as to the level of information they provide, with teacher-child interactions providing microlevel information about the type and quality of teacher-child interactions, inventories of classroom activities providing more general information about what teachers do with children, and lesson plan reviews providing information about what teachers intend to accomplish with children in a given week or on a given day. Among the three areas, the area of teacher-child interactions contains more measures to choose from than do the other two areas. Evaluators also must

determine whether inventories of classroom activities and reviews of lesson plans are more suitable as fidelity measures or as intermediate outcome measures.

In the area of assessing the type, quality, and frequency of teacher-child interactions, the Observational Record of the Caregiving Environment (ORCE; NICHD Research Network 1996), the Child-Caregiver Observation System (C-COS; Boller and Sprachman 1998; a streamlined adaptation of the ORCE), and the Adult Involvement Scale (Howes and Stewart 1987) provide child-focused data collected during alternating observation and recording periods over the course of an observation lasting between two and four hours.[2] These measures were developed because researchers hypothesized that classroom- or setting-level measures are not sensitive enough to capture variations in the quality of the experience of individual children. By choosing one child or a few children to observe in relation to the teachers, the resulting measure provides a detailed account of the frequency and quality of teacher interaction with focus children. These types of measures assess enhancements targeting change at the teacher-child interaction level.

One challenge related to using intensive teacher-child interactions measures in a large-scale study is that they require a great deal of training and reliability testing resources. The training tapes used as part of the NICHD study have not been available to researchers outside the research network, so any study attempting to use that measure would have to develop training and reliability testing tapes. The C-COS training and reliability testing tapes are available to researchers, but the tapes include examples of settings for children two and three years old. In order to use the C-COS with older children, tapes of preschool interactions would have to be developed. The types of behaviors assessed by the C-COS are narrower than those assessed by the ORCE and are primarily in the area of teacher talk to the focus child and child interactions with materials. The Adult Involvement Scale has videotaped training materials and the authors are developing a revised version with additional training resources. Inter-rater reliability is established using live testing.

The Arnett Caregiver Interaction Scale (CIS, Arnett 1989) is a widely used measure of teacher sensitivity, harshness, punitiveness, and detachment. It is usually conducted by focusing on one teacher and rating aspects of the teacher's interaction with all children. The measure, which can easily be administered as part of a broader quality observation, yields high levels of inter-rater reliability with modest training time (two hours of lecture and item-by-item review followed by a practice observation). One challenge is that it does not produce a great deal of variability in scores (Boller 2003); recently researchers have reported that it is less powerful than other measures in predicting child outcomes.

[2] The developers of the ORCE recommend two 44-minute observation periods.

Two other types of measures, inventories of teacher activities and reviews of lesson plans, could be used to assess intermediate outcomes. Although a few large-scale studies have used these types of measures, their predictive validity is not clearly documented.[3]

Early Childhood Environment Quality. There are a number of potential intermediate outcome measures to choose from that focus on the quality of children's experiences in early childhood settings. Given that some Head Start programs deliver child development services primarily through care provided in family child care settings, we include measures of quality provided in classrooms and in family child care homes. Generally, the measures in this area are comprehensive, capturing both the frequency of activities and type and number of materials available to children, in addition to ratings of how well the teachers conduct the activities and use the materials. We included measures that have been widely used in national studies, such as the ECERS-R (used in FACES and PCER), the Family Day Care Rating Scale, and the Assessment Profile (used in FACES and PCER). Although the ELLCO focuses on the literacy environment, we included it as a global measure of the classroom environment. The High/Scope Program Quality Assessment (2003) was developed to serve as a measure of the quality of care provided in center-based settings and to provide programs using the High/Scope curriculum with a way to assess fidelity. Evaluators using this measure to assess fidelity should choose a different measure to assess intermediate outcomes.

Global measures of quality require extensive training to meet inter-rater reliability standards. The observations last two to four hours; the time is needed to get a good sample of the activities and review materials in the classroom. Most of the measures have established internal consistency reliability, strong concurrent validity with other measures, and modest to strong correlations with child outcomes (termed "predictive validity").

One of the benefits of including a widely used measure like the ECERS-R to assess intermediate outcomes is that it provides comparability to other studies, perhaps most importantly to FACES. At each of the evaluation stages, evaluators can use as a benchmark comparisons of intermediate classroom quality outcomes with national Head Start data on classroom quality.

Parent Knowledge and Skills, Behavior, and Home Environment Quality

Evaluators will need intermediate outcome measures of parent knowledge and skills, parent behavior with their children, and home environment quality when enhancements are focused on parenting. Among these types of enhancements are family literacy initiatives, and curriculum implementation with a parent education or home environment component.

[3] As part of the FACES study, teachers reported the frequency of 18 different activities (such as computer time and naming colors), but findings on how this measure relates to other measures of classroom quality and to child outcomes have not yet been published.

Parent Knowledge and Skills. Evaluating a quality enhancement initiative focused on parent skills and knowledge may involve varying the level of parent training. For instance, Head Start staff could train parents on developing certain skills. First, Head Start staff would receive training on methods of working with parents to enhance family literacy. Next, Head Start education coordinators or teachers would train parents on specific methods of enhancing literacy, such as dialogic reading techniques. This intervention could potentially change the home environment and behavior of parents with their children.

