

Child Outcomes and Classroom Quality in FACES 2009

OPRE Report 2012-37a

SEPTEMBER 2012



DISCLAIMER:

The views expressed in this publication do not necessarily reflect the views or policies of the Office of Planning, Research and Evaluation, the Administration for Children and Families, or the U.S. Department of Health and Human Services.

This report and other reports sponsored by the Office of Planning, Research and Evaluation are available at <http://www.acf.hhs.gov/programs/opre/index.html>.

ACKNOWLEDGMENTS:

The authors would like to express their appreciation to our Project Officer Maria Woolverton and other federal staff at OPRE and the Office of Head Start. We thank the Mathematica team, including Annalee Kelly, Barbara Carlson, Anne Bloomenthal, Jennifer McNulty, Elizabeth Malone, Sara Skidmore, Melissa Dugger, Marcia Comly Rigby, Brian Takei, John Carsley, Erin Slyne, Ann Ponti, Dan O'Connor, Felicia Hurwitz, Kevin Manbodh, Anca Dumitrescu, Timothy Bruursema, Katherine Burnett, Kristina Rall, Serge Lukashanets, Katherine Bencio, Miriam Loewenberg, Cheri Vogel, Amanda Bernhardt, Laura Bernstein, Kimberly Ruffin, Alfreda Holmes, as well as Francene Barbour, Joan Gutierrez, Patti Vinci, and Jenny Smith at the Survey Operations Center and all of the Mathematica field and telephone staff who collected the data. The research brief also benefited from careful editing by Betty Teller. We are also grateful for the contributions of our partners at Juarez and Associates and the Educational Testing Service. Most of all, we offer our gratitude to the staff, families and children of the 60 FACES 2009 programs across the country, who once again opened their doors and shared their time with us.

Child Outcomes and Classroom Quality in FACES 2009

OPRE Report 2012-37a

SEPTEMBER 2012

Submitted to:

Maria Woolverton, *Project Officer*
Office of Planning, Research, and Evaluation
Administration for Children and Families
U.S. Department of Health and Human Services

Submitted by:

Emily Moiduddin
Nikki Aikens
Louisa Tarullo
Jerry West
Yange Xue
Mathematica Policy Research, Inc.

Project Director:

Jerry West, Mathematica Policy Research

Contract Number: HHSP23320092900YC
Mathematica Reference Number: 06573.122

This report is in the public domain. Permission to reproduce is not necessary.

Suggested citation:

Moiduddin, E., Aikens, N., Tarullo, L., West, J., Xue, Y. (2012). Child Outcomes and Classroom Quality in FACES 2009. OPRE Report 2012-37a. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.



EXECUTIVE SUMMARY

This report provides a portrait of children who entered Head Start for the first time in fall 2009 and completed a year in the program in spring 2010. It also describes their family backgrounds and the classrooms and programs that serve them. Data are drawn from the 2009 cohort of the Head Start Family and Child Experiences Survey (FACES), a periodic, longitudinal study of program performance. Successive samples of Head Start children, their families, classrooms, and programs provide descriptive information at the national level on the population served; staff qualifications, credentials, and opinions; Head Start classroom practices and quality; and child and family outcomes. Previous FACES cohorts were initiated in 1997, 2000, 2003 and 2006.

Methods

Sample. A total of 3,349 newly enrolled 3- and 4-year-old children participated in FACES in fall 2009. A total of 3,022 children were eligible for the spring 2010 followup and 89 percent participated.

Data Collection. In both fall 2009 and spring 2010, children in the study were administered a battery of direct child assessments, and their parents and teachers were interviewed. In spring 2010, observations were conducted in 370 Head Start classrooms. We use data from the parent interviews to describe children's backgrounds and home environments, and we use data from the direct child assessments to report on children's cognitive and physical outcomes at the beginning and end of their first year in Head Start. Parent and teacher ratings provide information about children's social skills, approaches to learning, problem behaviors, and academic and nonacademic accomplishments during the Head Start year. We use teacher interview data to describe children's first classroom experiences in Head Start and classroom observation data to describe classroom quality.

Population Estimates. The data used to report on child and family characteristics and child outcomes were weighted to represent the population of children entering Head Start for the first time in fall 2009 who were still enrolled in spring 2010, or to represent the teachers and classrooms supporting them. We

describe differences across groups or changes from fall to spring that are statistically significant.

Head Start Families

Thirty-six percent of children completing their first year of Head Start are Hispanic/Latino, 34 percent are African American, and 22 percent are White. Twenty-six percent of Head Start children live in households where a language other than English is the primary language spoken to them, and are considered dual language learners (DLLs). Spanish is by far the most prevalent non-English language and is the primary language spoken to 24 percent of children at home. Ninety-two percent of DLLs are from Spanish-speaking homes.

Families of many children completing their first year of Head Start face a number of economic risks. Thirty-two percent of children live with a parent who is unemployed or who has less than a high school education, and 63 percent live in a household with total income at or below the federal poverty threshold. Half of Head Start children live with a single mother.

Despite these challenges, Head Start parents actively support their children's learning, health, and well-being. Family members of most children regularly engage with them in activities that support learning. In the spring, 77 percent of families read to the child three times a week or more, and 92 percent or more reported engaging the child in the past week in activities such as teaching letters, words, or numbers (98 percent); playing with toys or games indoors (98 percent); playing a game or sport or exercising together (92 percent); or counting different things (92 percent). On average, families eat dinner together more than five nights per week. Almost all children have health insurance (97 percent), have a regular health care provider (93 percent), and have had a medical and dental check-up in the past year (98 percent and 95 percent, respectively). On the other hand, a minority of children eat fruits and vegetables at least twice a day (42 percent and 37 percent, respectively), and 19 percent of children watch more than two hours of television each day.

Parents reported high levels of satisfaction with their own and their children's experiences in Head Start at the end of their first year in the program; 80 to 90 percent of children's parents are "very satisfied" with

Head Start in terms of the support it provides for children and families in most areas (for example, helping the child to grow and develop and being open to parents' ideas and participation). The two areas where parents are less likely to be very satisfied are identifying and providing services for the family (71 percent) and helping parents become more involved in community groups (65 percent).

Children's Cognitive, Social-Emotional, and Physical Development and Health

Cognitive Development. We used direct child assessments to measure children's cognitive abilities in language, literacy, and mathematics. The language in which children were assessed is based on their home language and their proficiency in English, determined through a brief screening of English-language skills. Here, we summarize findings for children who were assessed in English and for those who were assessed in Spanish.

Child assessments used in FACES can be scored in multiple ways, each of which can be used to address different types of questions about children's skills and development. Our discussion of children's cognitive skills focuses on standard scores, which allow for examinations of progress relative to a group of peers of the same age. These scores have a mean of 100 and a standard deviation of 15 points. Scores above or below the mean indicate that the child's skills are more or less advanced than those of their same-age peers, respectively. With the exception of letter-word knowledge, Head Start children assessed in English score below norms (that is, below the mean of 100) across cognitive areas, including language, literacy, and mathematics, in both the fall and spring of their first program year; however, in all areas these children progress at a rate greater than their same-age peers. Letter-word scores for Head Start children increase from a mean of 96.4 in the fall to 102.2 in the spring, above the national average. In receptive vocabulary, children progress from 87.3 in the fall to 90.7 in the spring, and in expressive vocabulary from 81.6 in the fall to 84.6 in the spring, still a full standard deviation below the mean for same-age peers. In the area of early writing, children progress from a mean of 94.6 in the fall to 97.4 in the spring, and early math scores increase from 90.0 in the fall to 92.4 in the spring.

By spring, children assessed in Spanish score below the mean for a sample of same-age peers across all

developmental areas. Furthermore, children assessed in Spanish make progress toward the mean only in letter-word knowledge, with scores increasing from 81.1 in the fall to 87.3 in the spring. Their scores in English receptive vocabulary are more than two standard deviations below the national mean in the fall and spring (62.2 by the spring), and their scores in Spanish receptive vocabulary are almost one standard deviation below the mean (85.6 in the spring). In expressive vocabulary, the skills of Spanish-speaking children can be compared to those of their English-speaking peers and their bilingual peers. In both cases, scores are well below the mean; by the spring, these children score two standard deviations below English-speaking peers (70.0) and one standard deviation below bilingual peers (87.6). In both fall and spring, these children score about 90 in the area of early writing, and in the spring, they score 84.2 in early math.

Social-Emotional Development. FACES 2009 uses measures from a variety of sources to provide multiple perspectives on children's positive and challenging behaviors that may affect their ability to learn and interact with peers and adults. We report on children's social-emotional outcomes and approaches to learning on criterion-referenced measures using raw scores (standard scores are not available). Raw scores allow for measurement of change or growth in performance over time. They are an indicator of absolute, rather than relative, performance. We also report on children's executive functioning using the percentage of correct responses on a pencil-tapping task.

According to multiple sources, children show growth in their social skills during their first Head Start year. Based on teacher reports, children's positive social skills scores increase from 15 to 17 points (on a scale of 0 to 24) from fall to spring. Teachers also rated children as having fewer problem behaviors by the spring, including hyperactive behaviors, with scores decreasing from 4.7 in the fall to 4.4 in the spring (on a scale of 0 to 36, with lower scores indicating fewer behavior problems). Finally, teachers reported more positive approaches to learning; children's scores increase from 1.6 to 1.9 (on a scale of 0 to 3). Children are also able to control their first impulse and follow directions on a pencil-tapping task more consistently by spring, suggesting an improvement of executive functioning; on this direct assessment

completed by children entering Head Start as 4-year-olds, children responded correctly 43 percent of the time in the fall and 61 percent of the time in the spring.

Physical Development and Health. At the end of the Head Start year, parents reported 81 percent of children to be in excellent or very good physical health, and there are no differences in these reports between program entry and the end of the program year. Fourteen percent of children have an identified disability, the majority of which are reported to be speech or language impairments. In addition, more than one-third of children (37 percent) are overweight or obese at the end of the first program year. Overall, the percentage of children who are overweight or obese does not change from the fall to the spring. However, at the end of the year more 4-year-olds than 3-year-olds, more boys than girls, and more Hispanic/Latino children than those of other racial/ethnic backgrounds are obese.

Head Start Classrooms and Programs

Head Start teachers bring many years of experience to the classroom, with 68 percent having taught in Head Start for five years or more. A large majority of teachers—85 percent—have at least an associate’s (A.A.) degree, and 50 percent have at least a bachelor’s (B.A.) degree. Teachers’ attitudes appear consistent with developmentally appropriate practice in preschool classrooms, with scores averaging 8.0 out of 10 points. Teachers also reported engaging in a variety of language, literacy, and mathematics activities daily or almost daily, such as working on letter naming, writing letters, discussing new words, counting out loud, and working with geometric and counting materials. Teachers’ mental health status could affect their classroom behaviors and interactions with children. Although most teachers did not report symptoms of depression, 11 percent reported symptoms of at least moderate depression.

Head Start classrooms fall well within professional guidelines and Head Start Program Performance Standards for group size (14.2 children) and child-adult ratio (6.2 children to each adult). On average, classrooms score in the minimal-to-good range for classroom materials and arrangement (4.0 on a scale of 1 to 7) and for the quality of teacher-child interactions (4.7 on a scale of 1 to 7), as measured by the Early Childhood Environment Rating Scale—Revised. On the Classroom Assessment Scoring

System (CLASS), instructional support was rated in the low range (2.3 on a scale of 1 to 7) and emotional support and classroom organization in the middle range (5.3 and 4.7, respectively, on a scale of 1 to 7)—a pattern consistent with other studies.

Correlates of Classroom Quality, Teacher Attitudes, and Children’s Developmental Status

We examined associations among teacher and classroom characteristics and child outcomes. We found few associations between teacher characteristics and either classroom quality or teacher attitudes. Shifting to child outcomes, we found that only two child outcome measures consistently relate to classroom quality: children’s letter-word scores have positive, linear associations with both CLASS Instructional Support and Language Modeling, such that higher observed quality is associated with higher letter-word scores, and teacher-reported social skills have a positive, linear association with CLASS Classroom Organization, such that higher observed quality is associated with higher social skills.

We also asked whether the relationship between quality and outcomes differs in higher quality versus lower quality classrooms—whether there is a “threshold effect.” We identified a threshold in only two associations: receptive vocabulary with CLASS Instructional Support and problem behaviors with CLASS Positive Climate. In both cases, associations of observed quality and children’s scores are marginally stronger in higher quality classrooms. In other words, when classrooms are of higher observed quality, there is a stronger association between children’s scores on receptive vocabulary and the quality of instructional support, and a stronger association between children’s social skills and the quality of the emotional climate.

In general, these findings are consistent with the broader research literature that examines linkages between child outcomes and classroom quality: we identified a handful of modest, linear associations but also found evidence of nonlinear associations for certain quality-outcome pairs, including threshold effects.

INTRODUCTION

This report provides a portrait of children who entered Head Start for the first time in fall 2009 and completed a year in the program in spring 2010, as well as of their family backgrounds and the classrooms and programs that serve them. Data are drawn from the Head Start Family and Child Experiences Survey (FACES), which was first launched in 1997 as a periodic, longitudinal study of program performance. Successive nationally representative samples of Head Start children, their families, classrooms, and programs provide descriptive information on the population served; staff qualifications, credentials, and opinions; Head Start classroom practices and quality measures; and child and family outcomes. FACES includes a battery of child assessments across many developmental domains; interviews with children's parents, teachers, and program managers; and observations of classroom quality. In 2008, the Office of Planning, Research and Evaluation (OPRE) in the DHHS Administration for Children and Families (ACF) funded Mathematica Policy Research and its partners—Educational Testing Service and Juárez and Associates—to design and conduct FACES 2009.

FACES 2009 is the fifth in a series of national cohort studies—previous cohorts were initiated in 1997, 2000, 2003, and 2006. The FACES 2009 child sample was selected to represent 3- and 4-year-old children as they entered their first year of the program, drawing on participants from 60 selected programs from across the country.¹

CONCEPTUAL MODEL AND FRAMEWORK

The conceptual framework for FACES 2009 illustrates the complex interrelationships that help shape the developmental trajectories of children in Head Start (Figure 1). The child's place is primary and constitutes the central core of the relationships depicted in the figure; fostering his or her progress toward school readiness, broadly construed, is Head Start's ultimate goal. The family context—health, economic, and educational resources, as well as cultural factors—forms the first ring of influences surrounding the child. Membership in the Head Start community is reflected in the child's classroom and teachers and the wider Head Start program, all of which influence the quality of the early childhood learning experience. Factors affecting the child's development and well-being also include teacher credentials, classroom quality, and

program management. Finally, community, state, and national policy decisions, depicted in the outer ring, also affect the life of a Head Start child. These multidimensional contexts guide all aspects of the FACES study, from the selection of measures to the multilevel analyses needed to fully address program and policy issues in today's Head Start program.

The Head Start experience is designed to promote immediate, short- and long-term goals for children and families. For children, the experience includes preschool education, health screenings and examinations, nutritionally adequate meals, and opportunities to develop social-emotional skills that support school readiness. For parents, the experience involves opportunities to participate in policy and program decisions. The program provides parents with chances to participate in the classroom and strives to encourage their active involvement in the education and development of their children. Head Start seeks to promote adult literacy and further parent education, where needed and appropriate, and to provide opportunities for careers and training in early childhood education. The program also seeks to promote family self-sufficiency through provision of case management, assessment, referral, and crisis intervention services. Head Start acts as an advocate for necessary family-focused social services through interagency coordination and agreements.

Measurement of these child and family outcomes, both during the program years and through followup at the end of kindergarten, allows fuller understanding of Head Start's efforts to prepare children and their parents for the school experience.

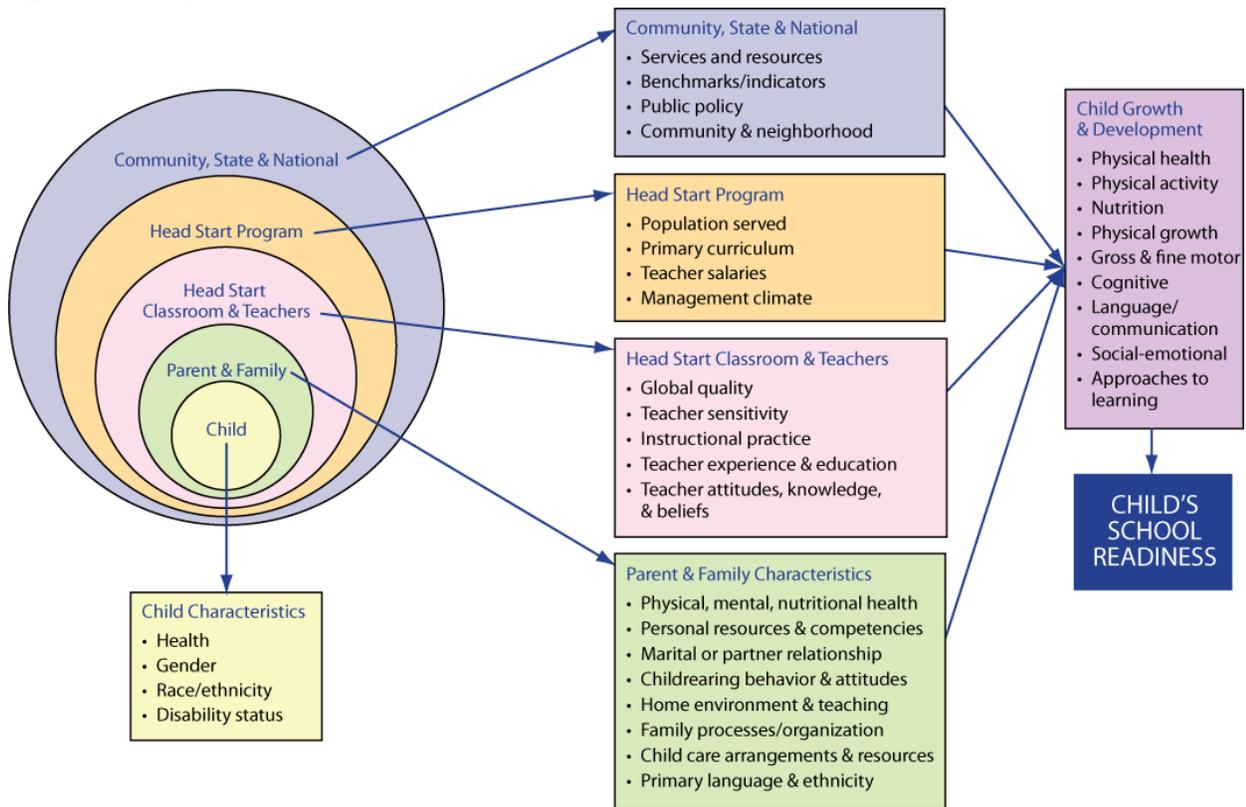
RESEARCH QUESTIONS FOR THIS REPORT

This report presents a portrait of children and families as they complete a year in the program. The report addresses these central research questions:

1. Children and Families Served by Head Start

- What are the demographic characteristics of the population of children and families served by Head Start as they complete a first year in the program? For example, what proportion of children are dual language learners (DLLs)?
- What family routines and learning activities are reported by Head Start parents at the end of the program year? How do these vary from reports at program entry?

Figure 1. Conceptual Model for FACES 2009



- What proportion of children have identified disabilities as they complete a year in the program and what categories of need do they have?
- What is the physical health status of children as they complete a year in the program, including height, weight, and body mass index (BMI)? How do these vary from their entering assessments?
- What are the cognitive and social skills of Head Start children as they complete a year in the program? How do these compare to their entering skills and behaviors?
- How do children's skills, progress, and developmental status vary by age, gender, race/ethnicity, and level of family risk?

2. Head Start Programs, Classrooms, and Services

- What are the characteristics and qualifications of Head Start teachers?
- What is the reported mental health status of Head Start teachers?

- How can we describe teachers' knowledge and beliefs about early childhood development and educational practice?
- What specific curricular and instructional approaches are being used in Head Start classrooms?
- What is the observed quality of Head Start classrooms, and how is it related to teacher and classroom characteristics and children's gains in school readiness skills?

The next section of the report provides background on the study methodology and sample. The following section offers information on the children's characteristics, family demographics, and home environment, including language background, educational activities in the home, family routines, and socioeconomic risk status. It includes information on parent involvement in Head Start and level of satisfaction with their own and their children's Head Start experiences. The next section chronicles children's developmental progress over the Head Start year, considering whether these outcomes vary

by age, gender, race/ethnicity, or risk status. Changes in children's skills and development during the program year reflect a range of influences including maturation, program and family influences, and other influences in children's lives. Presented next are the characteristics of teachers and classrooms, including measures of observed quality. Finally, the last section examines associations among teacher and classroom characteristics and child outcomes.

METHODS

The FACES 2009 sample provides information at the national level about Head Start programs, centers, classrooms, and the children and families they serve. A sample of Head Start programs was selected from the 2007–2008 Head Start Program Information Report (PIR),² with approximately two centers per program and three classrooms per center selected for participation. Within each classroom, an average of eight newly enrolled 3- and 4-year-old children was randomly selected for the study. Sixty programs, 129 centers, 486 classrooms, 439 teachers, and 3,349 children participated in the study in fall 2009.³ Children in the study were administered a battery of direct child assessments, their parents and teachers were interviewed, and interviews were conducted with the directors of the programs and centers in the sample and with education coordinators.

In spring 2010, data were collected again for the group of children who were completing their first year of the Head Start program.⁴ Data were collected over a four-month period (March–June).⁵ Mathematica data collection teams assessed the children at their Head Start centers, interviewed the children's lead teachers, and observed their classrooms. Children's parents were interviewed by phone⁶ and teachers were asked to complete a set of ratings for each sampled child in their classroom using either a web-based or a paper instrument.