Methods similar to those described above for assessing teacher knowledge and skills would be used to capture information on parent knowledge and skills—parent interviews or self-administered questionnaires on personal behavior. Measures of parent knowledge and skills are limited in comparison to measures of parent behavior and interaction with children. There are few measures of parent knowledge and skills that have been used in large-scale studies and that have known reliability and validity with low-income parents. FACES includes a scale of parenting control and warmth, in addition to questions about use of household rules and disciplinary practices. Other measures of interest would depend on the type of enhancement under consideration. For instance, for a family literacy enhancement, measures would include what parents know about how often children should be read to and what some strategies are for supporting language development at home. Evaluators could also rate parent skills in targeted areas based on live observations in the home or code them from videotapes.

Parent Behavior. Enhancements focused on changing parent behavior may require intermediate outcome measures of the quality of the parent-child relationship, the type and frequency of parenting activities, or how parents and children spend their time. A parent-focused enhancement that targets increasing parent use of technology to support children's letter-sound correspondence and phonemic awareness (for example, using handheld computer games at home), may indirectly affect reading frequency, talking and singing songs together, and the number of library visits. An enhancement focused on supporting children's verbal expression of emotions rather than acting out may affect parent responsiveness to the child, attentiveness to the child's nonverbal signals of distress or negative affect, and the overall quality of interactions among family members. In these areas, evaluators of enhancements have a somewhat smaller set of measures to choose from with proven reliability and validity in large-scale studies than are available for teacher behavior.

Self-report measures of parenting activities are widely used in large-scale studies because they are easy to administer and have been demonstrated to relate to child outcomes. Most studies of school readiness include an assessment of the extent to which parents read on their own and to their children, as well as how often they conduct a range of other types of activities with them (for example, the FACES parent interview includes questions about the activities parents do with their children that was adapted from the National Household Education Survey). Time use diaries and interviews also provide information about the types of activities parents and children do together. The predictive validity of these measures is not well established. Often these types of measures are used descriptively by

conditions (as they exist before random assignment) associated with different patterns of program impacts found across sites (*across*-site moderators). For example, if programs in an evaluation vary in their level of implementation of the enhancement, researchers can examine impacts by level of implementation (fully implemented versus not fully implemented). If the group of programs that were fully implemented have larger impacts than the other group of programs, evaluators can conclude that implementation is associated with impacts, but, because programs were not randomly assigned to a given level of implementation, they cannot conclude that full implementation caused the larger impacts.

Selecting Moderator Constructs and Measures

What variables are likely to operate as moderators of the impacts of Head Start enhancements? That is, in what key subgroups of children, and for Head Start programs with what characteristics, might we anticipate differential impacts? Like constructs reflecting child outcomes and intermediate outcomes, constructs conceptualized to be moderators of a Head Start enhancement's impacts on children should be grounded in a logic model specifying how and why impacts might be conditional on child, parent, program, or site characteristics.

With respect to subgroups of children that may be differentially affected by the enhancement, a compensatory perspective suggests that children at greatest risk may benefit most from a Head Start enhancement. On the other hand, a risk perspective suggests that children and/or families with many or severe risks may be less able—compared with their lower risk peers in the enhancement group—to mobilize and benefit from the enhancement because of a limited ability to supplement Head Start activities with related experiences or activities in the home. "Risks" can reflect limited:

- Time (for example, children with full-time employed mothers may not benefit as much from a family-focused enhancement as children with part-time employed mothers)
- Social capital (for example, children with few supportive adults may not benefit as much from a family-focused enhancement as children with many supportive adults in their lives; alternatively, the Head Start enhancement may benefit children with fewer supportive adults by compensating for them)
- Human capital (for example, children with less educated parents may benefit more from a classroom-based enhancement than children with more educated parents because it compensates for a lack of cognitive stimulation at home)
- Material resources (for example, children in families with fewer resources for educationally-stimulating materials may not benefit as much as children in families with more resources)
- Psychological resources (for example, children with depressed mothers may not benefit as much as children with mothers who are not depressed)

Appendix Table B.13 lists family, parent, and child characteristics from FACES that may be used to categorize children at various levels of risk. This list includes measures of children's developmental outcomes; evaluation designs that collect data on child outcomes prior to random assignment allow evaluators to assess whether the enhancement was more or less effective for, say, children with developmental delays versus children who do not have a delay. In addition, because program evaluators are often interested in whether program impacts differ for children, parents, or families with various demographic characteristics (for example, parent's age, race/ethnicity, child gender), demographic characteristics not necessarily reflective of risk are also included in this list.

With respect to program characteristics that might moderate enhancement impacts, one can think in terms of conditions that support or diminish the effectiveness of the enhancement. For example, program impacts may be larger in Head Start programs with a greater percentage of classroom staff with degrees in early childhood, with a greater percentage of classroom staff holding favorable views of the enhancement, or with a greater percentage of center directors with management training. Appendix Table B.14 lists a set of program and teacher characteristics from the Program Information Report (PIR), an annual report to the Head Start Bureau that Head Start grantees and delegates must complete each year, and FACES that may serve as moderators of enhancement impacts in cross-site impact analyses.