A total of 3,022 children were eligible for the spring 2010 followup⁷ and 89 percent participated.⁸ Child assessments were completed for 95 percent of these children and 86 percent of their parents were interviewed. A set of teacher ratings was completed for 96 percent of the children and interviews with 99 percent of children's lead teachers were conducted.⁹ In spring 2010, Mathematica staff also completed observations in 370 Head Start classrooms.¹⁰ Data

from the direct child assessments are used here to report on children's cognitive and physical outcomes at the beginning and end of their first year in Head Start. Parent and teacher ratings provide information about children's social skills, approaches to learning, problem behaviors, and academic and nonacademic accomplishments during the Head Start year. Assessor ratings are another source of information about children's social-emotional outcomes. We use parent interview data to describe children's backgrounds and home environments; teacher interview data to describe children's first Head Start classroom experiences; and classroom observation data to describe Head Start classroom quality.

Direct Child Assessments. The spring battery of direct child assessments, like the fall battery, included a set of standardized preschool assessments designed to measure children's cognitive outcomes (language, literacy, and mathematics) and physical outcomes (height and weight) through an untimed, one-on-one assessment of each child. The actual measures used are described below, where we report on children's cognitive scores at the end of the Head Start year and changes in scores over children's first year in the program.

Except for a few differences, the procedures used to administer the direct child assessments were the same as those used in the fall. It began with a language screening to determine whether children from households where English was not the primary spoken language should be assessed in English, Spanish, or administered an abbreviated battery that included the Peabody Picture Vocabulary Test, Fourth Edition (PPVT-4) (Dunn and Dunn 2006) and the measurement of height and weight.¹¹ However, if a child had been assessed in English in the fall, he or she was assessed in English in the spring regardless of his or her spring score on the language screener.¹² The assessments themselves used the same standard materials that were used in the fall such as stimulus and response pages from the PPVT-4 and Woodcock-Johnson Tests of Achievement, Third Edition (WJ III) (Woodcock et al. 2001) measures. Computer-assisted personal interviewing (CAPI) was used when administering the assessments to facilitate the movement from one measure to the next without the assessor having to calculate stopping or starting points (that is, basals and ceilings). Assessors read the questions and instructions from a computer

screen. The child responded by pointing to the correct answers on the assessment easel or giving a verbal response. Assessors entered the responses into a laptop computer using software that ensured all basal and ceiling rules were followed.

Parent Interviews. FACES 2009, using CAPI, collected information from Head Start parents in a variety of areas including characteristics of households (such as income, number of adult household members, languages spoken in the home) and household members (including age, race/ethnicity, and relationship to study child).¹³ Information was also collected on aspects of the child's home life, children's child care arrangements, and parents' ratings of their children's social skills, problem behaviors, and language, literacy, and math accomplishments. New to the spring interview were questions that asked parents about (1) their involvement and satisfaction with Head Start, (2) access to and use of community services and sources of social support, (3) outdoor spaces near their home where their child could play, and (4) household members' use of alcohol, tobacco, and drugs.

Teacher Interviews and Teacher-Child Reports. In spring 2010, FACES 2009 again conducted interviews with lead teachers about their educational backgrounds, professional experience, and credentials, using CAPI. Teachers reported on scheduled learning activities in their classrooms and estimated the amount of time spent on both teacher-directed and child-selected activities in a typical day, as well as frequency of various languages, literacy development, and math activities. Teachers were asked whether they have a primary curriculum guiding their classroom activities and, if so, they were asked about the number of hours of training they received on the curriculum and who provided this training. They were also asked about program management, including their views on program policies and procedures. In the spring interview, teachers were asked about their interactions with parents and how they go about communicating with those who speak a language other than English. They were also asked whether they have a regular mentor, experiences with that mentor, and involvement in training or technical assistance during this program year.

As in the fall, using a web-based Teacher-Child Report (TCR) form, lead teachers were asked to rate each FACES child in their classroom on a set of items assessing the child's accomplishments, cooperative classroom behavior, behavior problems, and their approaches to learning.¹⁴ Teachers also provided reports of children's health, developmental conditions, and absences during the program year.

Interviewer Ratings. At the end of the fall and spring one-on-one assessments, the assessor completed a set of rating scales evaluating the child's behavior in the assessment situation. Four subscales from the Leiter-R Examiner Rating Scales were used for FACES 2009: (1) attention, (2) organization/impulse control, (3) activity level, and (4) sociability. The 27 items and four subscales make up the cognitive/social scale.

Classroom Observations. In FACES 2009, measures of the classroom environment were obtained from a four-hour observation conducted in the spring. The protocols included an abbreviated form of the Early Childhood Environment Rating Scale-Revised (ECERS-R; Harms et al. 1998) comprised of 21 items and the Classroom Assessment Scoring System (CLASS; Pianta et al. 2008).¹⁵ Classroom observers also collected information on child-adult ratios and group sizes. Observer ratings are used to produce a set of scores that capture the quality of Head Start classrooms as well as indicators of classroom resources and teacher-child interactions. More information on the ECERS-R and CLASS measures is found in the classroom observation section of this report.

Population Estimates. The statistics found in this report are estimates of key characteristics of the population of newly entering Head Start children who were still enrolled in the program in spring 2010, their parents and families, and Head Start teachers and classrooms.¹⁶ The data used to report on child and family characteristics and child outcomes were weighted to represent all children entering Head Start for the first time in fall 2009 who were still enrolled in spring 2010.¹⁷ Teacher data were weighted to represent all teachers serving children who entered Head Start for the first time in fall 2009 and who were still enrolled in their classrooms in spring 2010.

Classroom observation data were weighted to represent all classrooms in spring 2010 that were serving children entering Head Start for the first time in fall 2009. All group differences and associations cited in the text of the report are statistically significant at the $p \leq .05$ level unless otherwise noted.

CHILD AND FAMILY DEMOGRAPHICS, PARENTING, AND THE HOME ENVIRONMENT

Head Start serves a diverse population of low-income children and their families. In recognition of the important role that family plays in a child's development, Head Start has made the family a cornerstone in its framework. The FACES 2009 parent interview collected a wealth of information on the family and household environment of children entering Head Start. This section presents key findings on characteristics of children, household demographics, the home learning environment and parenting practices, family health care and health status, child care, parent involvement and satisfaction with Head Start, and services received by Head Start families for the population of children who entered Head Start in fall 2009 and were still enrolled in the program in spring 2010. Child characteristics and household demographics are drawn from information collected in fall 2009, and all other outcomes are based on information collected in spring 2010. We also describe changes in parent practices and children's activities during the Head Start year, drawing on information collected in both fall 2009 and spring 2010.

Characteristics of Children

In spring 2010, approximately 89 percent of children who were newly enrolled in Head Start programs in fall 2009 were still enrolled and completing their first year of Head Start. The demographic characteristics of the children still enrolled in spring look very much like those who had entered the program in the fall.¹⁸

Sixty-one percent of children completing a year of Head Start were 3 years old when they first entered the program in the fall, and the others were 4 years old or older. They are evenly divided between boys and girls (50 percent in each group).

Thirty-six percent of children completing their first year of Head Start are Hispanic/Latino, and about one third (34 percent) African American.¹⁹ Three-year-olds are more likely to be African American (36 percent) than

are 4-year-olds (31 percent), while 4-year-olds are more likely to be Hispanic/Latino (39 percent) than are 3-year-olds (34 percent).

Thirteen percent of children had participated in Early Head Start. Among children enrolled in Head Start in spring 2010, those who entered the program as 3-year-olds were more likely to have participated in Early Head Start than those who entered at age 4, (15 and 11 percent, respectively).

Family Environment

Twenty-six percent of Head Start children completing their first year of the program live in households where a language other than English is the primary language spoken to them. Spanish is by far the most prevalent non-English language and is the primary language spoken to 24 percent of children at home. Ninety-two percent of DLL children are from Spanish-speaking homes. Among children completing their first year of Head Start, those who entered as 4-year-olds are somewhat more likely to be spoken to primarily in Spanish (27 percent) than are 3-year-olds (23 percent). Among children in households where a non-English language is the primary language spoken to the child, 60 percent are read to only in a language other than English and 20 percent watch television programs only in a non-English language.

Family structure. Overall, the family structure of children completing a year of Head Start looks the same as the group of children who entered the program in the fall. Most children (95 percent) live with at least one of their biological or adoptive parents.²⁰ Half (50 percent) live with a single mother. Seven percent live apart from their mother, and slightly more than half (55 percent) do not live with their father.

Forty-three percent of children live with both biological/adoptive parents. Just over one-quarter (27 percent) of children live in households with their married parents, and 15 percent live in households where their parents are cohabiting. Three-year-olds are more likely than 4-year-olds to live with both parents (45 and 39 percent, respectively).

Parent education, employment, and income. Sixty-eight percent of children have one or more parents with at least a high school diploma or GED living with them. Among children who live with their mother, 64 percent of mothers have at least a high school

diploma or GED. Among children living with their father, 53 percent of fathers have at least a high school diploma or GED. This includes children who live just with their father or with both parents.

Almost half (48 percent) of all children completing a year of Head Start have at least one parent who is working full time and living with them. Thirty-two percent of children are living with a parent who is not employed, including 18 percent who live with at least one parent who is unemployed/looking for work. For children who live with their mother, 27 percent of mothers are working full time, and another 21 percent are working part time. Among children living with their fathers, 72 percent of fathers are employed; most fathers (57 percent) are working full time.

Sixty-three percent of children completing their first year of Head Start live in households where the total household income is at or below the federal poverty threshold. Ninety-three percent of children live in households where total income is less than or equal to 185 percent of the poverty threshold.

Cumulative socioeconomic risk. Coming from a low-income family, single-parent household, or having parents who did not complete high school are identified as risk factors for poor developmental and educational outcomes.²¹ Children with one of these risk factors are more likely to have the others, and research has shown that having more than one risk factor can have negative consequences for children's development and school readiness skills.²² In FACES 2009, a socioeconomic risk index was created as a measure of cumulative family risk. The number of risks is based on three characteristics of children's living circumstances: (1) whether the child resides in a single-parent household, (2) whether the household income is below the federal poverty threshold, and (3) whether the child's mother has less than a high school diploma.²³

Fourteen percent of children completing a year of Head Start have none of these three risk factors, and the majority (74 percent) have one or two family risks. Fifty-one percent of Head Start children have more

than one family risk and 12 percent have all three. Three-year-olds are more likely to have no risks than are 4-year-olds (16 percent and 12 percent, respectively).

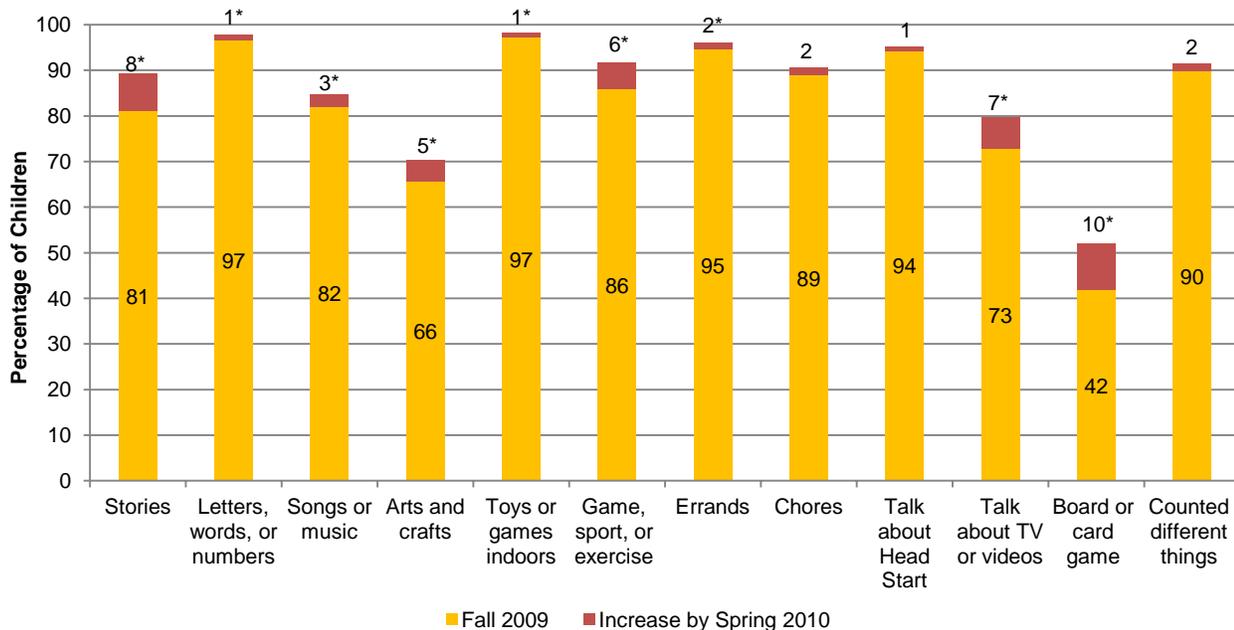
Home Learning Activities and Parenting Practices

Home learning activities. Head Start children participate in a variety of learning activities with their families, both in and outside of the home. One common activity is being read to by a parent or family member. The percentage of children who are read to three or more times a week (77 percent by the spring) did not change significantly between fall and spring. This level of reading did increase for some subgroups however. The percentage of Hispanic/Latino children read to three or more times a week increased from 65 percent in the fall to 70 percent in the spring, and the percentage of children who live in primarily non-English households who are read to three or more times a week rose from 60 to 66 percent.

For most other activities, the percentages of children who had engaged in the activity with a family member in the prior week increased between fall 2009 and spring 2010. As shown in Figure 2, the largest increases (5 to 10 percentage points) were in being told a story; working on arts and crafts; playing games, sports, or exercising; talking about TV programs; and playing a board game or card game with family members. Smaller increases (3 percentage points or less) occurred for teaching a child letters, words, or numbers; playing with toys or games indoors; teaching a child songs or music, and taking a child on errands. None of the activities show a decline between the fall and spring.

The percentages of children who had engaged in activities with family members outside the home also increased between the fall and spring for most types of activities. As shown in Figure 3, the largest increases (4 percentage points or more) were in going to a library, movie, live performance, or museum/historical site; talking about family history or ethnic heritage; or attending an event sponsored by a community group.

Figure 2. Family Members' Activities with Child in Past Week: Fall 2009 to Spring 2010

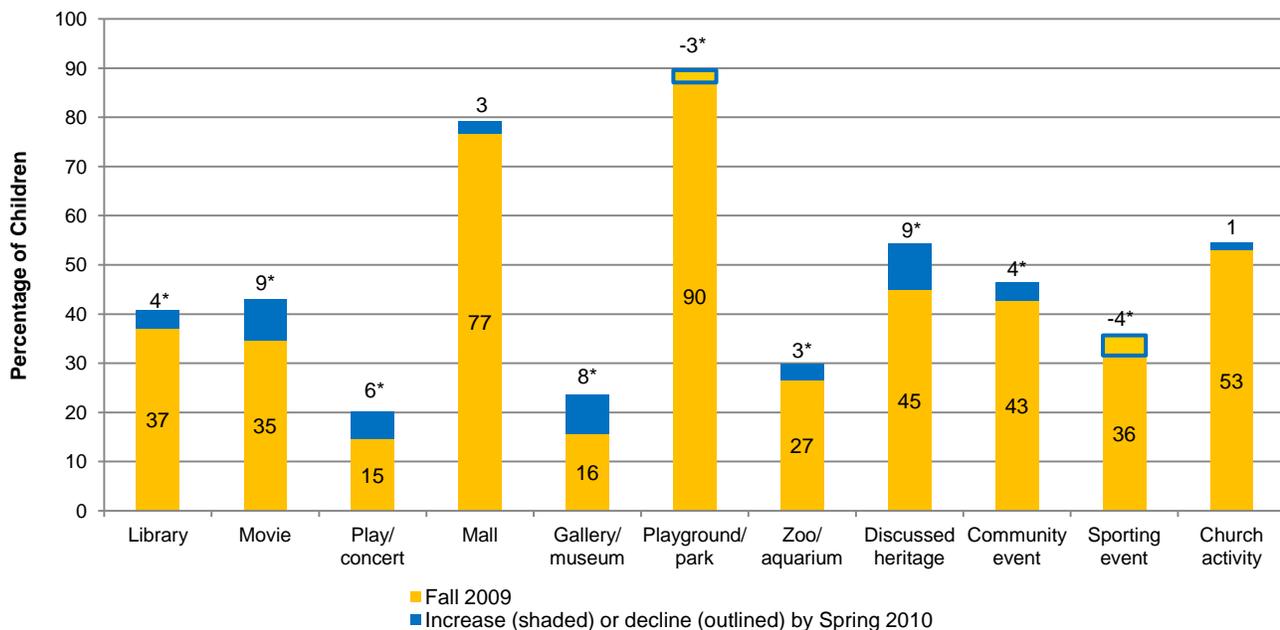


Source: Fall 2009 and Spring 2010 FACES Parent Interview.

Note: Statistics are weighted to represent all children who entered Head Start for the first time in fall 2009 and were still enrolled in spring 2010.

*Asterisk indicates that the difference between the fall and spring scores is statistically significant at the $p \leq .05$ level.

Figure 3. Family Members' Activities with Child in Past Month: Fall 2009 to Spring 2010



Source: Fall 2009 and Spring 2010 FACES Parent Interview.

Note: Statistics are weighted to represent all children who entered Head Start for the first time in fall 2009 and were still enrolled in spring 2010.

*Asterisk indicates that the difference between the fall and spring scores is statistically significant at the $p \leq .05$ level.

Physical activity and nutrition. Parental encouragement of physical activity and good nutritional choices at home can contribute to children's developing healthy habits and combating childhood obesity (Koplan et al. 2005). FACES measured physical activity through questions about frequency of outdoor play. The percentage of children who spend more than two hours playing outside on a typical weekday increased from 28 percent in the fall to 36 percent in the spring. This change was accompanied by a decrease in the percentage spending no time playing outside (from 18 percent to 9 percent) between fall and spring. Of course, it is possible that these changes may reflect seasonal variation in the weather between the two surveys. The pattern of change was similar among 3- and 4-year-old children, although the decline in the percentage of children spending less than one hour or no time outside (-15 percentage points) and the increase in the percentage spending one or more hours outside (15 percentage points) is especially large among 4-year-olds.

The American Academy of Pediatrics' guidelines recommend that children should watch no more than two hours of entertainment media a day (American Academy of Pediatrics 2001). The percentage of Head Start children who watch more than two hours of television on a typical weekday is similar in the fall (18 percent) and spring (19 percent). The percentage of children who watch videos or DVDs on a typical weekday increased between the fall and spring (from 65 percent in the fall to 71 percent in the spring) particularly for 4-year-olds (11 percentage points).

Many children also spend time using computers. A larger percentage of children have access to computers in the spring (67 percent) than in the fall (61 percent), particularly for 3-year-olds (for 4-year-olds, the change was not statistically significant). In addition, the percentage of children who spend time playing computer games on a typical weekday increased from 54 percent in the fall to 62 percent in the spring. The increase is especially large among 4-year-olds (from 56 percent in the fall to 69 percent in the spring).

FACES also asked parents to report on food their child consumed over the past week, concentrating on types of foods that are particularly salient for young children such as milk, soda, salty snacks, sweets, fast food, fruits, and vegetables. Parent reports were

compared to thresholds of consumption that indicate more healthy nutritional choices.²⁴ The percentage of children who ate no fast food in the prior week decreased from 33 percent in the fall to 29 percent in the spring, particularly among 4-year-olds (-6 percentage points). The percentage of children who drank milk at least twice a day also declined (from 66 percent in the fall to 63 percent in the spring). The percentage of children eating fruit at least twice a day in the prior week increased from 38 percent in the fall to 42 percent in the spring.

Child health care. The inclusion of health and wellness services in Heads Start's comprehensive mission reflects the recognition that health care practices can directly influence a child's well-being and development. More than 9 in 10 Head Start children completing their first year (93 percent) have a regular health care provider. Almost all (98 percent) have had a regular medical checkup in the past year and 95 percent have seen a dentist during that time.

In spring 2010, nearly all Head Start children (97 percent) have some type of health insurance. This percentage is higher than the nationwide insurance coverage rate among children of 90 percent in 2009 (Kaiser Commission on Medicaid and the Uninsured 2011). Among children who have insurance, 13 percent are covered only by private insurance and 30 percent are covered only by some type of government insurance.²⁵ Many parents (57 percent) report that their child is covered by both private and government insurance.

Discipline. FACES also asked about specific disciplinary practices and routines in the home. The percentages of children who were spanked in the prior week were similar in the fall and spring (33 percent and 31 percent, respectively). More than twice as many parents used time outs than used spanking in the fall and spring (71 and 72 percent, respectively); the percentage using time outs did not change significantly between fall and spring.

Household routines. Parents' interactions with their children at home, including the rules and routines that parents establish for their children, set the stage for socialization at school. Studies have found time spent eating meals together as a family to be associated with fewer behavior problems (Hofferth and Sandberg 2001) and that dinner table conversation supports

literacy development (Beals and Snow 2006). The average number of days per week that the Head Start family eats dinner together (5.4 days) is the same in the fall and spring, and the percentages that eat dinner together at least 5 nights per week are similar (69 percent in the fall and 72 percent in the spring). The average number of nights in the past week that children brushed their teeth before going to bed increased, but only slightly, between the fall (5.7 nights) and spring (5.8 nights). The number of children brushing their teeth at least five nights a week increased from 76 percent in the fall to 80 percent in the spring.

According to the National Sleep Foundation, children between the ages of 3 and 5 years old need to sleep 11 to 13 hours each day.²⁶ Research has identified associations between the quality and quantity of children's sleep and cognitive and social-emotional outcomes (Gaylor et al. 2010; Hofferth and Sandburg 2001), findings echoed in an analysis of FACES 2006 data (Atkins-Burnett and Aikens 2011). A large majority of Head Start children have a regular bedtime in both the fall (89 percent) and spring (88 percent). The amount of time children spent sleeping each night increased, but only slightly, between the fall and spring, from 10.4 hours to 10.5 hours.

Parent mental health. The mental health status of children's caregivers can affect a child's health, well-being, and development (Meadows et al. 2007; Downey and Coyne 1990). In spring 2010, parents were asked a set of questions from the Center for Epidemiological Studies Depression (CES-D) Scale (short form; Ross et al. 1983). Nearly two-thirds of parents (63 percent) report no symptoms of depression. Six percent of parents report symptoms of severe depression, 21 percent report symptoms of mild depression, and another 9 percent report symptoms of moderate depression.

Time at Head Start and Child Care

The National Household Education Surveys (NHES) 2005 Early Childhood Program Participation Survey found that almost three-quarters (73 percent) of children ages 3–5 who have not yet entered school are in some type of nonparental care arrangement, and that these children spend an average of 28 hours per week in such care (Iruka and Carver 2006). Because many children (37 percent) attend half-day Head Start programs, and even full-day programs

may not match the schedules of some working parents, FACES asked about child care arrangements before and after Head Start.

Children completing their first year of Head Start spend an average of 26 hours per week in their Head Start program. Among those in child care in addition to their Head Start program, children receive an average of 15 hours per week in before- or after-care. Three-year-old children spend slightly more time in other care than 4-year-olds (15.6 hours and 14 hours, respectively). Among all Head Start children, the amount of time they spend in any type of care—including both Head Start and child care—averages 31 hours per week. For those children who are in both Head Start and child care, the total average hours per week in care is 41 hours.

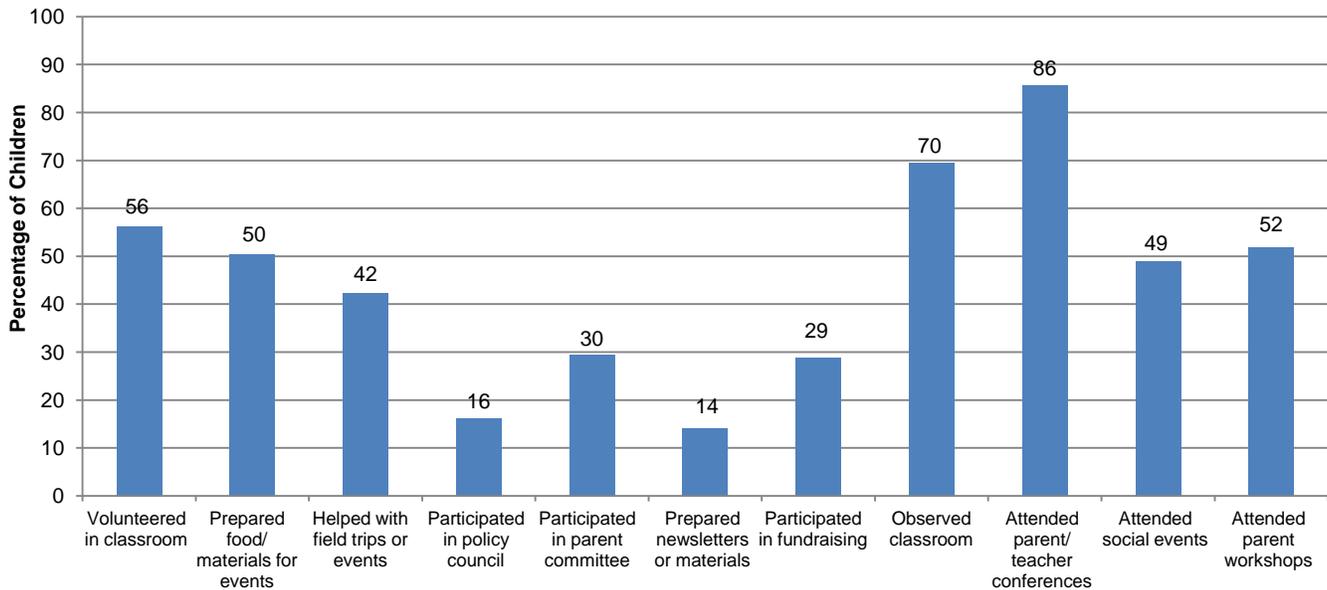
Child Care Outside of Head Start

In spring 2010, 40 percent of children are cared for by someone other than their parents before or after Head Start, and care in addition to Head Start is equally common among 3-year-old and 4-year-old children. Outside of Head Start, children spend their time in a variety of arrangements. Focusing on the care arrangement in which children spend the most time, relative care is the most common type of care children received before or after Head Start (received by 28 percent of all children, and 71 percent of those in any type of care). Only 8 percent of all Head Start children (20 percent of those in any child care) are cared for in a center-based program in addition to Head Start, and 3 percent (7 percent of those in any child care) are cared for in a non-relative home-based setting.

Parent Involvement in Head Start

Parents get involved in their children's Head Start programs in a variety of ways during the first Head Start year. As shown in Figure 4, and consistent with FACES 2006 (Aikens et al. 2010), the most common types of involvement in spring 2010 include attending a parent/teacher conference (86 percent) and observing in the classroom (70 percent). Fifty percent or more of parents also reported volunteering in the classroom (56 percent), attending parent education meetings or workshops (52 percent), and preparing food or materials for special events (50 percent). Activities that occur with the lowest frequency include preparing or distributing newsletters (14 percent) and participating in Head Start policy council (16 percent).

Figure 4. Parent Involvement in Head Start: Spring 2010



Source: Spring 2010 FACES Parent Interview.

Note: Statistics are weighted to represent all children who entered Head Start for the first time in fall 2009 and were still enrolled in spring 2010.

Parents of 3-year-olds and 4-year-olds report getting involved in similar ways, although parents of 4-year-olds are more likely than parents of 3-year-olds to attend parent-teacher conferences (89 percent and 84 percent, respectively) and prepare food or materials for special events (54 percent and 48 percent, respectively).

Parents also report a number of barriers to getting involved in Head Start, the most common of which are work schedules (reported by parents of 54 percent of children), a need for child care (30 percent), and school or training schedules (22 percent). Parents of 10 to 20 percent of children report barriers including a need for transportation, not knowing others at Head Start, health problems, not having enough opportunities to participate, and the opportunities provided being of no interest.

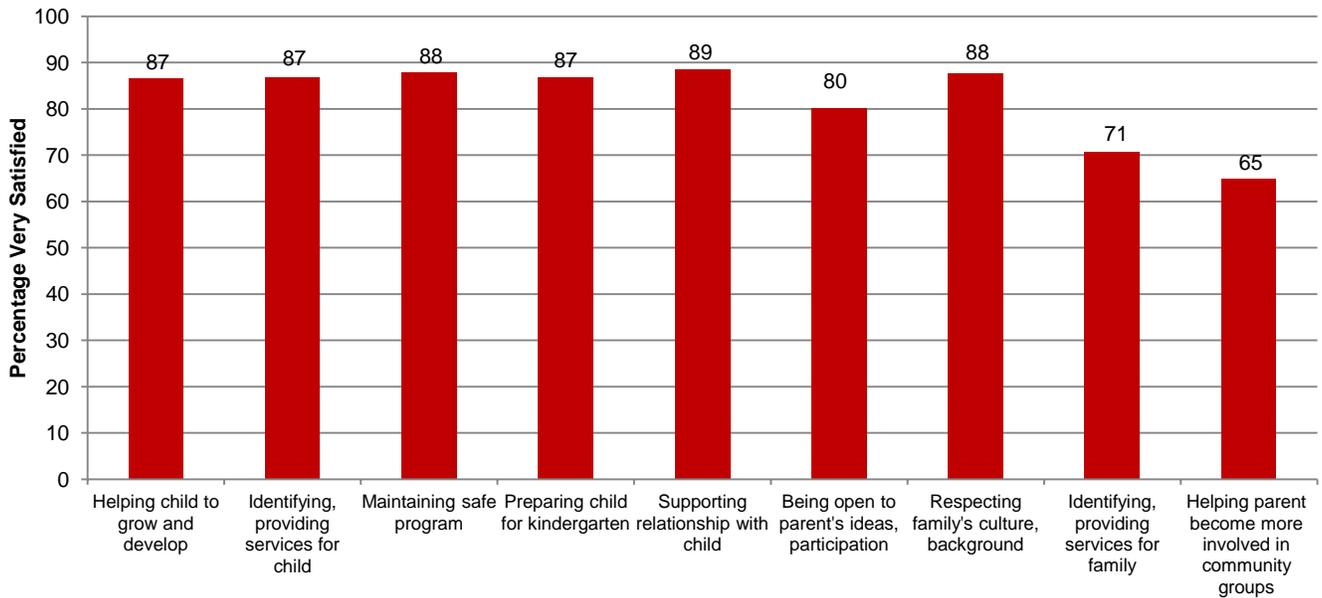
Satisfaction with Head Start

Parents report high levels of satisfaction with Head Start, consistent with findings for earlier FACES cohorts (ACF 2003; Aikens et al. 2010). As shown in Figure 5, parents of between 80 and 90 percent of children report they are "very satisfied" with Head

Start in terms of support for children and families in most areas. The two areas where parents are least likely to be very satisfied are identifying and providing services for the family (71 percent) and helping parents become more involved in community groups (65 percent). There are few differences in satisfaction for parents of 3-year-old and 4-year-old children. Parents of DLL children report high levels of satisfaction with how Head Start is helping the child to develop English language skills (parents of 89 percent of DLL children are very satisfied) and to develop skills in their home language (77 percent).

The parent interview included a list of positive experiences of children and parents with Head Start and asked parents whether each item described their and their child's experiences with the program "never," "sometimes," "often," or "always." As shown in Figure 6, for each item, a majority of parents of both 3-year-old and 4-year-old children reported that this was "always" their own and their child's experience. The frequency of parents responding that these things "always" occur falls below 80 percent for only one item—whether the child gets a lot of individual attention, where the frequency is 62 percent.

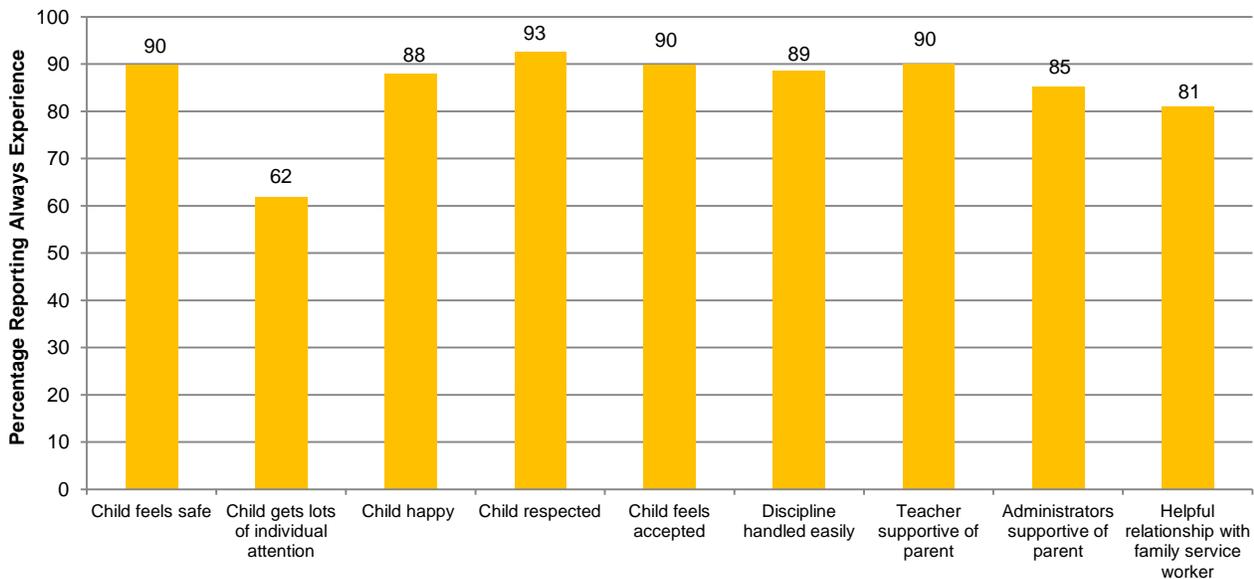
Figure 5. Parent Satisfaction with Head Start: Spring 2010



Source: Spring 2010 FACES Parent Interview.

Note: Statistics are weighted to represent all children who entered Head Start for the first time in fall 2009 and were still enrolled in spring 2010.

Figure 6. Positive Experiences with Head Start: Spring 2010



Source: Spring 2010 FACES Parent Interview.

Note: Statistics are weighted to represent all children who entered Head Start for the first time in fall 2009 and were still enrolled in spring 2010.

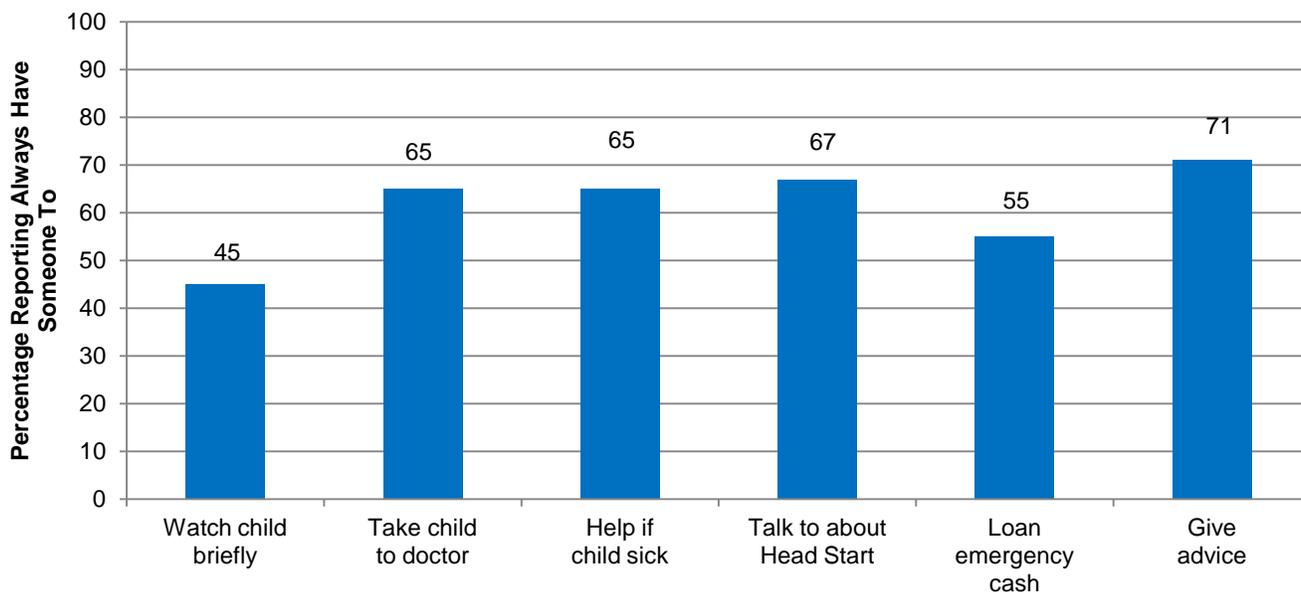
The majority of parents of both 3-year-old and 4-year-old children report that the program does “very well” at providing various kinds of information that can support family engagement in children’s learning; parents of more than 80 percent of children report that Head Start does “very well” at letting them know between conferences how their children are doing; helping them to understand what children of their own child’s age are like; making them aware of chances to volunteer; and providing workshops or advice about how to help children learn at home. Parents of 73 percent of children report that Head Start does “very well” at providing information on community services. The majority of parents (74 percent) who report a

language other than English is spoken at home also report that the program does “very well” at understanding the needs of families who do not speak English.

Social Support

FACES asked parents about the types of social support they receive from various sources, including Head Start. As shown in Figure 7, parents of the majority of children report that they can always find support to meet various needs with one exception—only 45 percent of parents report it is “always true” that they can find someone to watch their child so they can run an errand.

Figure 7. Social Support Parents Receive: Spring 2010



Source: Spring 2010 FACES Parent Interview.

Note: Statistics are weighted to represent all children who entered Head Start for the first time in fall 2009 and were still enrolled in spring 2010.

Parents are most likely to report that they find family members “very helpful” (87 percent) in terms of providing assistance in raising their child over the past month, followed by professionals other than Head Start staff (58 percent). Fifty-five percent of parents report finding Head Start staff very helpful, and 47 percent report finding friends very helpful.

Service Receipt

FACES asked parents about different community or government services that household members may be receiving, including assistance with school or job

training, English as a second language (ESL) classes, and various types of counseling. With the exception of dental or orthodontic or medical care, parents of 7 percent or less of Head Start children report household members obtaining various community services during the 2009–2010 program year. Eighteen percent report receiving dental or orthodontic care and 14 percent report receiving medical care. There are no differences in the percentage of children who are 3 or 4 years old whose parents report they receive various community services.

Among those who receive community services, the majority (between 55 and 79 percent) report that Head Start helped to make them aware of or obtain a service for help finding a job, help going to school or college, ESL classes, transportation to or from work, child care, alcohol or drug treatment or counseling, help dealing with family violence, help or counseling for other family problems, and dental or orthodontic care. Parents are least likely to report that Head Start helped make them aware of or obtain help with housing (22 percent), advice from a lawyer (20 percent), or medical care (20 percent).

In response to a separate question, parents of 12 percent of children completing their first year of Head Start reported that the program had helped them find a regular health care provider for their child. For those who received help, Head Start most often provided information on health care providers (67 percent) and made referrals to them (22 percent), but provided health care directly (4 percent) less often.

Also in response to a separate question, 32 percent of mothers and 18 percent of fathers take programs, courses, classes, or workshops during their child's first year of Head Start. Among those enrolled, 15 percent of mothers and 9 percent of fathers received help from Head Start to take or locate programs courses, classes, or workshops.

SUMMARY

Families of many children completing their first year of Head Start face a number of economic risks. Approximately one-third of children live with a parent who has less than a high school education or an unemployed parent while almost two-thirds live in a household with total income at or below the federal poverty threshold. Half of Head Start children live with a single mother. Despite these challenges, Head Start parents actively support children's learning and well-being. Family members of a majority of children regularly engage with children in activities that support learning, including reading to the child three times a week or more. On average, families eat dinner together more than five nights per week. Almost all children have health insurance, a regular health care provider, and have had a medical and dental check-up in the past year. On the other hand, a minority of children eat fruits and vegetables at least twice a day, and almost 20 percent of children watch more than

two hours of television each day. Parents report high levels of satisfaction with their own and their children's experiences in Head Start at the end of their first year in the program.

CHILD OUTCOMES

In this section we describe the cognitive, social-emotional, and health and physical development of children who entered the Head Start program in fall 2009 and were still enrolled in spring 2010. We report on children's skills on norm- and criterion-referenced measures using raw scores, W scores, item response theory (IRT)-based scores, and standard scores. Standard scores provide information on children's performance relative to same-age peers. These scores have a mean of 100 and a standard deviation of 15. An increase in a standard score signifies that the child is making progress relative to peers. Raw, W, and IRT scores allow for measurement of change or growth in performance over time. However, they are an indicator of absolute, rather than relative, performance.

In the area of children's cognitive development, we first describe the skills of children who were assessed in English in the fall and spring, followed by those assessed in Spanish at both waves.²⁷ We then describe children's social skills, problem behaviors, approaches to learning, and executive functioning during the first program year. Finally, we report on several aspects of children's health and physical development. Throughout this section, we describe the outcomes for all children and then provide descriptive information by important subgroups, including age, gender, race/ethnicity, and family risk status. As noted previously, changes in children's skills and development during the program year reflect a range of influences including maturation, program and family influences, and other influences in children's lives.