The impacts of the same Head Start enhancement model may also vary in different sites, or geographic locations. (Examples of such site-level moderators include the percent of the population living below the poverty line.)

Moderators must reflect conditions before random assignment. Therefore, moderators can be measured at baseline or, if they reflect static variables like gender or race, at any wave of data collection. This approach provides some flexibility in terms of respondent burden. If administrative records contain valid and reliable measures of moderators (such as gender, race, and disability status), evaluators can lessen respondent burden by obtaining this information from administrative data rather than from surveys.

Because moderators must be measured before random assignment, variables conceptualized as moderators of an enhancement's impacts should be fairly stable or at least not be affected by the enhancement. Otherwise, the interpretation of its role as a moderator may be unclear. For example, while a Head Start enhancement aimed at increasing parents' literacy-related activities with the child may work especially well among parents who have high educational expectations for their child, such expectations may themselves be affected by the Head Start enhancement. If they are, then are parents' pre-enhancement expectations—the moderating variable—still relevant? While researchers can use such malleable pre-random assignment characteristics to conduct subgroup analyses, interpreting the findings would be difficult.

REFERENCES

Aber, L., M. Rodriguez, W. Michel, and Y. Shoda. "Mother-Toddler Interactive Experiences as Precursors of Preschool Self-Regulation." Poster presented at the biennial meeting of the Society for Research. Indianapolis, IN, April 1995.

Abbott-Shim, M. and A. Sibley. *Assessment Profile for Early Childhood Programs: Research Version.* Georgia: Quality Assist, Inc., 1998.

Achenbach, Thomas M and Leslie A. Rescorla. *Manual for the ASEBA Preschool Forms and Profiles.* Burlington, VT: University of Vermont Department of Psychiatry, 2000.

Administration on Children, Youth and Families. A Descriptive Study of the Head Start Health Component, Volume II: Technical Report. Washington, DC: Department of Health and Human Services, 1996.

Alborough, J. *Dónde Está mi Osito?* (M. Castro, Trans.). Compton, CA: Santillana, 1992.

Alborough, J. *Where's my Teddy?* Cambridge, MA: Candlewick Press, 1992.

Arnett, J. "Caregivers in Day-Care Centers: Does Training Matter?" *Journal of Applied Developmental Psychology,* vol. 10, 1989, pp. 541-552.

Barkley, R.A. "The Assessment of Attention in Children." In *Frames of Reference for the Assessment of Learning Disabilities: New Views on Measurement Issues,* edited by G.R. Lyon. Baltimore, MD: Brookes Publishing Co., 1994, pp. 69-102.

Barkley, R.A., G.J. DuPaul, and M.B. McMurray. "A Comprehensive Evaluation of Attention Deficit Disorder With and Without Hyperactivity as Defined by Research Criteria." *Journal of Learning Disabilities,* vol. 23, 1990, pp. 279-283.

Barkley, R.A., G.M. Grodzinsky, and G.J. DuPaul. "Frontal Lobe Functions in Attention Deficit Disorder With and Without Hyperactivity: A Review and Research Report." *Journal of Abnormal Child Psychology,* vol. 20, 1992, pp. 163-188.

Baron, R.M., and D.A. Kenny. "The Moderator-Mediator Variable Distinction in Social Psychological Research: Conceptual, Strategic and Statistical Considerations." *Journal of Personality and Social Psychology,* vol. 51, 1986, pp. 1173-1182.

Block, J.H., and J. Block. "The Role of Ego-Control and Ego-Resiliency in the Organization of Behavior." In *Minnesota Symposia on Child Psychology,* vol. 13, edited by W.A. Collins. Hillsdale, NJ: Erlbaum, 1980, pp. 39-101.

Boller, Kimberly. "The Child-Caregiver Observation System." Presentation at the Annual Meeting of the Child Care Policy Research Consortium, Washington, DC, April 9, 2003.

Boller, K., S. Sprachman, and the Early Head Start Research Consortium. *The Child-Caregiver Observation System Instructors Manual.* Princeton, NJ: Mathematica Policy Research, 1998.

Brown, M. W. *Goodnight Moon.* New York, NY: Harper Collins, 1947.

Brown, M. W. *Buenas Noches Luna.* (T. M. Lawer, Trans.). New York, NY: Harper Collins, 1947.

Brownell, R. *Expressive One-Word Picture Vocabulary Test-2000 Edition (EOWPVT-III).* Novato, CA: Academic Therapy Publications, 2000.

Burts, D.C., C. H. Hart, R. Charlesworth, and L. Kirk. "A Comparison of Frequencies of Stress Behaviors Observed in Kindergarten Children in Classrooms with Developmentally Appropriate versus Developmentally Inappropriate Practices." *Early Childhood Research Quarterly*, 1990, vol. 5, pp. 407-423.

Caldwell, Bettye M., and Robert H. Bradley. Home Observation for Measurement of the Environment: Administration Manual*, Revised Edition.* Unpublished manuscript. Little Rock: University of Arkansas at Little Rock, 1984.