Child Cognitive Development

Instruments used. To assess children's skills and knowledge, norm- and criterion-referenced measures of language, writing, and math development were directly administered to the children. Receptive and expressive vocabulary were measured using the Peabody Picture Vocabulary Test, Fourth Edition (PPVT-4) (Dunn and Dunn 2006) and the Expressive

One-Word Picture Vocabulary Test (EOWPVT; EOWPVT-Spanish-Bilingual Edition [SBE]; Brownell 2000).²⁸ In addition, the Test de Vocabulario Imágenes Peabody (TVIP; Dunn et al. 1986) was used to measure children's receptive vocabulary in Spanish.²⁹ The assessment battery also measured children's letter-word knowledge and skills in applied problems and writing, using the Letter-Word Identification, Applied Problems, and Spelling subtests from the Woodcock Johnson III (WJ III; Woodcock et al. 2001) and the Bateria III Woodcock-Muñoz (WM III) Tests of Achievement (Woodcock et al. 2004).³⁰ A supplemental set of math items from the Early Childhood Longitudinal Study-Birth (ECLS-B) and ECLS-Kindergarten (ECLS-K) math assessments was used to assess a broader set of skills than is captured by Applied Problems.³¹ Similarly, to tap the skills of children who had progressed beyond letter knowledge but had not yet acquired sight words, a supplemental set of letter-sound items from the ECLS-B were included.³²

Language screening. In fall 2009, the direct child assessment began with a screening to determine whether children who primarily spoke a language other than English at home should be assessed in English, Spanish, or with a short assessment of vocabulary and height and weight measurements. Two subtests from the Preschool Language Assessment Survey 2000 (*preLAS* 2000; Duncan and DeAvila 1998), Simon Says and Art Show, were used as screening tools. All children were also administered the PPVT-4 to measure their English receptive vocabulary and the EOWPVT or EOWPVT-SBE to measure their expressive vocabulary. Following administration of both measures, children whose home language³³ was Spanish and who made five consecutive errors on Simon Says and Art Show were then routed to the Spanish-language cognitive assessment.³⁴ Similarly, a child who made five consecutive errors on both the Simon Says and Art Show and primarily spoke a language other than English or Spanish was routed out of the cognitive assessment following administration of the vocabulary

measures and was weighed and measured for height. Children who passed the language screener and whose primary home language was a language other than English received the cognitive assessment battery in English. Those from homes in which English was primarily spoken were administered the cognitive assessment battery in English, regardless of their scores on the language screener. In the spring, an adapted version of the screening procedure was used. All children were administered the Simon Says task of the *preLAS* 2000 as a warm-up to the assessment. Following this task (and the receptive and expressive vocabulary measures), those who primarily spoke English at home and those who had passed the language screener in the fall were routed to the English version of the assessment. All other children were administered both Simon Says and Art Show, and, as in the fall, performance on both tasks was used to determine whether these children should be assessed in English, Spanish, or administered a short assessment of vocabulary and height and weight measurements. Table 1 presents the routing procedures for the assessment based on a child's home language and performance on the screener.

Children assessed in English.³⁵ With the exception of letter-word knowledge, children score below national norms on measures of language, literacy, and math development in both fall and spring of their first year of Head Start (Table 2). However, across areas, children make progress toward norms during the year. For example, they gain approximately 3 standard score points in expressive and English receptive vocabulary. They also gain nearly 3 points in early writing, scoring near norms at the end of the program year. They gain almost 6 standard score points in the area of letter-word knowledge during this period and score above the national average in this area by spring (mean = 102.2; Figure 8). In all other areas, despite making progress towards norms during the program year, children remain below norms at the end of the program year. In fact, in expressive vocabulary, they score one standard deviation below norms in spring.

Table 1. FACES 2009 Language Routing Assessment Paths

English	Home Language			
	Spanish		Other	
	English Path	Spanish Path	English Path	Non-English Path
Language Screener (Simon Says and Art Show)	Language Screener (Simon Says and Art Show)	Language Screener (Simon Says and Art Show)	Language Screener (Simon Says and Art Show)	Language Screener (Simon Says and Art Show)
PPVT-4	PPVT-4	PPVT-4	PPVT-4	PPVT-4
EOWPVT	EOWPVT-SBE (conceptually scored)	EOWPVT-SBE (conceptually scored)	EOWPVT	EOWPVT
--	TVIP	TVIP	--	--
WJ III (Spelling, Letter-Word Identification, Applied Problems)	WJ III (Spelling, Letter-Word Identification, Applied Problems)	WM III (Spelling, Letter-Word Identification, Applied Problems)	WJ III (Spelling, Letter-Word Identification, Applied Problems)	--
ECLS-B Letter-Sounds Task ^a	ECLS-B Letter-Sounds Task ^a	--	ECLS-B Letter-Sounds Task ^a	--
ECLS Math	ECLS Math	ECLS Math (Spanish translation available)	ECLS Math	--
Executive Functioning Pencil Tapping Task ^b	Executive Functioning Pencil Tapping Task ^b	Executive Functioning Pencil Tapping Task ^b (Spanish translation available)	Executive Functioning Pencil Tapping Task ^b	--
Height and Weight	Height and Weight	Height and Weight	Height and Weight	Height and Weight

^a This task is administered only to children who meet a certain threshold on the WJ III Letter-Word Identification subtest. Therefore, it is only available for children assessed in English.

^b This task is administered only to children age 4 and older.

Table 2. Mean Standard Scores for FACES Direct Assessment Measures for Children Taking the Assessment in English at Both Waves: Fall 2009, Spring 2010, Fall-Spring Change

Scales (standard scores)	n	Fall 2009	Spring 2010	Fall-Spring Change
		Mean	Mean	Mean
PPVT-4	2,109	87.3	90.7	3.4*
EOWPVT	2,139	81.6	84.6	3.0*
WJ III: Letter-Word Identification	2,007	96.4	102.2	5.8*
WJ III: Spelling	2,041	94.6	97.4	2.8*
WJ III: Applied Problems	1,800	90.0	92.4	2.4*

Source: Fall 2009 and Spring 2010 FACES Direct Child Assessment.

Note: Statistics are weighted to represent all children who entered Head Start for the first time in fall 2009 and were still enrolled in spring 2010.

*Asterisk indicates that the difference between fall and spring scores is statistically significant at the $p \leq .05$ level.

Figure 8. Children’s Mean Woodcock Johnson III Letter-Word Standard Scores, Spring 2010



Source: Spring 2010 FACES Direct Child Assessment.

Note: Statistics are weighted to represent all children who entered Head Start for the first time in fall 2009 and were still enrolled in spring 2010.

Reported scores are for children assessed in English in fall 2009 and spring 2010.

*Within each category (age, gender, race/ethnicity), asterisks indicate subgroups that are statistically significant from other subgroups in the category at the $p \leq .05$ level.

Cross-cohort comparisons (Aikens et al. 2010) show that gains across areas during the first Head Start year are similar across the FACES 2003, 2006, and 2009 cohorts.³⁶ Children made significant progress towards norms in letter-word knowledge in FACES 2003³⁷ (+4.9 points), FACES 2006 (+5.6 points), and FACES 2009 (+5.8 points). Gains towards norms in children’s English receptive vocabulary skills³⁸ and applied problems were also similar across the three cohorts. While the gap between Head Start children’s scores and national norms in early writing did not narrow significantly in FACES 2003 or 2006, children in FACES 2009 did make significant progress towards norms in this area (+2.8 points).³⁹

On the ECLS math items, children also make progress during the program year and can correctly answer more items in the spring. For example, while only 29 percent of children in Head Start are able to identify numbers and shapes at the start of the program year, by the spring slightly more than half are able to do so (51 percent).⁴⁰ For comparison, in the ECLS–B national sample, 63 percent of preschool children demonstrated these skills.⁴¹

Age. With the exception of letter-word knowledge—where both groups score at or above norms in the spring—both 3- and 4-year-olds assessed in English score below norms across areas in the fall and spring (Table 3). Both 3- and 4-year-old children make significant progress towards norms in receptive vocabulary (+2.7 points and +4.6 points, respectively), expressive vocabulary (+3.2 points and +2.4 points, respectively), letter-word knowledge (+6.5 points and +5.1 points, respectively), and early writing (+2.9 points and +2.9 points, respectively). Only 3-year-olds make significant progress towards norms in applied problems (+2.6 points). Notably, among children who took the assessment in English, 3-year-olds perform closer to their same-age peers nationally than 4-year-olds across measures.

On the ECLS math items, both 3- and 4-year-old children make progress during the program year and can correctly answer more items in the spring. For example, while less than 20 percent of 3-year-olds are able to identify numbers and shapes at the start of the program year, by spring 39 percent are able to do so. The percentage increases from 45 percent to 70 percent among 4-year-olds. Across areas, 4-year-olds demonstrate more advanced skills in absolute terms.

Table 3. Mean Standard Scores for FACES Direct Assessment Measures by Age for Children Taking the Assessment in English at Both Waves: Fall 2009, Spring 2010, Fall–Spring Change

Scales (Standard Scores)	n	3-Year-Olds ^a			4-Year-Olds ^a			
		Fall 2009	Spring 2010	Fall-Spring Change	n	Fall 2009	Spring 2010	Fall-Spring Change
PPVT-4	1,245	88.1	90.8	2.7*	798	86.4	91.0	4.6*
EOWPVT	1,269	81.9	85.2	3.2*	804	81.5	83.9	2.4*
WJ III: Letter-Word Identification	1,151	97.9	104.4	6.5*	793	94.2	99.3	5.1*
WJ III: Spelling	1,179	94.7	97.5	2.9*	796	94.5	97.4	2.9*
WJ III: Applied Problems	986	91.0	93.6	2.6*	759	88.9	91.2	2.3

Source: Fall 2009 and Spring 2010 FACES Direct Child Assessment, Fall 2009 FACES Parent Interview.

Note: Statistics are weighted to represent all children who entered Head Start for the first time in fall 2009 and were still enrolled in spring 2010.

Child age is derived from the Fall 2009 FACES Parent Interview.

*Asterisk indicates that the difference between fall and spring scores is statistically significant at the $p < .05$ level.

^aAge as of September 1, 2009.

Gender. With the exception of letter-word knowledge and early writing skills, girls score below national norms across measures in both fall and spring of their first year of Head Start. However, across areas, including receptive vocabulary, expressive vocabulary, letter-word knowledge, early writing, and applied problems, girls make progress toward these norms during the program year. In fact, girls assessed in English gain 6 standard score points in letter-word knowledge during this period and score above the national average in this area by spring (103.6; Figure 8). They also score at the national mean in early writing by the spring (99.3). In contrast, boys make progress towards norms in the areas of receptive vocabulary (+3.6 points), expressive vocabulary (+3.3 points), and letter-word knowledge (+6.0 points) only. Unlike girls, they score below norms across measures at both the beginning and end of the year. The one exception is letter-word knowledge, in which boys score at norms in the spring. Across literacy and math outcomes, girls perform better, relative to same-age peers, than boys in the spring.

On the ECLS math items, both girls and boys can correctly answer more items by the spring. For example, while 28 percent of girls are able to identify numbers and shapes at the start of the program year, by spring 48 percent are able to do so. The

percentage increases from 25 percent to 44 percent among boys.

Race/ethnicity. With the exception of letter-word knowledge, children from all racial/ethnic backgrounds score below national norms across measures of language, literacy, and math development in both the fall and spring of their first year of Head Start. Across groups, in the area of letter-word knowledge children make progress toward norms during the year. In fact, children from all racial/ethnic groups score at or above norms in letter-word knowledge by the end of the program year (Figure 8), with Hispanic/Latino children (+6.4 standard score points) and African American children making the greatest gains (+7.4 standard score points). White and African American children make progress in English receptive vocabulary. African American children make the greatest progress in early writing during the program year (+3.6 standard score points), while only Hispanic/Latino children make progress in applied problems (+3.6 standard score points). Both African American and Hispanic/Latino children make progress towards English speaking peers in their expressive vocabulary.

On the ECLS math items, all children were able to correctly answer more items by the end of the program year. For example, while 34 percent of White

children are able to demonstrate number and shape skills at the start of the program year, by spring 57 percent are able to do so. The percentage increases from 25 percent to 46 percent among African American children, from 28 percent to 54 percent among Hispanic/Latino children, and from 33 percent to 56 percent among “other” race children.

African American children perform better, relative to same-age peers, than White and Hispanic/Latino children in the spring. White and African American children also score higher in ECLS math, English receptive vocabulary, and expressive vocabulary than do Hispanic/Latino children in the spring. Finally, White children score better than African American and Hispanic/Latino children in applied problems at the end of the program year.

Family risk. With the exception of letter-word knowledge, regardless of number of family risks,⁴² children score below national norms across measures of language, literacy, and math development in both the fall and spring of their first year of Head Start. In fact, children with no (+ 6.2 points), one (+ 6.2 points), and two or more risks (+5.4 points) make progress towards norms in the area of letter-word knowledge, and by the end of the program year, children in each of these groups score at or above norms (100.2, 104.1, and 99.5, respectively). Regardless of the number of family risks, children also make progress toward norms during the year in the area of expressive vocabulary. During the program year, children with no, one, and two or more family risks gain 3.4, 2.4, and 3.5 standard score points, respectively. However, all groups remain below norms at the end of the program year in this area. Children with either one or two or more family risks make progress towards norms in English receptive vocabulary (+3.4 and +4.1 points, respectively) and early writing (+3.0 and +2.7 points) during the program year.

On the ECLS math items, all children correctly answer more items by the spring. While 32 percent of children

with no family risks are able to identify numbers and shapes at the start of the program year, by the spring about 54 percent are able to do so. The percentage increases from 28 percent to 50 percent among children with one risk and from 23 percent to 41 percent among children with two or more risks.

In the spring, children with no or one family risk have better letter-word knowledge, applied problems, and early writing skills, relative to same-age peers, than those with two or more family risks. Similar patterns emerged in children’s expressive vocabulary and English receptive vocabulary in the spring, and children with two or more risks also score lower in math than those with fewer family risks.

Children assessed in Spanish. Children who take the assessment in Spanish at the beginning and end of their first Head Start year score below norms across measures of language, literacy, and math development in both the fall and spring (Table 4). Children assessed in Spanish at both time points make progress towards norms only in the area of letter-word knowledge during the year. In fact, children gain about 6 standard score points in this area during the year, or about one-third of a standard deviation. However, despite these gains and unlike children assessed in English, these children still score below norms in letter-word knowledge in the spring.⁴³

Children’s language development. As previously described, FACES 2009 assesses the expressive vocabulary and English receptive vocabulary of all children regardless of home language and screener performance, providing an opportunity to understand the language development of all children. For children from Spanish-speaking households, FACES assesses their receptive vocabulary in English and Spanish, as well as their expressive vocabulary across both languages.⁴⁴ The former provides an opportunity to understand the language development of this group of children both in English and in their home language, while the latter depicts development across languages.

Table 4. Mean Standard Scores for FACES Direct Assessment Measures for Children Taking the Assessment in Spanish at Both Waves: Fall 2009, Spring 2010, Fall–Spring Change

Scales (Standard Scores)	n	Fall 2009	Spring 2010	Fall–Spring Change
		Mean	Mean	Mean
PPVT-4 ^a	67	61.4	62.2	0.8
TVIP	197	83.8	85.6	1.8
EOWPVT ^a	208	66.3	70.0	3.6
EOWPVT-SBE ^b	63	84.5	87.6	3.2
WM III: Letter Word Identification	104	81.1	87.3	6.2*
WM III: Spelling	190	90.0	89.5	-0.5
WM III: Applied Problems	146	82.3	84.2	1.9

Source: Fall 2009 and Spring 2010 FACES Direct Child Assessment.

Note: Statistics are weighted to represent all children who entered Head Start for the first time in fall 2009 and were still enrolled in spring 2010.

* Asterisk indicates that the difference between fall and spring scores is statistically significant at the $p \leq .05$ level.

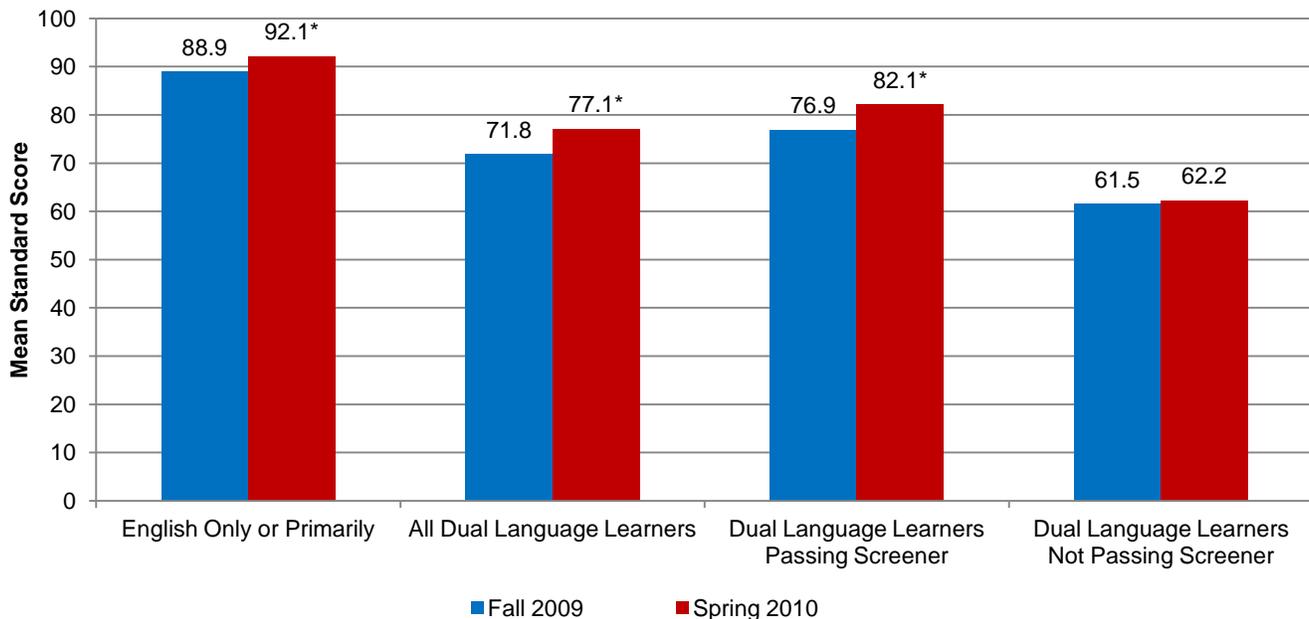
^a These scores, normed against children in the U.S. nationally, allow for comparison of children’s skills with English-speaking peers. Scores in this table are for children who did not pass the language screener threshold and took the remainder of the assessment in Spanish at both waves.

^b The EOWPVT-SBE standard scores allow for comparison of children’s vocabulary skills relative to all young Hispanic children nationally. These standard scores are only available for children age 4 and older at the time of assessment. In this table, we only report EOWPVT-SBE standard scores for children who entered the program at age 4.

Across language groups and regardless of performance on the language screener, children score behind peers in their English receptive vocabulary in the fall and the spring, but most make progress towards norms in this area during the program year. Not surprisingly, children from households where a language other than English is spoken, including Spanish and Other languages, score considerably lower in English receptive vocabulary skills in both fall and spring than children who come from households where English is the primary language spoken in the home (Figure 9). Children who are from households where a language other than English is spoken and unable to pass the language screener have the lowest scores at both time points and do not make progress toward norms during the program year. All other children make progress toward norms. We see similar patterns in expressive vocabulary, with those unable to pass the language screener performing most poorly in fall and spring and not making progress towards peers during the program year.

When examining Spanish receptive vocabulary skills among children from households where Spanish is spoken, there are only small differences based on children’s ability to pass the language screener. That is, both children who do and do not demonstrate enough English proficiency to be assessed in English have similar Spanish receptive vocabulary skills in fall and spring. Neither group makes progress relative to peers during the program year. Finally, when looking at expressive vocabulary among children from households where Spanish is spoken, those who are able to pass the language screener have stronger skills than those who do not in both fall and spring. Neither group makes statistically significant progress toward norms during the program year, however. Regardless of performance on the screener, both groups of children from households where Spanish is spoken score closer to the Spanish-bilingual norms than to the English norms on expressive vocabulary.

Figure 9. Mean PPVT-4 Standard Scores by Language of Assessment: Fall 2009-Spring 2010



Source: Fall 2009 and Spring 2010 FACES Direct Child Assessment.

Note: Statistics are weighted to represent all children who entered Head Start for the first time in fall 2009 and were still enrolled in spring 2010.

Dual language learners are children identified as having a primary home language other than English for the purposes of the direct assessment. It includes children from homes where Spanish or Other languages are primarily spoken.

*Asterisk indicates that the difference between the fall and spring scores is statistically significant at the $p \leq .05$ level.

Child Social-Emotional Development

FACES 2009 uses measures from a variety of sources—teacher, parent, assessor, and direct assessment—to provide multiple perspectives on children’s positive and challenging behaviors that may affect their ability to learn and interact with peers and adults. Using items taken from the Behavior Problems Index (Peterson and Zill 1986), Personal Maturity Scale (Entwisle et al. 1997), and Social Skills Rating System (Gresham and Elliott 1990), we present teacher reports of children’s cooperative classroom behavior, such as making friends easily and waiting their turn in games or other activities, as well as problem behaviors in the classroom, such as being very restless and unable to sit still or disrupting ongoing activities. We also present teachers’ ratings of children’s approaches to learning, using the ECLS–K Approaches to Learning scale (U.S. Department of Education 2002).⁴⁵

Parents, using a similar set of items as teachers, also reported on children’s social skills and problem

behaviors in the home environment. Using the Leiter International Performance Scale-Revised Examiner Rating Scale (Leiter-R; Roid and Miller 1997), FACES assessors rated children’s behaviors during the assessment situation in such areas as attention, organization and impulse control, activity level, and sociability. Finally, for FACES 2009, a pencil tapping task (Blair 2002; Diamond and Taylor 1996; Smith-Donald et al. 2007) was added to capture 4-year-old children’s executive functioning. As with cognitive measures, we describe the skills and behaviors of all children, and then of important groups of children (that is, by children’s age at program entry, gender, race/ethnicity, and number of family risks).

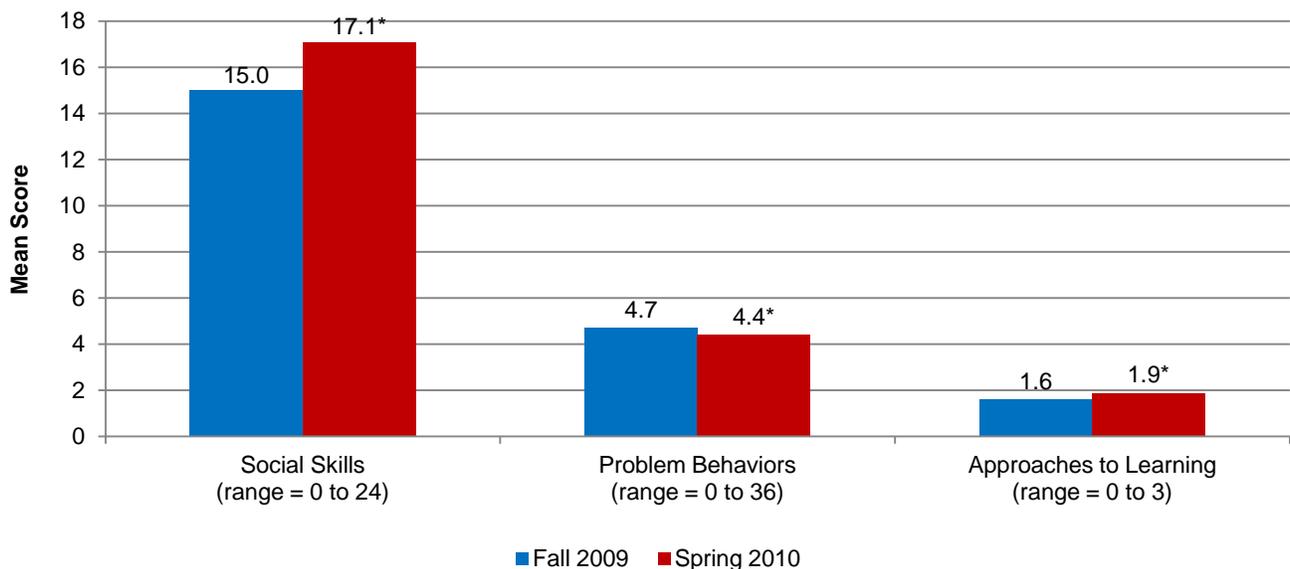
With the exception of Leiter-R standard scores, we report on children’s social-emotional outcomes and approaches to learning on criterion-referenced measures using raw scores. While standard scores provide information on children’s performance relative to same-age peers, raw scores allow for measurement of change or growth in performance

over time. They are an indicator of absolute, rather than relative, performance. As noted previously, changes in children’s skills and development during the program year reflect a range of influences including maturation, program and family influences, and other influences in children’s lives.

Both teachers and parents report that children show growth in social skills during their first Head Start year. Consistent with prior FACES cohorts (Aikens et al. 2010; Zill et al. 2006), teachers report that children demonstrate better social skills and cooperative classroom behavior by the end of the program year (Figure 10). They also rate children as having fewer problem behaviors by the spring, as well as more positive approaches to learning. As in FACES 1997, 2000, and 2006, teachers rate children as showing fewer hyperactive behaviors⁴⁶ by spring. They do not report differences in children’s aggressive and withdrawn behaviors between program entry and the end of the program year. This is similar to patterns found in FACES 1997, 2000, and 2006.⁴⁷ On the Approaches to Learning scale in the fall, on average, teachers rate children as “sometimes” to “often” (mean of 1.6 out of 3) demonstrating skills related to

attention, persistence, adaptability, and independence. They report higher scores for children in the spring (1.9 out of 3), suggesting that on average children “often” demonstrate these skills by the end of the program year. Parents report that children demonstrate more social skills and positive approaches to learning, on average, in the spring than in the fall, but they do not report differences in problem behaviors between program entry and the end of the program year. Assessors also report differences in children’s behavior during the assessment situation between program entry and the end of the program year, including the child’s attention, organization and impulse control, activity level, and sociability. Finally, children are able to inhibit their initial impulse and respond correctly to more trials on the pencil tapping task by the end of the program year than in the fall. In fact, in comparison to responding correctly 43 percent of the time at the beginning of the year, children are able to do so 61 percent of the time in the spring. Head Start children in a small study (mean age = 60.5 months) responded correctly across a similar percentage of trials (59 percent; Smith-Donald et al. 2007).

Figure 10. Teacher Reports of Social-Emotional Development: Fall 2009-Spring 2010



Source: Fall 2009 and Spring 2010 FACES Teacher Child Report.

Note: Statistics are weighted to represent all children who entered Head Start for the first time in fall 2009 and were still enrolled in spring 2010.

*Asterisk indicates that the difference between the fall and spring scores is statistically significant at the $p \leq .05$ level.

Age. Teachers and parents report that both 3- and 4-year-old children demonstrate more social skills and positive approaches to learning, on average, by the end of the program year. Parents do not report 3- or 4-year-old children as demonstrating fewer problem behaviors by spring, and teachers only report those who entered at age 3 as demonstrating fewer hyperactive behaviors, on average, in spring. Assessors rate both 3- and 4-year-olds as demonstrating better social/cognitive skills by the end of the program year, including attention and organization/impulse control. With the exception of teacher-reported withdrawn behaviors and parent-reported problem behaviors, 4-year-olds are reported as having better social-emotional outcomes than 3-year-olds in the spring.

Gender. Teachers and parents report that both boys and girls demonstrate more social skills and more positive approaches to learning by spring. However, teachers only report girls as having fewer problem behaviors and hyperactive behaviors on average by the end of the program year. Assessors rate both boys and girls as demonstrating better social/cognitive skills by the end of the program year. Both boys and girls are able to inhibit their initial impulse and respond correctly on the pencil tapping task more frequently by the end of the program year than they were in the fall. In fact, by the spring, boys respond correctly 59 percent of the time and girls are able to do so 64 percent of the time. Across measures, girls are reported as having better social-emotional outcomes than boys in the spring.

Race/ethnicity. Teachers report that all children, regardless of race/ethnicity, demonstrate more social skills on average by the end of the program year. However, they report only Hispanic/Latino children as having fewer problem behaviors and hyperactive behaviors in the spring. Only parents of Hispanic/Latino children report that their children demonstrate more social skills and positive approaches to learning, on average, in the spring. Assessors rate all children, regardless of race/ethnicity, as demonstrating better social/cognitive skills by the end of the program year. Regardless of race/ethnicity, children are able to inhibit their initial impulse and respond correctly across more trials on the pencil tapping task by the end of the program year. In the spring, teachers report Hispanic/Latino children as having more social skills than African

American children and fewer problem behaviors, aggressive behaviors, hyperactive behaviors, and withdrawn behaviors than African American and White children. Teachers report African American children as demonstrating fewer positive approaches to learning behaviors than White and Hispanic/Latino children. Parents report African American children as having fewer social skills than White and Hispanic/Latino children in the spring, and, in contrast to teacher reports, parents report Hispanic/Latino children as having more problem behaviors than White and African American peers. Assessors also rate African American children as having fewer social/cognitive skills than White and Hispanic/Latino children in the spring, including organization/impulse control. Finally, White children responded correctly across more trials on the pencil tapping task in the spring than African American and Hispanic/Latino children.

Family risk. Teachers report that all children demonstrate more social skills and positive approaches to learning and fewer problem behaviors, on average, by the end of the program year. They also report children with no family risks and those with two or more risks as demonstrating fewer hyperactive behaviors in the spring. Parents report that children with one and two or more risks demonstrate more social skills and positive approaches to learning on average in the spring. Assessors rate all children, regardless of number of family risks, as demonstrating better social/cognitive skills by the end of the program year. Regardless of the number of family risks, children are able to inhibit their initial impulse and respond correctly across more trials on the pencil tapping task by the end of the program year. In the spring, parents report children with two or more family risks as having more problem behaviors than parents of children with one or no risks. Assessors also rate children with two or more family risks as having fewer social/cognitive skills than children with one or no risks in the spring, including organization/impulse control.

Child Health and Physical Development

Parents and teachers reported on several aspects of children's health and physical development, including disability status and health and developmental conditions or concerns. As in FACES 2006, each child's height and weight were measured to support

analyses of overweight, obesity, or underweight status. In this section, we provide information from teacher reports on children’s status in spring 2010, including child disability and Individualized Education Plan (IEP) or Individualized Family Service Plan (IFSP) status. From parent reports, we present child health status during the program year. We also report on Head Start children’s height and weight and BMI in the fall and spring of the program year.

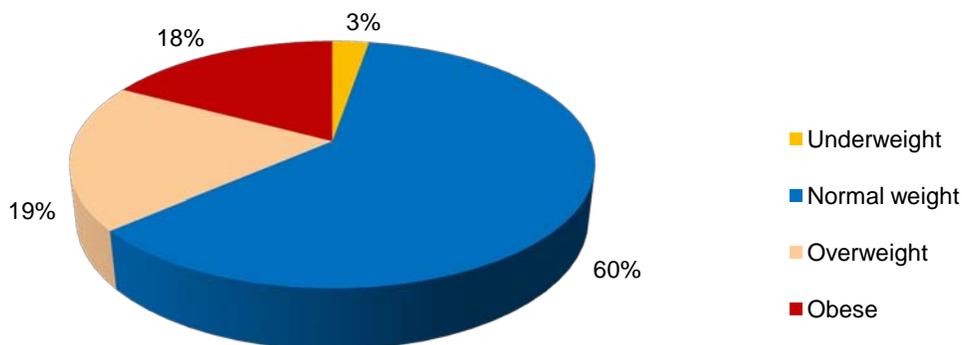
Approximately 14 percent of children are reported by their teacher as having a disability in the spring. Among Head Start children identified by teachers as having a disability, speech and language impairments (77 percent) and cognitive impairments (24 percent) are the most common. About a quarter of those identified as having a disability are reported by teachers to have more than one impairment (27 percent). About two-thirds of Head Start children with an identified disability have an IEP or IFSP in the spring (67 percent).

As in the fall, 81 percent of children are rated as having “excellent” or “very good” health by their parents in the spring. Only a small percentage of children are reported as having “fair” or “poor” general health at the end of the program year (6 percent). This

finding is consistent with prior FACES cohorts. Across age, gender, racial/ethnic, and risk factor groups, there were no changes in reports of children’s health status during the program year. However, at the end of the program year, fewer 4-year-old than 3-year-old children, boys than girls, Hispanic/Latino children than those of other racial/ethnic backgrounds, and children with two or more risks than those with one or no risks are rated as having “excellent” or “very good” health.

As in the fall, Head Start children have an average BMI⁴⁸ that is above average for their age range at the end of the program year. In fact, about 19 percent of children are overweight, and 18 percent are obese (Figure 11).⁴⁹ A similar percentage (18 percent) of preschoolers in the ECLS–B were obese (Anderson and Whitaker 2009). On average, children grew about one inch and gained nearly three pounds during their first Head Start year. There were no changes in their BMI between the beginning and end of the year for the population as a whole and across all subgroups (age, gender, race/ethnicity, and number of risk factors). However, at the end of the program year, more 4-year-old than 3-year-old children, boys than girls, and Hispanic/Latino children than those of other racial/ethnic backgrounds are obese.

Figure 11. Child BMI Categories: Spring 2010



Source: Spring 2010 FACES Direct Child Assessment.

Note: Statistics are weighted to represent all children who entered Head Start for the first time in fall 2009 and were still enrolled in spring 2010.

SUMMARY

Child Cognitive, Social-Emotional, and Physical Development

With the exception of letter-word knowledge, Head Start children assessed in English score below norms

across developmental areas, including language, literacy, and mathematics in both the fall and spring of their first program year. Children assessed in Spanish remain below norms across all developmental areas by the spring. Furthermore, children assessed in Spanish make progress towards norms only in the

area of letter-word knowledge, while those assessed in English make progress toward norms across areas.

Children are able to inhibit their initial impulse and respond correctly across more trials on the pencil tapping task by the end of the program year, suggesting improvements in children's executive functioning skills. Both teachers and parents report that children show growth in their social skills during their first Head Start year. Teachers also rate children as having fewer problem behaviors by the spring, including hyperactive behaviors, as well as more positive approaches to learning. Based on the child's behavior during the direct assessment, assessors rate children as demonstrating better social/cognitive skills at the end of the year, including attention and organization/impulse control.

At the end of the Head Start year, 14 percent of children have an identified disability, with the majority of these reported to be speech or language impairments. In addition, more than one-third of children (37 percent) are overweight or obese at the end of the first program year. Although children grow about one inch and gain nearly three pounds during their first Head Start year, there are no changes in their BMI between the beginning and end of the year. However, there are differences in children's weight outcomes in the spring: more 4-year-old than 3-year-old children, boys than girls, and Hispanic/Latino children than those of other racial/ethnic backgrounds are obese. Finally, on average, parents report children to be in excellent or very good physical health, and there are no differences in these reports between program entry and the end of the program year. As with child obesity, there are differences by important subgroups in ratings of children's health status. At the end of the program year fewer 4-year-old than 3-year-old children, boys than girls, Hispanic/Latino children than those of other racial/ethnic backgrounds, and children with two or more risks than those with one or no risks are rated as having "excellent" or "very good" health.

HEAD START TEACHERS AND CLASSROOMS

The FACES conceptual framework envisions a relationship between the provision of quality early care and educational services and children's developmental outcomes. To examine teacher and

classroom characteristics that may relate to the quality of services for children and parents, FACES conducted interviews with lead teachers in each classroom where children who were new to Head Start in fall 2009 were still enrolled in spring 2010. Information on the characteristics, credentials, and beliefs of teachers as well as reported classroom activities, curricula, and assessments were collected. As discussed in greater detail below, in spring 2010, FACES also conducted classroom observations of the quality of classroom materials and resources and teacher-child interactions.

Teacher Characteristics

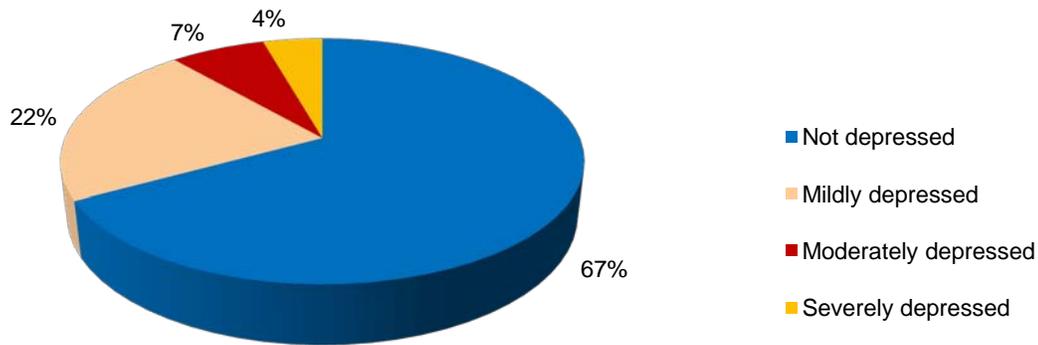
Most Head Start lead teachers are female (99 percent) and 57 percent are between the ages of 30 and 49. Five percent are 60 or older and approaching retirement. Forty-two percent of Head Start teachers are White, 31 percent are African-American, and 22 percent are Hispanic/Latino.

The average lead teacher has been in the Head Start classroom for nearly nine years, and more than two-thirds (67 percent) have five or more years of experience teaching in Head Start. The average annual salary is \$28,527. Eighty-five percent of Head Start teachers have at least an associate's (A.A.) degree, and half (50 percent) have at least a bachelor's (B.A.) degree.⁵⁰

In addition to their degrees, many Head Start lead teachers have specific training in child development and teaching. Almost half (47 percent) pursued a field of study that included early childhood education and 95 percent of teachers who have completed at least some college enrolled in at least six courses in early childhood education. Twenty-eight percent of teachers completed a course on DLLs. Forty-seven percent of Head Start teachers report having a Child Development Associate (CDA) certificate, 43 percent a teaching certificate or license, and 36 percent are currently enrolled in teacher-related training.

As teachers' mental health status could affect their classroom behaviors and interactions with children, teachers were asked about their depressive symptoms using the short (12-item) form of the CES-D. As shown in Figure 12, most Head Start teachers (67 percent) do not report symptoms of depression. However, 4 percent of teachers report symptoms of severe depression, and another 7 percent report symptoms of moderate depression.

Figure 12. Depressive Symptoms Among Head Start Teachers: Spring 2010



Source: Spring 2010 FACES Teacher Interview.

Note: Statistics are weighted to represent all teachers serving children who entered Head Start for the first time in fall 2009 and were still enrolled in their classrooms in spring 2010.

FACES measured teacher beliefs and attitudes using a 24-item Teacher Beliefs Scale (Burts et al. 1990) that consists of statements worded to reflect positive attitudes and knowledge of generally accepted practices in preschool settings, or to reflect a lack of such attitudes and knowledge. Scores regarding attitudes toward developmentally appropriate practice average 8.0 out of 10 overall; higher scores indicate attitudes consistent with practice appropriate for working with young children. Teacher’s scores on child-initiated practice are 4.5 out of 5 and on didactic, strictly teacher-initiated practice, 2.5 out of 5. Teachers scoring higher on developmentally appropriate practices are likely to endorse such items as, “Head Start classroom activities should be responsive to individual differences in development” and to disagree with such items as “Children should work silently and alone on seatwork.” The former item contributes to the child-initiated practice scale and the latter to the didactic practice scale. Teachers who score higher on developmentally appropriate practice are also likely to score higher on child-initiated practice.

Head Start teachers are generally positive about their profession and their program. On a measure of management climate (Lambert et al. 1999; Lambert 2002), teachers rate their programs relatively positively (on average, 3.7 on a 5-point scale). The scale asks teachers to rate the strength of their agreement with statements about the program such as, “Provides enough assistance to teachers in the classroom” and “Promotes teamwork among

teachers.” Teachers also report high levels of satisfaction. Ninety-four percent of lead teachers agree or strongly agree with the statement, “I really enjoy my present teaching job,” while 98 percent agree or strongly agree that “I am certain I am making a difference in the lives of the children I teach.” Eighty-eight percent strongly agree they would again choose teaching as a career. The average teacher satisfaction score is 4.5 out of 5 points.

Classroom Educational Environment

We asked teachers to report on the curricula and assessment tools they use in the classroom. Fifty-three percent of all Head Start teachers report using the Creative Curriculum as their primary curriculum, whether they use only one curriculum or a combination of curricula. The High/Scope Curriculum is also common, with 15 percent of teachers reporting its use. Nine percent of teachers report using another widely available curriculum, and only 2 percent of teachers report using a locally designed curriculum. Other types of curricula⁵¹ are used by 22 percent of teachers.

The assessment tools that teachers report using follow a similar pattern. More than one-third (38 percent) of Head Start teachers report using the Creative Curriculum as their primary assessment tool. Smaller percentages report using High/Scope Child Observation Record (COR) (9 percent), Learning Accomplishment Profile (LAP) System (8 percent), and Desired Results Developmental Profile (DRDP)

(7 percent) assessment tools. Another 6 percent report using a program designed assessment tool.

Using curriculum and assessment tools that are aligned with each other and with program standards is considered advantageous to supporting children's learning (National Research Council 2008). Of the Head Start teachers who report using the Creative Curriculum, 68 percent use the Creative Curriculum assessment tool and 32 percent use a different tool. Thirty percent of teachers overall report using both the Creative Curriculum and its assessment tool. Six percent of teachers use both the High/Scope curriculum and the High/Scope COR.

Teachers also reported class sizes, including the number of children at each age. On average, teachers report serving 17 children in the classroom and that by the spring more than three-quarters (79 percent) of the children in their classrooms are 4 years old or older. Based on these reports, in the spring, 75 percent of all Head Start classrooms are mixed-age classrooms. Teachers report that classes meet, on average, 4.6 days each week for 5.8 hours each day.

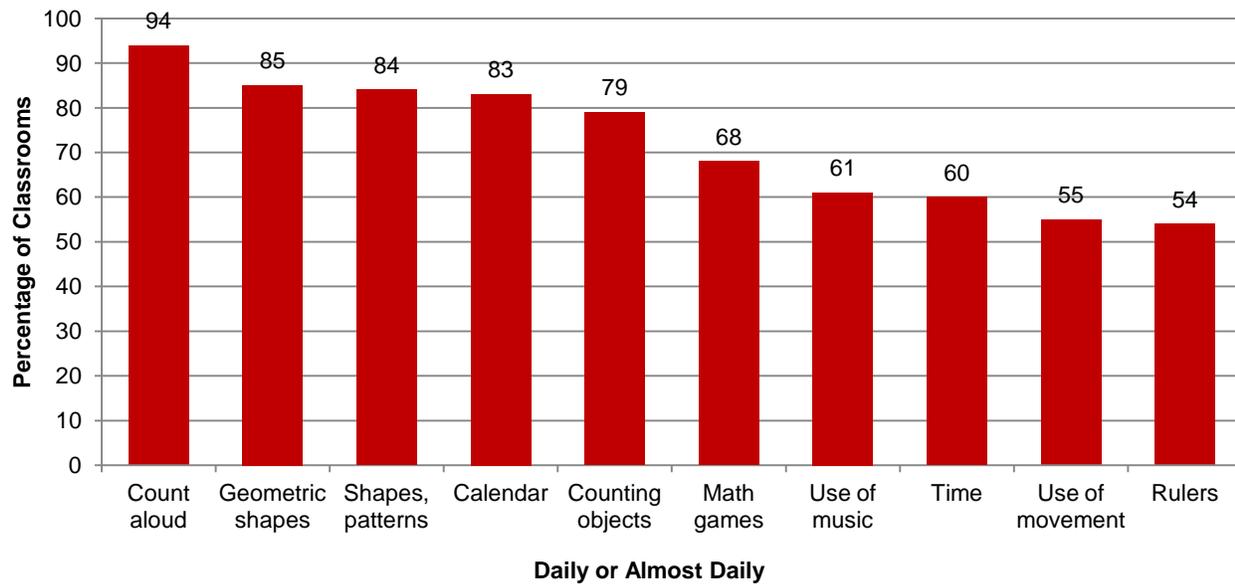
FACES 2009 asked teachers about the types and frequency of learning activities in early literacy and mathematics that were commonly used in their classrooms. For most of these, more than half of teachers report engaging in these activities daily or almost daily. The most common reading and language activities (reported as occurring daily or almost daily in 75 percent or more of classrooms) include working on letter naming, practicing writing letters, discussing new words, working on phonics, listening to the teacher read stories where children can see the print, learning about conventions of print, and writing names. Activities occurring less frequently, although still occurring daily or almost daily in at least 50 percent of classrooms, include dictating stories to an adult, retelling stories, learning about rhyming words and word families, and learning about common prepositions. Only listening to the teacher read stories where children do not see print is reported as occurring less often than daily or almost daily by a majority of teachers.