Campbell, S.B., E.W. Pierce, C.L. March, L.J. Ewing, and E.K. Szumowski. "Hard-to-Manage Preschool Boys: Symptomatic Behavior Across Contexts and Time." *Child Development,* vol. 65, 1994, pp. 836-851.

ChildTrends. "Research Brief: Preventing Problems vs. Promoting the Positive: What Do We Want for Our Children?" Washington, DC: ChildTrends, May 2000.

Clark, K.E., and G.W. Ladd. "Connectedness and Autonomy Support in Parent-Child Relationships: Links to Children's Socioemotional Orientation and Peer Relationships." *Developmental Psychology,* vol. 36, 2000, pp. 485-498.

Clements, Douglas H., J. Sarama, and A.M. Dibiase (eds.). *Engaging Young Children in Mathematics: Standards for Early Childhood Mathematics Education.* Mahwah, NJ: Lawrence Erlbaum Associates, 2004.

Cost, Quality, and Child Outcomes Study Team. Cost, Quality, and Child Outcomes in Child Care Centers: Executive Summary. Denver: University of Colorado at Denver, January 1995.

Dodge, K., G. Petit, C. McClaskey, and M. Brown. "Social Competence in Children." *Monographs of Society for Research in Child Development,* vol. 51, serial no. 213, 1986.

Duncan, S.E., and E.A. De Avila. *PreLAS 2000*. Monterey, CA: CTB/McGraw-Hill, 1998.

Duncan, S.E., and E. De Avila. *PreLAS 2000 Manual Para Examinadores: Español Forma C*. Monterey, CA: CTB/McGraw-Hill, 1998.

Duncan, S., and E. De Avila. *Pre-LAS*. Monterey, CA: CTB McGraw-Hill, 1985.

Dunn, L.M., and L.M. Dunn. *Peabody Picture and Vocabulary Test, Third Edition. Examiner's Manual and Norms Booklet*. Circle Pines, MN: American Guidance Service, 1997.

Dunn, L.M., E.R. Padilla, D.E. Lugo, and L.M. Dunn. *Test de Vocabulario en Imagenes Peabody*. Circle Pines, MN: American Guidance Service, 1986.

Dunn, L.M., E.R. Padilla, D.E. Lugo, and L.M. Dunn. "Examiner's Manual." Test De Vocabulario en Imagenes Peabody: Adaptacion HispanoAmericana. English-Language Edition of the Manual. Circle Pines, MN: American Guidance Service, 1986.

Early Childhood Research Institute on Measuring Growth and Development. *Picture Naming Individual Growth and Development Indicator*. Minneapolis: University of Minnesota, Center for Early Education and Development, 2003.

Early Head Start Research Consortium. Staff Survey. 1997, 1999. Available at [http://www.acf.hhs.gov/programs/opre/ehs/ehs_resrch/index.html#instru]. Accessed April 6, 2005.

Egeland B., and M. Hiester. *Teaching Task Rating Scales*. Minneapolis, MN: University of Minnesota, Institute of Child Development, 1993.

FACES Research Team. "Behavior Problems Scale." In *Head Start FACES: Longitudinal Findings on Program Performance: Third Progress Report*. Washington, DC: U.S. Department of Health and Human Services, Administration on Children, Youth and Families, Head Start Bureau, 2001, pp. 129-133.

FACES Research Team. "Classroom Activities." In *Head Start FACES: Longitudinal Findings on Program Performance: Third Progress Report*. Washington, DC: U.S. Department of Health and Human Services, Administration on Children, Youth and Families, Head Start Bureau, 2001, pp. 129-133.

FACES Research Team. "Color Names and Counting." In *Head Start FACES: Longitudinal Findings on Program Performance: Third Progress Report.* Washington, DC: U.S. Department of Health and Human Services, Administration on Children, Youth and Families, Head Start Bureau, 2001, pp. 129-133.

FACES Research Team. "Cooperative Classroom Behavior." In *Head Start FACES: Longitudinal Findings on Program Performance: Third Progress Report.* Washington, DC: U.S. Department of Health and Human Services, Administration on Children, Youth and Families, Head Start Bureau, 2001, pp. 129-133.

FACES Research Team. "Social Awareness Tasks." In *Head Start FACES: Longitudinal Findings on Program Performance: Third Progress Report.* Washington, DC: U.S. Department of Health and Human Services, Administration on Children, Youth and Families, Head Start Bureau, 2001, pp. 129-133.

FACES Research Team. "Story and Print Concepts." In *Head Start FACES: Longitudinal Findings on Program Performance: Third Progress Report.* Washington, DC: U.S. Department of Health and Human Services, Administration on Children, Youth and Families, Head Start Bureau, 2001, pp. 129-133.

FACES Research Team. "Your Child's Accomplishments." Modified from Mason, J.M., J. Stewart, C. Peterman, and D. Dunning. *Your Child's Accomplishments: Child Literacy Interest Questionnaire,* 1992. In *Head Start FACES: Longitudinal Findings on Program Performance: Third Progress Report.* Washington, DC: U.S. Department of Health and Human Services, Administration on Children, Youth and Families, Head Start Bureau, 2001, pp. 129-133.