Similarly, teachers report frequent math-related activities in their classrooms. As shown in Figure 13, the most common math activities (reported as occurring daily or almost daily in 75 percent or more of classrooms) include counting out loud, working with geometric and counting manipulatives, engaging in activities that are calendar-related and those that involve shapes and patterns. All other math activities addressed occur less often but still daily or almost daily in at least 50 percent of classrooms: playing math-related games, using music and creative movement or drama to understand math concepts, working with rulers or other measuring instruments, and engaging in activities related to telling time.

Classroom Observations

To measure quality of Head Start classrooms, FACES 2009 used the full Classroom Assessment Scoring System (CLASS; Pianta et al. 2008) in conjunction with a shortened version of the Early Childhood Environment Rating Scale-Revised (ECERS-R; Harms et al. 1998) in spring 2010. The CLASS measures classroom quality in terms of both instructional and social-emotional aspects of the environment across three domains of interaction: Instructional Support, Emotional Support, and Classroom Organization. Instructional Support measures the quality of instructional practices used in the classroom (its three dimensions include Concept Development, Quality of Feedback, and Language Modeling).⁵² Emotional Support measures the social and emotional functioning in the classroom (its three dimensions include Classroom Climate, Teacher Sensitivity, and Regard for Student Perspectives).⁵³ Finally, Classroom Organization measures the teacher's ability to organize the classroom to make efficient use of class time (its three dimensions include Behavior Management, Productive Use of Time, and Instructional Learning Formats).⁵⁴ The ECERS-R is a global rating of classroom quality based on structural features of the classroom. It has been used historically in FACES (starting with the earlier version of the ECERS [Harms and Clifford 1980] in the 1997 cohort) and in many other large-scale studies. For the first time in FACES, FACES 2009 used the short form of the ECERS-R that includes an abbreviated set of items based on findings in other large-scale studies.⁵⁵

Figure 13. Percentage of Classrooms Conducting Math Activities Daily or Almost Daily: Spring 2010



Source: Spring 2010 FACES Teacher Interview.

Note: Statistics are weighted to represent all classrooms serving children who entered Head Start for the first time in fall 2009 and were still enrolled in their classrooms in spring 2010.

Both the CLASS and the ECERS-R items are scored on a 7-point scale, with higher scores reflecting better quality care.

Classroom observations also included observer counts of child-adult ratios and group sizes. Observations were done in a representative sample of 370 classrooms of 3- and 4-year-old children in their first year of Head Start. Observations lasted for four hours, on average, and were typically completed in the mornings.⁵⁶

The National Association for the Education of Young Children (NAEYC) offers professional standards for appropriate child-teacher ratios (9 children per adult in classrooms with 3-year-olds and 10 children per adult in classrooms with 4-year-olds) and group sizes (18 or fewer children in classes with 3-year-olds and 20 or fewer children in classes with 4-year-olds). Head Start Program Performance Standards provide similar guidelines (8.5 children per adult in classrooms with 3-year-olds and 10 children per adult in classrooms with 4-year-olds, and a maximum group size of 17 or 20, respectively; see National Child Care Information and Technical Assistance Center [NCCIC] 2008). In spring 2010, the average Head Start classroom is observed to have 6.2 children to each adult, and the average

observed group size is 14.2,⁵⁷ falling well within professional guidelines and Head Start Program Performance Standards. Half of classrooms (49 percent) have observed group sizes of fewer than 15 children and the majority (97 percent) have ratios of 9 to 1 or lower.

As previously described, FACES 2009 used the ECERS-R short form. Thus, the FACES 2009 total score represents a different (smaller) set of items than the total scores from prior rounds and may not represent global quality in the same way. The average ECERS-R short form score was 4.3, and 75 percent of classrooms fell in the minimal to good range (between 3 and 5). Few classrooms (5 percent) scored below 3 (considered the threshold for minimal quality), and less than one percent scored above a 6 (considered excellent quality).

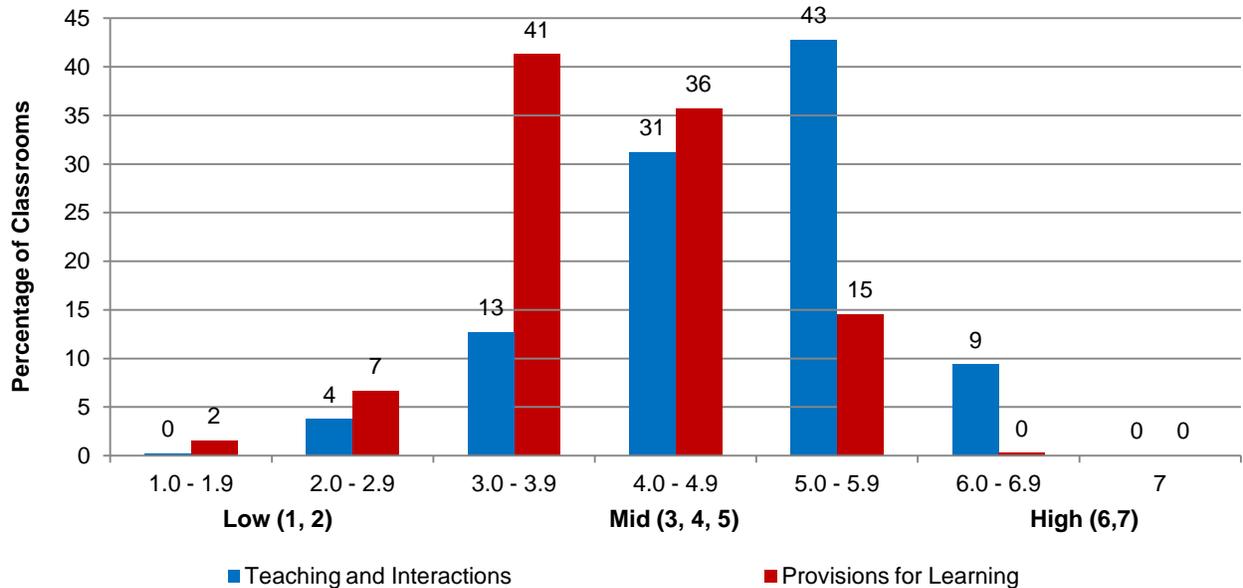
The short form of the ECERS-R addresses dimensions of quality identified in other large scale studies. Two factors reported in the Multi-State Study of Prekindergarten represent the key dimensions of quality tapped by the full ECERS-R (Clifford et al. 2005).⁵⁸ The Provisions for Learning subscale focuses on materials available in the classroom and the arrangement of classroom space, while the Teaching

and Interactions subscale focuses on the quality of teacher-child interactions. The two subscales reliably assess the areas of classroom quality most proximal to learning.

Head Start classrooms were more likely to score high on Teaching and Interactions (4.7) than Provisions for

Learning (4.0). Eighty percent scored between 3 and 5 out of a possible 7 on the Provisions for Learning subscale, and an additional 9 percent scored below 3. On the Teaching and Interactions subscale, half of observed classrooms scored between 3 and 5 (50 percent), 5 percent scored below a 3, and 45 percent scored above 5 (see Figure 14).

Figure 14. Percentage Distribution of Scores on ECERS-R Short Form Factors: Spring 2010



Source: Spring 2010 FACES Classroom Observation.

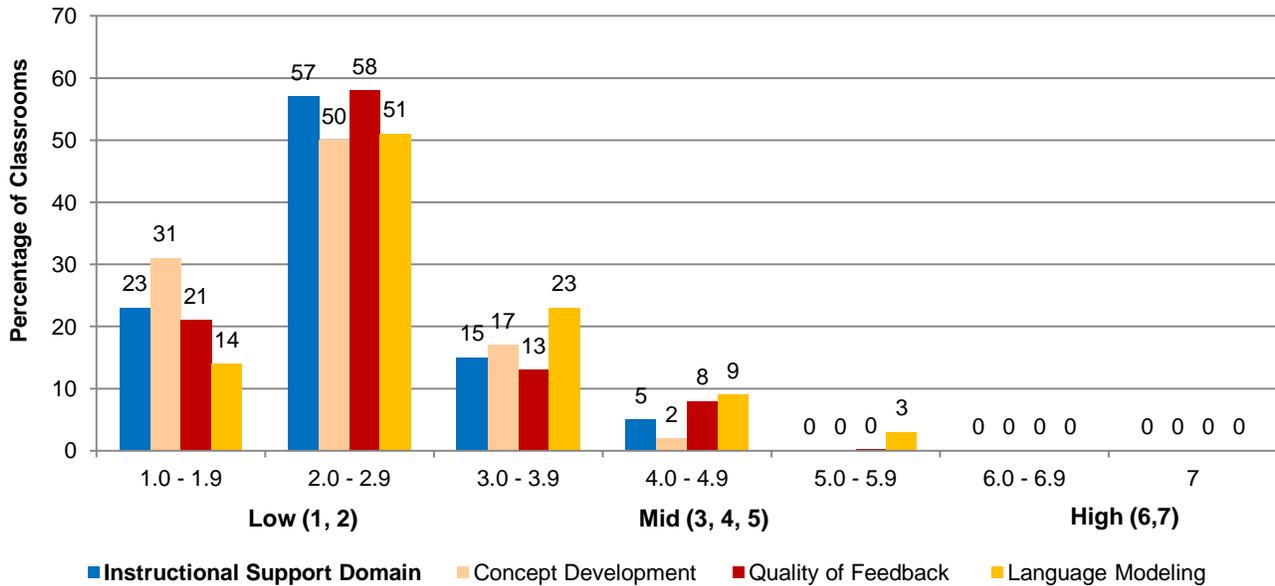
Note: Statistics are weighted to represent all classrooms serving children who entered Head Start for the first time in fall 2009 and were still enrolled in their classrooms in spring 2010.

While remaining in the minimal to good range, mean scores on these two subscales have changed over time. The mean Teaching and Interactions subscale score was stable from 2001 to 2004, declined between 2004 and 2007 and then increased between 2007 and 2010; scores were 5.4 in 2001, 5.0 in 2004, 4.1 in 2007, and 4.7 in 2010. On Provisions for Learning, scores declined between each cohort from 2001 to 2007 and then increased between 2007 and 2010; mean scores were 4.7 in spring 2001, 4.1 in 2004, 3.6 in 2007, and 4.0 in 2010.⁵⁹ Across cohorts, classrooms scored higher on Teaching and Interactions than Provisions for Learning. This pattern is similar to that reported in the NCEDL Study of Prekindergarten (Clifford et al. 2005).

On the Instructional Support domain of the CLASS, classrooms scored at the low end of the 7-point scale.

Average quality was 2.3, with the majority of classrooms (87 percent) rated in the low range (1 or 2 points). Thirteen percent of classrooms scored in the middle range on the domain (3, 4, or 5 points), and none scored in the high range (6 or 7 points). A larger percentage of classrooms scored in the low range in spring 2007 (96), when the average score was 1.9.⁶⁰ On the dimensions (subscales) of Instructional Support, scores on Concept Development (2.1) were lower than those for Language Modeling (2.5) and Quality of Feedback (2.3). Most classrooms scored in the low range on Concept Development (85 percent), Language Modeling (75 percent), and Quality of Feedback (84 percent), although 25 percent of the classrooms had a middle rating on Language Modeling (see Figure 15).

Figure 15. Percentage Distribution of Scores on CLASS Instructional Support Domain and Dimensions: Spring 2010



Source: Spring 2010 FACES Classroom Observation.

Note: Statistics are weighted to represent all classrooms serving children who entered Head Start for the first time in fall 2009 and were still enrolled in their classrooms in spring 2010.

On both the CLASS Emotional Support and Classroom Organization domains, classrooms scored in the middle range of the scale. Average quality was 5.3 and 4.7, respectively. Nearly all classrooms are rated in the middle range on both domains (94 and 98 percent, respectively).⁶¹ On the Emotional Support dimensions, classrooms were rated highest in Positive Climate (5.3), followed by Teacher Sensitivity (4.7), Regard for Student Perspectives (4.5), and Negative Climate (1.3)⁶² (see Figure 16). On the Classroom Organization dimension, classrooms were rated highest in Behavior Management (5.0) and Productivity (4.9) followed by Instructional Learning Formats (4.0) (see Figure 17).

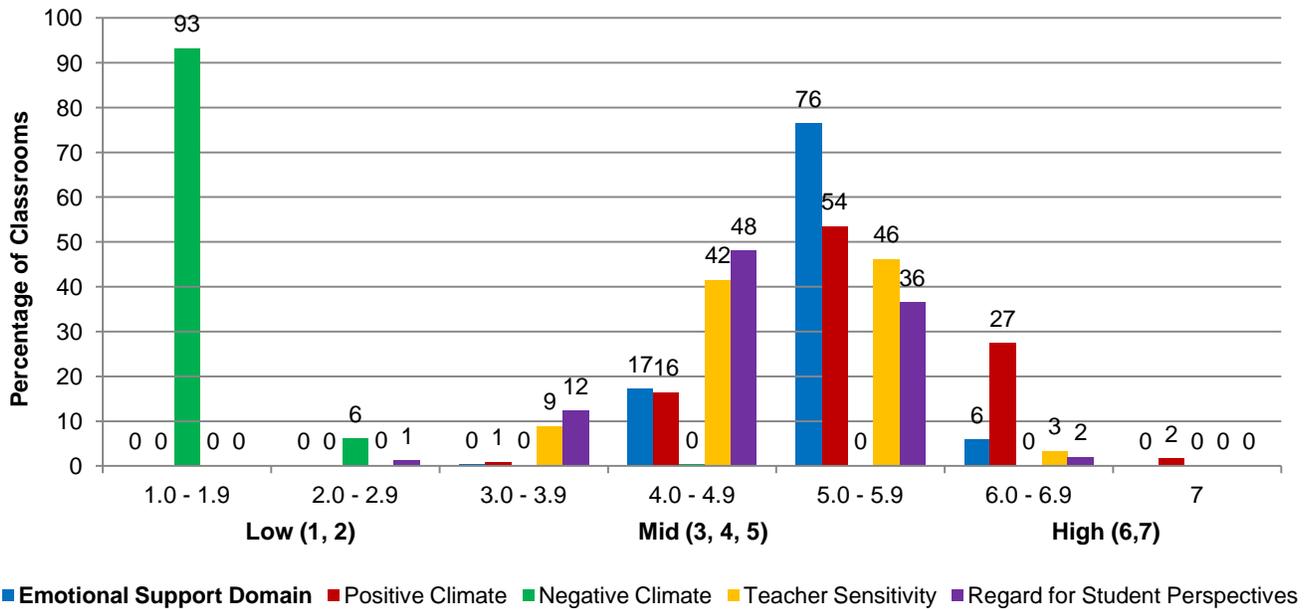
While the CLASS does not have normative data, the CLASS Technical Appendix (Pianta et al. 2008) reports mean scores from several large-scale studies that used the CLASS or its precursor, the Classroom Observation System. The mean scores in Head Start classrooms fall within or below ranges reported in these studies. Lower ratings (that is, in the 1 to 2 range) on the Instructional Support domain, relative to

the other CLASS domains, are consistent with findings reported in the CLASS appendix.

SUMMARY

Head Start teachers bring many years of experience to the classroom, with more than two-thirds having been in the classroom for five years or more. The vast majority of teachers have at least an associate’s (A.A.) degree, and half have at least a bachelor’s (B.A.) degree. Teachers’ attitudes appear consistent with developmentally appropriate practice, and they report engaging in a variety of language, literacy, and mathematics activities daily or almost daily. Classroom observations show that group sizes and child-adult ratios fall well within professional guidelines and Head Start Program Performance Standards. On average, classrooms score in the minimal to good range for classroom materials and arrangement and for the quality of teacher-child interactions as measured by the ECERS-R. Instructional support was rated in the low range and emotional support and classroom organization in the middle range on the CLASS, a pattern consistent with other studies.

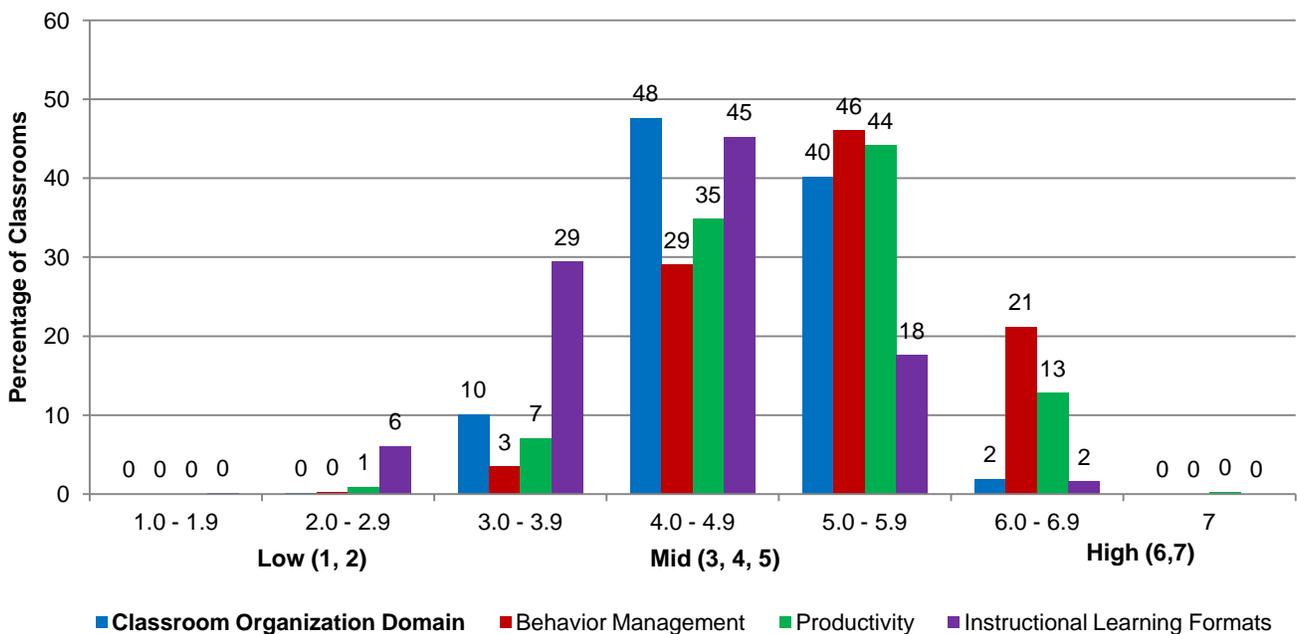
Figure 16. Percentage Distribution of Scores on CLASS Emotional Support Domain and Dimensions: Spring 2010



Source: Spring 2010 FACES Classroom Observation.

Note: Statistics are weighted to represent all classrooms serving children who entered Head Start for the first time in fall 2009 and were still enrolled in their classrooms in spring 2010.

Figure 17. Percentage Distribution of Scores on CLASS Classroom Organization Domain and Dimensions: Spring 2010



Source: Spring 2010 FACES Classroom Observation.

Note: Statistics are weighted to represent all classrooms serving children who entered Head Start for the first time in fall 2009 and were still enrolled in their classrooms in spring 2010.

CORRELATES OF CLASSROOM QUALITY AND TEACHER ATTITUDES

We now shift to addressing the question of how aspects of classroom quality relate to child progress over the course of the first Head Start year (fall 2009 to spring 2010). As a preliminary step, in this section of the report we examine associations between teacher and classroom characteristics to explore whether there are mediators linking teacher backgrounds with observed classroom quality. In the next section, we examine associations between teacher and classroom characteristics and child outcomes. Note that these associations should not be interpreted as causal relationships.

We used two-level hierarchical linear models (HLM), with classrooms nested within programs, to examine the teacher characteristics associated with classroom quality and teacher attitudes, controlling for characteristics of the program. The use of HLM recognizes that teachers/classrooms in the same program are not independent of each other because of shared resource levels, policies, and program practices. Similar to the descriptive findings, the analyses were weighted at each level to represent programs and classrooms serving children who entered Head Start for the first time in fall 2009 and were still enrolled in spring 2010. Observations with missing data on any of the covariates were excluded from the analyses.⁶³ These analyses replicate those conducted for the FACES 2006 cohort (Aikens et al. 2010).

As measures of classroom quality, we used the Teaching and Interactions and Provisions for Learning subscales from the short form of the ECERS-R. From the CLASS, we used the Instructional Support domain, the Language Modeling dimension, the Emotional Support domain, the Positive Climate dimension, and the Classroom Organization domain. We also examined correlates of teacher attitudes, including teachers' level of satisfaction with teaching as a career and their attitudes toward developmentally appropriate practice (DAP), as these may be mediators that link education levels or professional development with the quality of classroom/teacher practice.