FACES Research Team and Head Start Quality Research Consortium. "Your Child's Behavior (7-item)." In *Head Start FACES: Longitudinal Findings on Program Performance: Third Progress Report.* Washington, DC: U.S. Department of Health and Human Services, Administration on Children, Youth and Families, Head Start Bureau, 2001, pp. 129-133.

FACES Research Team and Head Start Quality Research Consortium. "Your Child's Behavior (12-item)." In *Head Start FACES: Longitudinal Findings on Program Performance: Third Progress Report.* Washington, DC: U.S. Department of Health and Human Services, Administration on Children, Youth and Families, Head Start Bureau, 2001, pp. 129-133.

Fantuzzo, J. W., E. Tighe, and S. Childs. "Family Involvement Questionnaire: A Multivariate Assessment of Family Participation in Early Childhood Education. *Journal of Educational Psychology,* vol. 92, 2000, pp. 367–376.

Feagans, L., K. Fendt, and D. Farran. "The Effects of Day Care Intervention on Teachers' Ratings of the Elementary School Discourse Skills in Disadvantaged Children." *International Journal of Behavioral Development,* vol. 18, 1995, pp. 243-261.

Feshbach, L. "Aggression-Conduct Problems, Attention-Deficits Hyperactivity, Play, and Social Cognition in Four-Year-Old Boys." *Dissertation Abstracts International,* vol. 50, 1990, p. 5878. (Also, doctoral dissertation, University of Washington, 1989.)

Jorde Bloom, P. *Early childhood work environment survey*. Evanston, IL: Early Childhood Professional Development Project, National-Louis University, 1993.

Kagan, S.L., E. Moore, and S. Bredekamp (Eds). *Reconsidering Children's Early Development and Learning: Toward Shared Beliefs and Vocabulary*. Washington, DC: National Educational Goals Panel, 1995.

Ladd, G.W. "Social Networks of Popular, Average, and Rejected Children in School Settings." *Merrill-Palmer Quarterly,* vol. 29, 1983, pp. 283-307.

Ladd, G.W., and J.M. Price. "Predicting Children's Social and School Adjustment Following the Transition from Preschool to Kindergarten." *Child Development,* vol. 58, 1987, pp. 1168-1189.

LaFreniere, P.J., and J.E. Dumas. "Social Competence and Behavior Evaluation in Children Ages 3 to 6 Years: The Short Form (SCBE-30)." *Psychological Assessment,* vol. 8, 1996, pp. 369-377.

LaParo, K.M., and R.C. Pianta. *Classroom Assessment Scoring System (CLASS). Pre-K Version, FACES*, University of Virginia, 2004.

Lamb, M.E. "Nonparental Child Care: Context, Quality, Correlates, and Consequences." In *Handbook of Child Psychology, Vol. 4: Child Psychology in Practice,* edited by W. Damon, I.E. Sigel, and K.A. Renninger. New York: Wiley, 1998.

Lambert, R. G., M. Abbott-Shim, and C. Oxford-Wright. *Policy and Program Management Inventory: Administrators Version.* Georgia State University, 1999. Available at http://education.uncc.edu/qrc/admin%20version2.pdf. Accessed March 30, 2005.

Lambert, R. G., M. Abbott-Shim, and C. Oxford-Wright. *Policy and Program Management Inventory: Teachers Version.* Georgia State University, 2001. Available at http://education.uncc.edu/qrc/teacher%20version.pdf. Accessed March 30, 2005.

Levine, S., F.F. Elzey, and M. Lewis. *California Preschool Social Competency Scale.* Palo Alto, CA: Consulting Psychologists Press, 1969.

Lonigan, C.L., R.K. Wagner, J.K. Torgesen, and C.A. Rashotte. *Preschool Comprehensive Test of Phonological & Print Processing*, 2002.

Lutz, M.N., J.F. Fantuzzo, and P. McDermott. *Adjustment Scales for Preschool Intervention.* (In press.)

Mather, N., and R.W. Woodcock. "Examiner's Manual." Woodcock-Johnson III Tests of Achievement. Itasca, IL: Riverside Publishing, 2001.

McCarthy, D. *McCarthy Scales of Children's Abilities*. San Antonio, TX: The Psychological Corporation, 1970, 1972.

McDermott, P.A., L.F. Green, J.M. Francis, and D.H. Stott. *Preschool Learning Behavior Scale*. Philadelphia, PA: Edumetric and Clinical Science, 2000.

Matheny, Adam P., Jr., Theodore D. Wachs, Jennifer L. Ludwig, and Kay Phillips. "Bringing Order Out of Chaos: Psychometric Characteristics of the Confusion, Hubbub, and Order Scale." *Journal of Applied Developmental Psychology*, vol. 16, 1995, pp. 429-444..

Minarik, E. H. *Little Bear's Wish. In E. H. Minarik Little Bear*. New York: NY: HarperCollins Publishers, Inc., 1985.