Independent variables at the teacher/classroom level included teacher education, teaching experience,

reported depressive symptoms, frequency of mentoring, and perceived management support. Perceived management support is a composite that reflects Head Start teachers' perceptions of support provided by program management to them and other teaching staff, based on 12 items from the Program Management Inventory (PMI; Lambert et al. 1999). The analyses of classroom quality also included DAP attitudes and teacher satisfaction with teaching as a career.

Program-level control variables in each of these analyses included program socioeconomic status (SES), as measured by the average household poverty ratio in a program; dual language learners (DLLs), as a percentage of the total enrollment; percentage of teachers using a curriculum and assessment from the same package; rate of teacher turnover in the prior program year; and adjusted mean teacher salary.

All outcomes were z-scored so that the coefficients may be interpreted as the change in the teacher or classroom outcome in standard deviation units for each one-point increase in the respective independent variable.⁶⁴

We find that teacher credentials—education and experience—are associated with the emotional climate of a classroom as well as with teacher satisfaction and attitudes. First, teacher education at the AA level (as compared to a high school education or less) is related to higher scores in the CLASS Positive Climate dimension (ES = .34) and higher teacher satisfaction (ES = 1.27); teacher education at the BA level or above is not related to climate or satisfaction. Having 20 or more years of teaching experience (as compared to 3 years or fewer) is also related to higher scores on DAP attitudes (attitudes consistent with practice appropriate for working with young children; ES = .58), while having 11–20 years of experience is related to higher teacher satisfaction (ES = .33), compared to less experienced teachers. Note that teacher experience may also have an indirect influence on classroom quality; in separate models, higher teacher satisfaction is related to higher ECERS-R Teaching and Interactions (ES = .20).

Teachers' perceptions of the management climate are associated with teacher and classroom outcomes. Teachers who perceived a more positive management

climate reported greater satisfaction with their work (ES = .30) than those with less favorable views. Surprisingly, higher levels of perceived positive management climate are associated with lower scores on CLASS Classroom Organization (ES = -.26).⁶⁵ No other teacher or classroom level characteristics are associated with classroom quality or teacher satisfaction and attitudes.

ASSOCIATIONS BETWEEN OBSERVED QUALITY AND CHILDREN'S DEVELOPMENTAL STATUS

As an element of program monitoring, Head Start assesses the quality of the classroom, focusing primarily on teacher-child interactions. Research on the relationship of classroom quality with child outcomes shows consistent although modest associations (Burchinal et al. 2011a; Zaslow et al. 2010). We conducted a series of analyses to assess the relationship between quality and outcomes in FACES 2009, taking two approaches. In the first approach, we used three-level HLM to explore associations between teacher and classroom characteristics and child outcomes. We were particularly interested in whether there are linear or nonlinear associations between classroom quality and outcomes. In the second approach, we explored the possibility of threshold effects—that the association between outcomes and quality may be stronger in higher than lower quality classrooms.⁶⁶ Recent analyses based on a number of large-scale studies have identified possible threshold effects (Burchinal et al. 2011a, 2010).⁶⁷ Note that associations identified in either approach should not be interpreted as indicating causal relationships.

In all analyses, we explored associations between measures of quality and child outcomes that are more closely aligned, as there is an emerging theory that the type of support associated with a particular quality feature is more likely to influence aligned outcomes (Zaslow et al. 2010). Thus, for child cognitive outcomes, we examined associations with CLASS Instructional Support, Language Modeling, and ECERS-R Teaching and Interactions. For social-emotional outcomes, we examined associations with CLASS Emotional Support and Positive Climate. For both cognitive and social-emotional outcomes, we examined associations with CLASS Classroom Organization, as this serves as an indicator of

teachers' ability to manage the context within which instructional interactions and emotional support take place (Pianta et al. 2008).

All analyses account for the clustering of children within classrooms and classrooms within programs because children in the same classroom and program share a common set of preschool experiences, so their outcomes are not independent.

CLASSROOM AND TEACHER CHARACTERISTICS AND CHILD OUTCOMES

We used three-level HLM to examine the associations between characteristics of teachers and classrooms and children's outcomes, controlling for child, family, and program characteristics. Similar to the descriptive findings, the analyses were weighted at each level. At the child level, analyses were weighted to represent all children who entered Head Start for the first time in fall 2009 and were still enrolled in spring 2010. At the classroom and program levels, analyses were weighted to represent classrooms and programs serving those children. Observations with missing data on any of the covariates were excluded from the analyses.⁶⁸ These analyses replicate those conducted for the FACES 2006 cohort (Aikens et al. 2010).

We estimated models of children's developmental status in the spring, controlling for their initial status as measured in the fall.⁶⁹ Outcomes included children's receptive and expressive language (PPVT-4,⁷⁰ EOWPVT); literacy (WJ Letter-Word Identification); mathematics (WJ Applied Problems); social-emotional development (teacher ratings of children's social skills, behavior problems, and approaches to learning); and executive functioning (children's performance on the pencil tapping task). The language, literacy, and mathematics outcomes in the models used equal-interval W- or GSV-scores.⁷¹ Social-emotional outcomes were measured with raw scores. W-, GSV, and raw scores are a marker of absolute, rather than relative, performance. All outcomes were z-scored so that the coefficients may be interpreted as the change in child outcome in standard deviation units for each one-point increase in the respective independent variable.

The teacher/classroom-level characteristics included aspects of quality aligned with particular outcomes (depending on the outcomes, ECERS-R Teaching and

Interactions and CLASS Instructional Support, Language Modeling, Emotional Support, Positive Climate, and Classroom Organization), teacher education, full-day class, mean peer abilities, variation in peer abilities, and teacher DAP attitudes. To test whether there is a nonlinear association between classroom quality and children's outcomes, we included both a linear and a quadratic term in the model and dropped the quadratic term if it was not significant.

The child/family-level control variables included child age at assessment, gender, race/ethnicity, household language, household poverty ratio, maternal education, maternal depressive symptoms, children's fall score on the same outcome, and time interval between the fall and spring assessments.⁷² The program-level control variables included program SES, as measured by the average household poverty ratio in a program; DLLs, as a percentage of the total enrollment; percentage of teachers using a curriculum and assessment from the same package; teacher turnover; and adjusted program mean salary.⁷³ The findings discussed here are from models that include the full set of control variables.

Children's Cognitive Outcomes

Teacher and classroom characteristics. Teachers' DAP attitudes are related to children's outcomes on the PPVT-4 (ES = .05) but not to other cognitive measures. Variation in peer abilities within classrooms is positively related to children's spring outcomes on the EOWPVT (ES = .09), meaning that the more variation in the scores of other children in the classroom, the better outcomes on this measure at the end of the program year. Peer abilities (both the mean level and variability) are unrelated to children's receptive vocabulary, letter-word knowledge, and applied problems skills in the spring. No other teacher- and classroom-level characteristics in the models were associated with children's cognitive outcomes.

Observed quality measures. Children's letter-word knowledge is positively associated with several quality measures, and all associations are linear. The higher the CLASS Instructional Support (ES = .16), Language Modeling (ES = .12), and Classroom Organization (ES = .12) score, the better the spring

letter-word knowledge, controlling for other variables in the model.

None of the classroom quality measures is significantly associated with children's receptive and expressive language or math outcomes.

Children's Social-Emotional Outcomes

Teacher and classroom characteristics. Teacher education is directly associated with one social-emotional outcome: a BA degree (relative to a high school degree) is associated with higher executive functioning scores, as measured by pencil tapping (ES = .11 - .30)⁷⁴. No other teacher- and classroom-level characteristics are associated with children's social-emotional outcomes.

Observed quality measures. Aspects of classroom quality predict two of the four social-emotional outcomes. First, children's social skills are positively associated with Classroom Organization (ES = .15), indicating that an organized classroom and efficient use of class time (including behavior management) relate to social skills. Second, children's executive functioning has a nonlinear association with CLASS Positive Climate (linear ES = .12, quadratic ES = .07). As Positive Climate scores increase from the minimum observed in Head Start classrooms (2.3) to approximately 5.5 points (slightly above the mean of 5.3), there was a decline in children's executive functioning scores; thereafter, executive functioning scores increased with Positive Climate. This indicates that executive functioning scores increase in only the highest quality classrooms, as measured by Positive Climate.

THRESHOLD EFFECTS FOR CLASSROOM QUALITY AND CHILD OUTCOMES

Identifying the level or threshold above or below which quality and outcomes show stronger associations can assist policymakers and programs as they set targets for program quality and aim to support child outcomes. We used three-level HLM to explore whether the relationship between quality and outcomes differs in higher quality versus lower quality classrooms. As a first step, we explored whether there are linear or nonlinear associations between classroom quality and outcomes; this step essentially repeated the HLM analyses described above. We

then conducted analyses that address the question of whether there are thresholds in the quality-outcome associations—whether the association between outcomes and quality may be stronger in higher than lower quality classrooms. These analyses replicate those conducted as part of the Child Care and Early Education Quality Features, Thresholds and Dosage and Child Outcomes: (Q-DOT) Study (Zaslow et al. 2010).⁷⁵ Using data from a number of large child care studies, including FACES 2006, those analyses identified threshold effects in the relationship between CLASS Instructional Support and both language and reading, and ECERS-R Teaching and Interactions and language (Burchinal et al. 2011b).

To test for threshold effects, we examined whether the relationship between quality and outcomes differed below and above certain cut-points used to define higher and lower quality classrooms.⁷⁶ Following the procedures used in Q-DOT, the cut-points were based on the developer's guidelines for defining moderately good quality and were adjusted when less than 25 percent of the classrooms were above or below that cut-off (as a reminder, scores for all of the quality measures range from 1 to 7). Thus, the cut-points were as follows:

- ECERS-R Teaching and Interactions: 4.5
- CLASS Emotional Support and Classroom Organization: 5.0
- CLASS Instructional Support and CLASS Language Modeling: 2.75

Like the first set of HLM analyses, the exploration of linear and nonlinear associations and threshold effects accounts for the clustering of children within classrooms and classrooms within programs. It also focuses on the alignment of quality measures and outcomes. However, to replicate analyses conducted as part of the Q-DOT study, these analyses departed from the procedures used in the earlier HLM analyses in four key ways. First, they included a smaller number of child and family control variables and omitted program and classroom control variables. Second, analyses were weighted at the child level only, not at the classroom or program level. Third, we used all observations in the analyses, rather than dropping observations with any missing data.⁷⁷ Fourth, as described below, the type of scores used

for some measures differed from those described earlier.

We estimated models of children's developmental status in the spring, controlling for their initial status measured in the fall.⁷⁸ We used all of the child cognitive and social-emotional outcome measures and included WJ Spelling and ECLS-B Math, as the latter two also were used in the Q-DOT analyses. We used standard scores⁷⁹ for the PPVT-4 and WJ measures in these analyses so that the coefficients could be interpreted as the change in standard score units for each one-point increase in the respective classroom quality variable. The EOWPVT and social-emotional outcomes were measured using raw scores,⁸⁰ and ECLS-B Math was measured using IRT scores.

The child/family-level control variables included gender, race/ethnicity, household language, maternal education, household poverty ratio, children's fall score on same outcome, and time interval between fall and spring assessments.⁸¹

We also present effect sizes, which can be interpreted as the standard deviation change in the child outcome associated with a standard deviation change in the respective independent variable. Consistent with the rest of this report, we considered a p-level less than .05 as indicating the association between quality and outcomes is statistically significant. As in the Q-DOT analyses, we noted marginal statistical significance of threshold effects at a p-level less than .10.

Children's Cognitive Outcomes

ECERS-R Teaching and Interactions. The results show a nonlinear association between Teaching and Interactions and WJ Spelling (linear ES = -.35, quadratic ES = .05). Generally, higher levels of Teaching and Interactions are associated with lower WJ Spelling scores; the negative association is weaker as the level of Teaching and Interactions increases. We did not identify any other associations or threshold effects for this measure of classroom quality.

CLASS Instructional Support. There is a positive association between Instructional Support and children's letter-word knowledge; higher levels of Instructional Support are associated with higher WJ

Letter Word Identification scores (ES = .06). The results indicate a threshold effect in the area of vocabulary: the positive association between Instructional Support and children's vocabulary, as measured by the PPVT-4, is marginally stronger in higher quality (ES = .09) than lower quality classrooms (ES = .01). We did not identify any other associations or threshold effects.

CLASS Language Modeling. Language Modeling is positively associated with children's letter-word knowledge; higher levels of Language Modeling are associated with higher WJ Letter Word Identification scores (ES = .06). There is also a nonlinear association between Language Modeling and children's math skills (linear ES = .14, quadratic ES = -.03). Higher levels of Language Modeling are associated with higher ECLS-B Math scores, but the positive association is weaker as the level of Language Modeling increases. We did not identify any other associations or threshold effects.

CLASS Classroom Organization. We did not identify any significant associations or threshold effects between CLASS Classroom Organization and children's cognitive outcomes.

Children's Social-Emotional Outcomes

CLASS Emotional Support. We did not identify any significant associations or threshold effects between

CLASS Emotional Support and children's social-emotional outcomes.

CLASS Positive Climate. The analysis of threshold effects shows that higher levels of Positive Climate are associated with fewer teacher-reported behavior problems, and the association is marginally stronger in higher quality (ES = -.10) than lower quality classrooms (ES = .02). We did not identify any other associations or threshold effects.

CLASS Classroom Organization. The results indicate positive linear associations between Classroom Organization and two of the social-emotional outcomes. Higher levels of Classroom Organization are associated with higher levels of social skills (ES = .06) and approaches to learning (ES = .01). We did not identify any other associations or threshold effects.

Summary. We found few associations between teacher characteristics and classroom quality and teacher attitudes. Shifting to child outcomes, Table 5 summarizes findings regarding the associations between quality and cognitive outcomes, and Table 6 presents the same information for social-emotional outcomes. Looking across the two tables suggests that there are limited linear associations and threshold effects between quality and outcomes in Head Start classrooms.

Table 5. Summary of Associations and Threshold Effects for Aligned Classroom Quality Measures and Child Cognitive Outcomes from Multivariate HLM Models

	PPVT-4	EOWPVT	WJ III: Letter-Word	WJ III: Applied Problems	WJ III: Spelling ^a	ECLS-B Math ^a
ECERS-R Teaching and Interactions					Nonlinear (1)	
CLASS Instructional Support	Threshold		Linear (2)			
CLASS Language Modeling			Linear (2)			Nonlinear (1)
CLASS Classroom Organization			Linear (1)			

Source: Fall 2009 and Spring 2010 FACES Direct Child Assessment, Fall 2009 FACES Parent Interview, Fall 2009 and Spring 2010 FACES Teacher Interview, Spring 2010 FACES Classroom Observation, and Fall 2009 FACES Center Director Interview.

Notes: Analyses are weighted to be representative of children who entered Head Start for the first time in fall 2009 and were still enrolled in spring 2010.

(1) Indicates statistically significant associations were identified with one of the two methods used to examine linear and non-linear associations.

(2) Indicates statistically significant associations were identified with two of the two methods used to examine linear and non-linear associations.

^a Only examined with one method.

Table 6. Summary of Associations and Threshold Effects for Aligned Classroom Quality Measures and Child Social-Emotional Outcomes from Multivariate HLM Models

	Social Skills	Problem Behaviors	ECLS-K Approaches to Learning	Pencil Tapping
CLASS Emotional Support				
CLASS Positive Climate		Threshold		Nonlinear (1)
CLASS Classroom Organization	Linear (2)		Linear (1)	

Source: Fall 2009 and Spring 2010 FACES Direct Child Assessment, Fall 2009 FACES Parent Interview, Fall 2009 and Spring 2010 FACES Teacher Interview, Spring 2010 FACES Classroom Observation, and Fall 2009 FACES Center Director Interview.

Notes: Analyses are weighted to be representative of children who entered Head Start for the first time in fall 2009 and were still enrolled in spring 2010.

(1) Indicates statistically significant associations were identified with one of the two methods used to examine linear and non-linear associations.

(2) Indicates statistically significant associations were identified with two of the two methods used to examine linear and non-linear associations.

As a reminder, we used two approaches to testing linear and nonlinear associations for child outcomes (with the exception of WJ Spelling and ECLS-B Math). For those outcomes in which we used both approaches, we found that only two child outcome measures consistently related to classroom quality (that is, there were statistically significant associations using both approaches):

- The WJ Letter Word has statistically significant, positive linear associations with both CLASS Instructional Support and Language Modeling.
- Teacher-reported social skills has a statistically significant, positive linear association with CLASS Classroom Organization.

We identified a threshold effect in only two associations: the PPVT-4 with CLASS Instructional Support and problem behaviors with CLASS Positive Climate. In both cases, associations are marginally stronger in higher quality classrooms.

In general, these findings are consistent with the broader research literature that examines linkages between child outcomes and classroom quality: we identified a handful of modest, linear associations but also found evidence of nonlinear associations for certain quality-outcome pairs, including threshold effects.⁸² Note that one of those threshold effects—in the association between language and CLASS Instructional Support—is consistent with findings from the Q-DOT study (Burchinal et al. 2011b). However, we find fewer associations than in recent meta-analyses.

REFERENCES

- Administration for Children and Families. “Head Start FACES 2000: A Whole-Child Perspective on Program Performance: Fourth Progress Report.” Washington, DC: U.S. Department of Health and Human Services, 2003.
- Aikens, N., E. Moiduddin, L. Tarullo, and J. West. “Data Tables for Child Outcomes and Classroom Quality in FACES 2009 Report.” OPRE Report 2012-37b. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, September 2012.
- Aikens, N, L. Tarullo, L. Hulse, C. Ross, J. West, and Y. Xue. “A Year in Head Start: Children, Families and Programs.” Report submitted to the U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research and Evaluation. Washington, DC: Mathematica Policy Research, October 2010.
- American Academy of Pediatrics, Committee on Public Education. “Children, Adolescents, and Television.” *Pediatrics*, vol. 107, no. 2, February 2001, pp. 423-426.
- Anderson, S.E., and R.C. Whitaker. “Prevalence of Obesity Among U.S. Preschool Children in Different Racial and Ethnic Groups.” *Archives of Pediatric and Adolescent Medicine*, vol. 163, no. 4, 2009, pp. 344–348.

- Atkins-Burnett, S., and N. Aikens. "Sweet Dreams: Linkages Between Sleep Quality and Quantity and Children's Developmental Outcomes in a Head Start Sample." Paper presented at the Society for Research in Child Development, Biennial Meeting, Montreal, March 2011.
- Beals, D. E. and C. E. Snow. "Mealtime Talk That Supports Literacy Development." *New Directions for Child and Adolescent Development*, vol. 111, 2006, pp. 51-66.
- Blair, C. "School Readiness: Integrating Cognition and Emotion in a Neurobiological Conceptualization of Children's Functioning at School Entry." *American Psychologist*, vol. 57, 2002, pp. 111–127.
- Brownell, R. "Expressive One-Word Picture Vocabulary Test." San Antonio, TX: Harcourt Assessment, Inc., 2000.
- Burchinal, M. R., K. Kainz, and Y. Cai. "How Well Are our Measures of Quality Predicting to Child Outcomes: A Meta-Analysis and Coordinated Analyses of Data from Large Scale Studies of Early Childhood Settings." In *Measuring Quality in Early Childhood Settings*, edited by M. Zaslow, I. Martinez-Beck, K. Tout, and T. Halle. Baltimore: Brookes Publishing, 2011a.
- Burchinal, M.R., N. Vandergrift, R. Pianta, and A.J. Mashburn. "Threshold Analysis of Association Between Child Care Quality and Child Outcomes for Low-Income Children in Prekindergarten Programs." *Early Childhood Research Quarterly*, vol. 25, no. 2, 2010, pp. 166–176.
- Burchinal, M.R., Y. Xue, and A.J. Mashburn. "Secondary Data Analysis Looking for Thresholds in Child Care Quality." Paper presented at the Society for Research in Child Development Biennial Meeting, Montreal, April 2011b.
- 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 Instructional Practices." *Early Childhood Research Quarterly*, vol. 5, 1990, pp. 407–423.
- Chernoff, J., K. Denton Flanagan, C. McPhee, and J. Park. "Preschool: First Findings from the Preschool Follow-Up of the Early Childhood Longitudinal Study, Birth Cohort (ECLS–B)." NCES 2008–025. Washington, DC: U.S. Department of Education, Institute of Education Sciences, NCES, 2007.
- Clifford, R.M., O. Barbarin, F. Chang, D.M. Early, D. Bryant, C. Howes, M. Burchinal, and R. Pianta. "What Is Pre-Kindergarten? Characteristics of Public Pre-Kindergarten Programs." *Applied Developmental Science*, vol. 9, 2005, pp. 126–143.
- Croninger, R.G., and V.E. Lee. "Social Capital and Dropping Out of High School: Benefits to At-Risk Students of Teachers' Support and Guidance." *Teachers College Record*, vol. 103, no. 4, 2001, pp. 548–541.
- Diamond, A., and C. Taylor. "Development of an Aspect of Executive Control: Development of the Abilities to Remember What I Said and to "Do as I Say, Not as I Do." *Developmental Psychobiology*, vol. 29, 1996, pp. 315–334.
- Downey, D.B., P.T. von Hippel, and B. Broh. "Are Schools the Great Equalizer? Cognitive Inequality During the Summer Months and the School Year." *American Sociological Review*, vol. 69, no. 5, 2004, pp. 613–635.
- Downey, G. and J.C. Coyne. Children of Depressed Parents: An Integrated Review." *Psychological Bulletin*, vol. 108, pp. 50–76, 1990.
- Duncan, S.E., and E.A. DeAvila. *Preschool Language Assessment Survey 2000 Examiner's Manual: English Forms C and D*. Monterey, CA: CTB/McGraw-Hill, 1998.
- Dunn, L.M., D.E. Lugo, E.R. Padilla, and L.M. Dunn. *Test de Vocabulario en Imagenes Peabody*. Circle Pines, MN: American Guidance Service, 1986.
- Dunn, L.M., and D.M. Dunn. *Peabody Picture Vocabulary Test. Fourth Edition*. Circle Pines, MS: American Guidance Service, 2006.