Minarik, E. H. *Los Deseos de Osito. In E. H. Minarik Little Bear*. New York: NY: HarperCollins Publishers, Inc., 1985.

Mirsky, A.G., B.J. Anthony, C.C. Duncan, M.B. Ahearn, and S.G. Kellam. "Analysis of the Elements of Attention: A Neuropsychological Approach." *Neurological Review,* vol. 2, 1991, pp. 109-145.

Mischel, W. "Processes in Delay of Gratification." In *Advances in Experimental Social Psychology, vol. 7,* edited by L. Berkowitz. San Diego, CA: Academic Press, 1974, pp. 249-292.

Mischel, W., Y. Shoda, and M. Rodriguez. "Delay of Gratification in Children." *Science,* vol. 244, 1989, pp. 933-938.

National Institute of Child Health and Human Development (NICHD) Study of Early Child Care and Youth Development (1992). "Child's Health (Section 2 of the Standardized Parent Interview)." Phase II instrument documentation. Available at [http://secc.rti.org/PhaseIIdata.cfm]. Accessed June 18, 2004.

National Institute of Child Health and Human Development (NICHD) Study of Early Child Care and Youth Development (1992). "Friends or Foes?" Phase II instrument documentation. Available at [http://secc.rti.org/PhaseIIdata.cfm]. Accessed June 18, 2004.

National Institute of Child Health and Human Development (NICHD) Study of Early Child Care and Youth Development (1992). "Friendship Interaction Coding." Phase II instrument documentation. Available at [http://secc.rti.org/PhaseIIdata.cfm]. Accessed June 18, 2004.

National Institute of Child Health and Human Development (NICHD) Study of Early Child Care and Youth Development (1992). "Height and Weight." Phase I Instrument Documentation. Available at [http://secc.rti.org/home.cfm]. Accessed April 1, 2005.

National Institute of Child Health and Human Development (NICHD) Study of Early Child Care and Youth Development. "Observational Record of the Caregiving Environment (ORCE)." 1996.

Woodcock, R.W., and A.F. Muñoz-Sandoval. *Batería Woodcock-Muñoz Pruebas de Aprovechamiento—evisada.* Chicago: Riverside, 1996.

Zaff, John, and Elizabeth Hair. "Positive Development of the Self: Self-Concept, Self-Esteem, and Identity." In M.H. Bornstein, L. Davidson, D.L. M. Keyes, K.A. Moore, & The Center for Child Well-Being (eds.). Well-Being: Positive Development Across the Life Course. New York: Erlbaum, 2003.

Zaslow, Martha, Tamara Halle, Natasha Cabrera, Julia Calkins, Lindsay Pitzer, and Nancy Geyelin Margie. "Child Outcome Measures in the Study of Child Care Quality: Overview and Next Steps." *Journal of Policy Analysis and Management*, under review, 2004.

Zill, N., M. Collins, and J. West. *Approaching Kindergarten: A Look at Preschoolers in the United States.* NCES Statistical Analysis Report 95-280. Washington, DC: U.S. Department of Education, Office of Educational Research and Improvement, 1995.

Zimmerman, I.L., V.G. Steiner, and R.E. Pond. *Preschool Language Scale (PLS-4).* San Antonio, TX: The Psychological Corporation, 2002.

APPENDIX A

THE HEAD START CHILD OUTCOMES FRAMEWORK AND SUMMARY OF MEASURES BY DOMAINS

THE HEAD START CHILD OUTCOMES FRAMEWORK[1]

LANGUAGE DEVELOPMENT

Listening and Understanding

- Demonstrates increasing ability to attend to and understand conversations, stories, songs, and poems
- Shows progress in understanding and following simple and multiple-step directions
- Understands an increasingly complex and varied vocabulary* [2]
- For non-English speaking children, progresses in listening to and understanding English*

Speaking and Communicating

- Develops increasing abilities to understand and use language to communicate information, experiences, ideas, feelings, opinions, needs, questions; and for other varied purposes*
- Progresses in abilities to initiate and respond appropriately in conversation and discussions with peers and adults
- Uses an increasingly complex and varied spoken vocabulary*

---

[1] From *The Head Start Bulletin*, Issue No. 76 (2003). Available at [www.headstartinfo.org/publications/hsbulletin76/hsb76_09.htm]

[2] Asterisks indicate the four domain elements and nine indicators that are legislatively mandated.

- Progresses in clarity of pronunciation and towards speaking in sentences of increasing length and grammatical complexity
- For non-English speaking children, progresses in speaking English*

LITERACY

Phonological Awareness

- Shows increasing ability to discriminate and identify sounds in spoken language
- Shows growing awareness of beginning and ending sounds of words
- Progresses in recognizing matching sounds and rhymes in familiar words, games, songs, stories, and poems
- Shows growing ability to hear and discriminate separate syllables in words
- Associates sounds with written words, such as awareness that different words begin with the same sound*

Book Knowledge and Appreciation*

- Shows growing interest and involvement in listening to and discussing a variety of fiction and nonfiction books and poetry
- Shows growing interest in reading-related activities, such as asking to have a favorite book read; choosing to look at books; drawing pictures based on stories; asking to take books home; going to the library; and engaging in pretend-reading with other children
- Demonstrates progress in abilities to retell and dictate stories from books and experiences, to act out stories in dramatic play, and to predict what will happen next in a story
- Progresses in learning how to handle and care for books; knowing to view one page at a time in sequence from front to back; and understanding that a book has a title, author, and illustrator

Print Awareness and Concepts

- Shows increasing awareness of print in classroom, home, and community settings
- Develops growing understanding of the different functions of forms of print such as signs, letters, newspapers, lists, messages, and menus
- Demonstrates increasing awareness of concepts of print, such as that reading in English moves from top to bottom and from left to right, that speech can be written down, and that print conveys a message

- Shows progress in recognizing the association between spoken and written words by following print as it is read aloud
- Recognizes a word as a unit of print, or awareness that letters are grouped to form words, and that words are separated by spaces*

Early Writing

- Develops understanding that writing is a way of communicating for a variety of purposes
- Begins to represent stories and experiences through pictures, dictation, and in play
- Experiments with a growing variety of writing tools and materials, such as pencils, crayons, and computers
- Progresses from using scribbles, shapes, or pictures to represent ideas, to using letter-like symbols, to copying or writing familiar words such as their own name

Alphabet Knowledge

- Shows progress in associating the names of letters with their shapes and sounds
- Increases in ability to notice the beginning letters in familiar words
- Identifies at least 10 letters of the alphabet, especially those in their own name*
- Knows that letters of the alphabet are a special category of visual graphics that can be individually named*

MATHEMATICS

Number and Operations*

- Demonstrates increasing interest and awareness of numbers and counting as a means for solving problems and determining quantity
- Begins to associate number concepts, vocabulary, quantities, and written numerals in meaningful ways
- Develops increasing ability to count in sequence to 10 and beyond
- Begins to make use of one-to-one correspondence in counting objects and matching groups of objects
- Begins to use language to compare numbers of objects with terms such as more, less, greater than, fewer than, and equal to
- Develops increased abilities to combine, separate, and name "how many" concrete objects

Geometry and Spatial Sense

- Begins to recognize, describe, compare, and name common shapes, their parts, and attributes
- Progresses in ability to put together and take apart shapes
- Begins to be able to determine whether two objects are the same size and shape
- Shows growth in matching, sorting, putting in a series, and regrouping objects according to one or two attributes such as color, shape, or size
- Builds an increasing understanding of directionality, order, and position of objects and of words such as up, down, over, under, top, bottom, inside, outside, in front, and behind.

Patterns and Measurement

- Enhances abilities to recognize, duplicate, and extend simple patterns using a variety of material.
- Shows increasing abilities to match, sort, put in a series, and regroup objects according to one or two attributes such as shape or size
- Begins to make comparisons between several objects based on a single attribute
- Shows progress in using standard and nonstandard measures for length and area of objects

SCIENCE

Scientific Skills and Methods

- Begins to use senses and a variety of tools and simple measuring devices to gather information, investigate materials, and observe processes and relationships
- Develops increased ability to observe and discuss common properties, differences, and comparisons between objects and materials
- Begins to participate in simple investigations to test observations, discuss and draw conclusions, and form generalizations
- Develops growing abilities to collect, describe, and record information through a variety of means including discussion, drawings, maps, and charts
- Begins to describe and discuss predictions, explanations, and generalizations based on past experience

Scientific Knowledge

- Expands knowledge of and abilities to observe, describe, and discuss the natural world, materials, living things, and natural processes
- Expands knowledge of and respect for their bodies and the environment
- Develops growing awareness of ideas and language related to attributes of time and temperature
- Shows increased awareness and beginning understanding of changes in materials and cause-effect relationships

CREATIVE ARTS

Music

- Participates with increasing interest and enjoyment in a variety of music activities including listening, singing, finger plays, games, and performances
- Experiments with a variety of musical instruments

Art

- Gains ability in using different art media and materials in a variety of ways for creative expression and representation
- Progresses in abilities to create drawings, paintings, models, and other art creations that are more detailed, creative, or realistic
- Develops growing abilities to plan, work independently, and demonstrate care and persistence in a variety of art projects
- Begins to understand and share opinions about artistic products and experiences

Movement

- Expresses through movement and dancing what is felt and heard in various tempos and musical styles
- Shows growth in moving in time to different patterns of beat and rhythm in music

Dramatic Play

- Participates in a variety of dramatic play activities that become more extended and complex

- Shows growing creativity and imagination in using materials and in assuming different roles in dramatic play situations

SOCIAL AND EMOTIONAL DEVELOPMENT

Self-Concept

- Begins to develop and express awareness of self in terms of specific abilities, characteristics, and preferences
- Develops growing capacity for independence in a range of activities, routines, and tasks
- Demonstrates growing confidence in a range of abilities and expresses pride in accomplishments

Self-Control

- Shows progress in expressing feelings, needs, and opinions in difficult situations and conflicts without harming themselves, others, or property
- Develops growing understanding of how their actions affect others and begins to accept the consequences of their actions
- Demonstrates increasing capacity to follow rules and routines and use materials purposefully, safely, and respectfully

Cooperation

- Increases abilities to sustain interactions with peers by helping, sharing, and discussion
- Shows increasing abilities to use compromise and discussion in working, playing, and resolving conflicts with peers
- Develops increasing abilities to give and take in interactions, to take turns in games or using materials, and to interact without being overly submissive or directive

Social Relationships

- Demonstrates increasing comfort in talking with and accepting guidance and directions from a range of familiar adults
- Shows progress in developing friendships with peers
- Progresses in responding sympathetically to peers who are in need, upset, hurt, or angry; and in expressing empathy or caring for others

Knowledge of Families and Communities

- Develops ability to identify personal characteristics, including gender and family composition
- Progresses in understanding similarities and respecting differences among people, such as genders, race, special needs, culture, language, and family structures
- Develop growing awareness of jobs and what is required to perform them
- Begins to express and understand concepts and language of geography in the contexts of the classroom, home, and community

APPROACHES TO LEARNING

Initiative and Curiosity

- Chooses to participate in an increasing variety of tasks and activities
- Develops increased ability to make independent choices
- Approaches tasks and activities with increased flexibility, imagination, and inventiveness
- Grows in eagerness to learn about and discuss a growing range of topics, ideas, and tasks

Engagement and Persistence

- Grows in abilities to persist in and complete a variety of tasks, activities, projects, and experiences
- Demonstrates increasing ability to set goals and develop and follow through on plans
- Shows growing capacity to maintain concentration over time on a task, question, set of directions or interactions, despite distractions and interruptions

Reasoning and Problem Solving

- Develops increasing ability to find more than one solution to a question, task, or problem
- Grows in recognizing and solving problems through active exploration, including trial and error, and interactions and discussions with peers and adults
- Develops increasing abilities to classify, compare and contrast objects, events, and experiences

PHYSICAL HEALTH AND DEVELOPMENT

Fine Motor Skills

- Develops growing strength, dexterity, and control needed to use tools such as scissors, paper punch, stapler, and hammer
- Grows in hand-eye coordination in building with blocks, putting together puzzles, reproducing shapes and patterns, stringing beads, and using scissors
- Progresses in abilities to use writing, drawing, and art tools, including pencils, markers, chalk, paint brushes, and various types of technology

Gross Motor Skills

- Shows increasing levels of proficiency, control, and balance in walking, climbing, running, jumping, hopping, skipping, marching, and galloping
- Demonstrates increasing abilities to coordinate movements in throwing, catching, kicking, bouncing balls, and using the slide and swing

Health Status and Practices

- Progresses in physical growth, strength, stamina, and flexibility
- Participates actively in games, outdoor play, and other forms of exercise that enhance physical fitness
- Shows growing independence in hygiene, nutrition, and personal care when eating, dressing, washing hands, brushing teeth, and toileting
- Builds awareness and ability to follow basic health and safety rules such as fire safety, traffic and pedestrian safety, and responding appropriately to potentially harmful objects, substances, and activities

Table A.1. List of Child Outcome Measures by Head Start Child Outcomes Framework Domains and Domain Elements

Measure	Language Development		Literacy				Mathematics			Science		Creative Arts				Social & Emotional Development					Approaches to Learning			Physical Health & Development		
	Listening & Understanding	Speaking & Communicating	Phonological Awareness	Book Knowledge & Appreciation	Print Awareness & Concepts	Early Writing	Alphabet knowledge	Number & Operations	Geometry & Spatial Sense	Patterns & Measurement	Scientific Skills & Methods	Scientific Knowledge	Music	Art	Movement	Dramatic Play	Self-Concept	Self-Control	Cooperation	Social Relationships	Knowledge of Families & Communities	Initiative & Curiosity	Engagement & Persistence	Reasoning & Problem Solving	Fine Motor Skills	Gross Motor Skills	Health Status & Practices
Attribution Bias Questionnaire																			X	X							
Batería Woodcock-Muæoz Pruebas de Aprovechamiento, Identificación de Letras y Palabras					X		X																				
Bateria Woodcock-Muæoz Pruebas de Aprovechamiento-Revisada, Dictado Test						X	X																		X		
Bateria Woodcock-Muæoz Pruebas de Aprovechamiento-Revisada, Problemas Aplicados								X																			
Behavior Assessment System for Children (BASC-2)																		X	X	X			X				
California Preschool Social Competency Scale																			X	X			X				
Caregiver-Teacher Report Form (C-TRF)																	X	X	X	X			X				
Child Behavior Checklist for Ages 1‰-5 (CBCL/1‰-5)-Parent Report																	X	X	X	X			X				
Child Health Record Review, including Hospitalizations, Illnesses, Health Problems, Growth, Dates of Physical Examinations, Dental Screenings, Immunizations, and Nutrition Information																											X
Children s Behavior Questionnaire (CBQ) Mother Report																		X	X	X							
Children s Behavior Questionnaire (CBQ) Caregiver Report																		X		X			X				
Child s Adaptive Language Inventory	X	X																									
Color Names and Counting								X																			
Counting Blocks								X																			
Delay of Gratification																		X									