- Entwisle, D.R., K.L. Alexander, and L.S. Olson. *Children, Schools, and Inequality*. Boulder, CO: Westview Press, 1997.
- Gaylor, E., X. Wei, and M.M. Burnham. "Associations Between Nighttime Sleep and Developmental Outcomes in a Nationally Representative Sample of Preschool-Age Children." Paper presented at 24th Annual Meeting of the Associated Professional Sleep Societies, San Antonio, Texas, June 5–9, 2010.
- Gresham, F.M., and S.N. Elliot. *Social Skills Rating System*. Circle Pines, MN: American Guidance Service, 1990.
- Harms, T. and R. Clifford. *The Early Childhood Environment Rating Scale (ECERS)*. New York NY: Teachers College Press, 1980.
- Harms, T., R. Clifford, and D. Cryer. *Early Childhood Environment Rating Scale—Revised (ECERS–R)*. New York, NY: Teachers College Press, 1998.
- Hofferth, S.L. and J.F. Sandberg. "How American Children Spend Their Time." *Journal of Marriage and the Family*, vol. 63, no. 2, 2001, pp. 295-308.
- Howes, C., M. Burchinal, R. Pianta, D. Bryant, D. Early, and R. Clifford. "Ready to Learn? Children's Pre-Academic Achievement in Pre-Kindergarten Programs." *Early Childhood Research Quarterly*, vol. 23, 2008, pp. 27–50.
- Hulseley, L., N. Aikens, A. Kopack, J. West, E. Moiduddin, and L. Tarullo. "Head Start Children, Families, and Programs: Present and Past Data from FACES." OPRE Report 2011-33a. Washington, DC: U.S. Department of Health and Human Services, December 2011.
- Iruka, I.U., and P.R. Carver. "Initial Results from the 2005 NHES Early Childhood Program Participation Survey." NCEES 2006-075. Washington, DC: National Center for Education Statistics, 2006.
- Kaiser Commission on Medicaid and the Uninsured. "Health Coverage of Children: The Role of Medicaid and CHIP." Washington, DC: Henry J. Kaiser Family Foundation, 2011. Available at <http://www.kff.org/medicaid/upload/7698-05.pdf>. Accessed September 10, 2011.
- Koplan, J.P., C.T. Liverman, and V.A. Kraak, Editors. *Preventing Childhood Obesity: Health in the Balance*. Washington, DC: Committee on Prevention of Obesity in Children and Youth, National Academies Press, 2005.
- Lambert, R. "Evaluating Management Climate in Head Start Programs: The Measurement Properties of the Policy and Program Management Inventory." *NHSA Dialog: A Research-to-Practice Journal for the Early Intervention Field*, vol. 6, no. 1, 2002, pp. 37–52.
- Lambert, R., M. Abbott-Shim, and C. Oxford-Wright. "Policy and Program Management Inventory, Teacher Version." Charlotte, NC: The Center for Educational Evaluation and Measurement, 1999.
- McDermott, P.A., L.F. Green, J.M. Francis, and D.H. Stott. *Preschool Learning Behaviors Scale*. Philadelphia: Edumetric and Clinical Science, 2000.
- Meadows, S.O., S. McLanahan, and J. Brooks-Gunn. "Parental Depression and Anxiety and Early Childhood Behavior Problems Across Family Types." *Journal of Marriage and Family*, vol. 69, pp. 1162–1177, 2007.
- National Child Care Information and Technical Assistance Center. "QRS Planning Tool: Standards for Early Childhood Education Programs, May 2008." Available at <http://nccic.acf.hhs.gov/poptopics/qrs-pt-ece.html>. Accessed September 15, 2011.
- National Research Council. *Early Childhood Assessment: Why, What, and How*, edited by C.E. Snow and S.B. Van Hemel. Washington, DC: The National Academies Press, 2008.
- NICHD Early Child Care Research Network and Duncan, G. "Does Quality of Child Care Affect Child Outcomes at Age 4 ½?" *Developmental Psychology*, vol. 39, no. 3, 2003, pp. 451–469.
- Pallas, A.M., G. Natriello, and E.L. McDill. "The Changing Nature of the Disadvantaged

- Population: Current Dimensions and Future Trends. *Educational Researcher*, vol. 18, no. 5, 1989, pp. 16–22.
- Peterson, J., and N. Zill. “Marital Disruption, Parent-Child Relationships, and Behavior Problems in Children.” *Journal of Marriage and the Family*, vol. 48, 1986, pp. 295–307.
- Pianta, R., K. LaParo, and B. Hamre. *The Classroom Assessment Scoring System Pre-K Manual*. Charlottesville, VA: University of Virginia, 2008.
- Rathbun, A., and J. West. *From Kindergarten Through Third Grade: Children’s Beginning School Experiences*. NCES 2004-007. Washington, DC: National Center for Education Statistics, 2004.
- Roid, G.H., and L.J. Miller. *Examiners Manual: Leiter International Performance Scale-Revised*. Chicago: Stoelting Co., 1997.
- Ross, C., J. Mirowsky, and J. Huber. “Center for Epidemiological Studies-Depression Scale Short Form.” *American Sociological Review*, vol. 48, no. 6, 1983, pp. 809–823.
- Smith-Donald, R., C. Raver, T. Hayes, and B. Richardson. “Preliminary Construct and Concurrent Validity of the Preschool Self-Regulation Assessment (PSRA) for Field-Based Research.” *Early Childhood Research Quarterly*, vol. 22, 2007, pp. 173–187.
- Tarullo, L., N. Aikens, E. Moiduddin, and J. West. “A Second Year in Head Start: Characteristics and Outcomes of Children Who Entered the Program at Age Three.” Washington, DC: U.S. Department of Health and Human Services, December 2010.
- U.S. Department of Education, National Center for Education Statistics. “Early Childhood Longitudinal Study—Kindergarten Class of 1998–99 (ECLS–K), Psychometric Report for Kindergarten Through First Grade.” NCES 2002–05. Washington, DC: U.S. Department of Education, Institute of Education Sciences, NCES, 2002.
- West, J., L. Tarullo, N. Aikens, L. Malone, and B.L. Carlson. “FACES 2009 Study Design.” OPRE Report 2011-9. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, June 2011.
- Woodcock, R.W., A.F. Muñoz-Sandoval, K. McGrew, N. Mather, and F. Schrank. *Batería III Woodcock-Muñoz*. Itasca, IL: Riverside Publishing, 2004.
- Woodcock, R.W., K. McGrew, and N. Mather. *Woodcock-Johnson III Tests of Achievement*. Itasca, IL: Riverside Publishing, 2001.
- Zaslow, M., R. Anderson, Z. Redd, J. Wessel, L. Tarullo, and M. Burchinal. “Quality Dosage, Thresholds, and Features in Early Childhood Settings: A Review of the Literature.” Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research and Evaluation, August 2010.
- Zill, N., A. Sorongon, K. Kim, C. Clark, and M. Woolverton. “FACES 2003 Research Brief: Children’s Outcomes and Program Quality in Head Start.” Washington, DC: U.S. Department of Health and Human Services, 2006.
- Zill, N., and J. West. *Findings from the Condition of Education 2000: Entering Kindergarten Children*. NCES 2001-035. Washington, DC: National Center for Education Statistics, 2001.

NOTES

¹ For detailed information on the FACES 2009 study design and measures, see West et al. 2011.

² Migrant and Seasonal Head Start (MSHS) programs, American Indian and Alaska Native (AI/AN) programs, programs in Puerto Rico and other U.S. territories, and programs not directly providing services to 3-, 4-, and 5-year-olds (such as Early Head Start) were excluded from the frame. The Office of Head Start provided information about any defunded (or soon-to-be defunded) programs before

sampling and these programs were then deleted from the sample frame.

³ Three of the 65 programs originally sampled were determined to be ineligible because we learned they were under provisional management or otherwise in financial jeopardy. In addition, two eligible programs declined to participate.

⁴ Children who were no longer enrolled in the program where they were sampled in fall 2009 and who were not enrolled in one of the other FACES 2009 programs were not included in the spring 2010 data collection.

⁵ The first visits to Head Start programs were in March 2010; however, parent interviews by telephone began in late February of that year.

⁶ Parents who did not have telephones, preferred not be called at home, or did not want to use their own cell phone minutes were offered the option of completing the interview by phone at their child's Head Start center or in a face-to-face interview with a member of the data collection staff. Only 2 percent of parent interviews were completed in person.

⁷ This total represents 81 percent of the children who were sampled and eligible for the fall 2009 baseline data collection.

⁸ These are all weighted marginal response rates, not accounting for prior stages of sampling and participation. The cumulative weighted response rates, which take into account the response rate for prior stages of the sample (such as, program, center, and child response rates), as well as fall 2009 consent rates, are by definition lower. The cumulative child response rate through spring 2010 is 82 percent. The corresponding response rates associated with completing the child assessments, parent interviews, and teacher ratings in spring 2010 are 78 percent, 71 percent, and 79 percent, respectively. At the teacher level, among participating classes, the marginal weighted response rate for the teacher interview was 99 percent. At the child level, among children whose parents gave consent, the rate for child assessments was 95 percent, the rate for parent interviews was 86

percent, and the rate for teacher-child reports was 94 percent.

⁹ The cumulative teacher interview response rate is 92 percent.

¹⁰ A total of 391 of 482 eligible classrooms were sampled for the classroom observations. The cumulative weighted response rate for the observations, which takes into account nonresponse at the program level, was 87 percent. To be eligible for observation, the classroom had to meet three criteria: (1) be in a center-based program (home-based services were not observed); (2) be one of the originally sampled classrooms (classrooms that children moved to in the spring were not eligible); and (3) have at least one sampled, eligible child whose parents gave consent.

¹¹ The screening process and cognitive assessment measures are described in the section on child outcomes.

¹² Simon Says, a subtest from the Preschool Language Assessment Survey 2000 (*preLAS 2000*; Duncan and DeAvila 1998), was used as a warm-up activity at that start of the assessment for this group of children.

¹³ The preferred respondent for the spring interview was the child's biological mother or the fall 2009 respondent. Ninety-five percent of the spring interviews were completed by the same respondent who had been interviewed in the fall (and 87 percent were the child's biological mother). For 4 percent of the children, the first parent interview was completed in spring; 96 percent completed the first parent interview in fall.

¹⁴ In spring 2010, 80 percent of TCRs (and 76 percent of all eligible cases) were completed using the web-based instrument with the balance completed using paper forms.

¹⁵ FACES 2006 used the full ECERS-R, the Instructional Support scale from CLASS, and the Arnett Caregiver Interaction Scale.

¹⁶ See Aikens et al. 2012 for the statistics found in this report. That volume includes a set of data tables designed to accompany this report.

¹⁷ Weights are used to compensate for the differential probabilities of selection at the sampling stage (for example, we selected programs, centers, and classrooms with probability proportional to size; and we selected a fixed number of children per classroom out of a variable number of eligible children) and to adjust for changes in children's eligibility status and the effects of nonresponse.

¹⁸ Information on the characteristics of children and their families in fall 2009 can be found in "Head Start Children, Families, and Programs: Present and Past Data from FACES" (Hulsey et al. 2011).

¹⁹ All references to African American refer to African American, Non-Hispanic.

²⁰ All references to mothers, fathers, or parents include both biological and adoptive parents.

²¹ Croninger and Lee 2001; Pallas et al. 1989; Rathbun and West 2004; Zill and West 2001.

²² Downey et al. 2004; Rathbun and West 2004; West et al. 2001.

²³ Each of these factors is based on parent reports in fall 2009 and given equal weight in the risk index.

²⁴ The nutritional guidelines were determined a priori, based on conversations with a member of an Office of Head Start expert panel. Guidelines call for drinking milk, eating fruit and vegetables at least twice a day; drinking no soda, sports drinks, or non-100 percent juice drinks; eating no fast food; and eating sweets and salty snacks less than once a day.

²⁵ Government insurance includes Medicaid, State Children's Health Insurance Program (SCHIP), military health care, Indian Health Service, and other government insurance programs. Previous FACES reports (Aikens et al. 2010, Tarullo et al. 2010) have presented percentages of children with SCHIP, Medicaid, and other government programs separately.

²⁶ <http://www.sleepfoundation.org/article/sleep-topics/children-and-sleep>. Accessed October 5, 2010.

²⁷ With the exception of vocabulary measures, we are unable to provide information on changes in the skills of children who changed their language of assessment between fall and spring (n = 247), as these children receive different assessment measures at each wave.

²⁸ The Expressive One-Word Picture Vocabulary Test: English and the Spanish-Bilingual Edition (EOWPVT; EOWPVT-SBE; Brownell 2000) measure the expressive vocabulary of children from English- and Spanish-speaking households, respectively. The EOWPVT provides a measure of children's expressive vocabulary relative to English-speaking peers nationally, while the EOWPVT-SBE reflects children's vocabulary skills relative to Spanish-bilingual and Spanish-dominant peers. The EOWPVT-SBE allows for conceptual scoring (that is, it provides prompts for both English and Spanish and accepts responses in each language and various Spanish dialects). All children take the same items but are scored as correct when they accurately identify an object, whether they label it in English or Spanish, thereby providing a picture of children's bilingual expressive vocabulary. In FACES, the EOWPVT-SBE was used with children whose primary home language is Spanish, while the EOWPVT was used with all other children.

²⁹ All children, regardless of home language or performance on the *pre*LAS, received the English receptive vocabulary measure, the PPVT-4, and the expressive vocabulary measure, the EOWPVT or EOWPVT-SBE. The TVIP is the Spanish-language version of the PPVT-4 and was used with children whose primary home language was Spanish, regardless of performance on the *pre*LAS. Thus, children whose parents spoke Spanish to them at home received the receptive vocabulary component of the battery in English (PPVT-4) as well as in Spanish (TVIP). They also received the Spanish-bilingual version of the EOWPVT. All other children received the PPVT-4 and the EOWPVT.

³⁰ The English assessment used the WJ III subtests and the Spanish assessment used the Bateria III WM subtests.

³¹ FACES used 23 mathematics items from the ECLS–B in fall and spring of the Head Start year(s) and an additional 7 items from the ECLS–K in kindergarten.

³² Because this measure requires a higher skill set for children and is administered to children passing a threshold on the letter-word subtest, only a subset of children receives it. In addition, it is only available in English. Scores are calibrated for the full English-assessed sample based on the subsample of children who are administered the full set of items. We report these scores in the set of data tables Child Outcomes and Classroom Quality in FACES 2009 (Aikens et al. 2012).

³³ For the direct assessment, home language was based on information provided on parent consent forms.

³⁴ As noted previously, regardless of performance on the language screener, children from homes where Spanish was primarily spoken were also administered the TVIP as a measure of their receptive Spanish vocabulary. Thus, these children received the receptive vocabulary component of the battery in English (PPVT-4) and Spanish (TVIP). They also received the Spanish-bilingual version of the EOWPVT.

³⁵ Some children were administered large sections of (or the majority of) the cognitive assessments in Spanish (or not assessed at all) in fall 2009 and then were assessed in English in spring 2010. Similarly, some children were unable to achieve a basal on the PPVT-4 in the fall but were able to by spring. Data in this section reflect the performance of children assessed in English in both fall 2009 and spring 2010. In addition, mean scores are only reported for those with valid scores at both occasions (for example, those who established a basal on the PPVT-4 at both waves). See Aikens et al. 2012 for the mean scores for all children assessed in spring 2010, regardless of language of assessment, child performance, or availability of valid scores in the fall. In this set of tables, children's mean spring 2010 scores are slightly lower (for example, 1 to 2 standard score points lower). Variability in children's scores is comparable for both sets of scores.

³⁶ Cross-cohort comparisons only include children in FACES 2003, 2006, and 2009, because in earlier cohorts, 3-year-old children were not assessed in all areas. In addition, WJ scores in FACES 2000 were drawn from the Woodcock-Johnson Psycho-Educational Battery-Revised edition (WJ-R). WJ scores in FACES 2003 were drawn from a hybrid version of the WJ–R and the WJ III. This allows for estimation of children's scores based on either the norms for the WJ–R or WJ III. FACES 2006 and 2009 used the WJ III.

³⁷ All analyses of the WJ III scores are based on original calculations using FACES 2003 data.

³⁸ PPVT scores in FACES 2003 were drawn from the Peabody Picture Vocabulary Test-Third edition (PPVT-III). FACES 2006 and 2009 used PPVT-4.

³⁹ Reported cross-cohort comparisons are drawn from previous FACES reporting (Aikens et al. 2010).

⁴⁰ This score is a proficiency probability score and indicates the probability that a child would have passed the proficiency level. It can be interpreted as the percentage of the population that has "mastered" this skill or skill set (for example, $.30 \times 100 = 30$ percent of Head Start children are able to demonstrate these skills at the start of the program year). These scores can take on any value from zero to one.

⁴¹ The ECLS–B preschool wave was intended to assess children in the fall, when most children would be about 48 through 57 months of age. However, the age at time of testing in the ECLS–B preschool wave ranged from approximately 3 years, 8 months to 5 years, 5 months (Chernoff et al. 2007). On average the FACES children in the fall 2009 round were assessed earlier in the program year than the ECLS–B sample, and their ages ranged from approximately 2 years, 7 months to 5 years, 8 months. Conversely, FACES children in the spring 2010 round were assessed later, with ages ranging from 3 years, 3 months to 6 years, 3 months.

⁴² The number of family risks is based on three family characteristics: whether the child resides in a single parent household, whether household income

is below the poverty threshold, and whether the mother has less than a high school diploma.

⁴³ Although scores on the WM III subtests were scaled to be comparable with the WJ III, early items on the Letter-Word Identification subtests suggest differences in the demands required of children. Specifically, children who are administered the WM III are required to provide verbal responses earlier in the assessment than those administered the WJ III, suggesting less “warm up” time and greater demands of these children.

⁴⁴ For children from homes where Spanish is primarily spoken, the expressive vocabulary assessment is conceptually scored. For these children, standard scores using both the EOWPVT-SBE and the EOWPVT norms are created. EOWPVT standard scores provide a measure of children's English expressive vocabulary relative to young children in the U.S., while the EOWPVT-SBE standard scores reflect children's bilingual (English and Spanish) vocabulary skills relative to young Hispanic children nationally.

⁴⁵ For the first time in FACES, teachers in FACES 2009 rated each child on the six items that comprise the Approaches to Learning scale from the ECLS–K (U.S. Department of Education 2002). Earlier FACES cohorts used the Preschool Learning Behavior Scale (PLBS) (McDermott et al. 2000) to assess children's approaches to learning.

⁴⁶ Similar declines in hyperactive behavior were not found in FACES 2003. Reported problem behavior and hyperactive scores in FACES 2009 were constructed using different items from those reported in cohorts prior to FACES 2006.

⁴⁷ Significant fall–spring declines were found in teacher-rated withdrawn behavior in FACES 2003.

⁴⁸ Body mass index (BMI) is the ratio of an individual's weight to height (weight in kilograms divided by squared height in meters) and can be used as an indicator of overweight and obese status. Calculation of BMI is specific to gender and age.

⁴⁹ The Centers for Disease Control and Prevention sets the criterion of “overweight” as being when the child's BMI score is from the 85th to 94th percentile for their age and gender, and of “obese” as being when the child's BMI is at or above the 95th percentile.

⁵⁰ By 2013, half of all Head Start teachers must have at least a B.A. degree in early childhood or a B.A. that incorporates relevant coursework and experience (Improving Head Start for School Readiness Act of 2007, P.L. 110-134).

⁵¹ “Other” types of curricula are those that do not have printed materials available for use in implementation or information on the goals related to the specific curriculum.

⁵² Concept Development refers to teachers' use of instructional discussions and activities to promote higher-order thinking in contrast to rote instruction. Quality of Feedback refers to the degree to which the teacher provides feedback that expands learning and understanding and encourages participation. Language Modeling refers to the quality and amount of teachers' use of language-stimulating and language-facilitation techniques.

⁵³ Classroom Climate addresses positive (warm, supportive, respectful) and negative (irritation, crying, threats to control behavior) aspects of the classroom climate. Teacher Sensitivity refers to responsiveness of teachers to children and support for exploration, learning, and social situations. Regard for Student Perspectives refers to teachers' interest in children's perspectives and support for their autonomy.

⁵⁴ Behavior Management refers to teachers' use of proactive, anticipatory techniques for managing behavior. Productive Use of Time addresses how the day is scheduled and whether time is used efficiently. Instructional Learning Formats refers to modes of instruction and engagement of children.

⁵⁵ The Multi-State Study of Prekindergarten (Clifford et al. 2005) identifies 21 items that represent the key dimensions of quality tapped by the 43 items on the full ECERS-R: Provisions for Learning and

Teaching and Interactions. These items are used in FACES 2009.

⁵⁶ Classroom observations were completed by observers trained and certified after meeting reliability standards showing proficiency to administer each instrument. Reliability was defined as being within one point of the gold standard observer on the scale or dimension score at least 80 percent of the time. Sixteen of the 17 classroom observer/gold standard observer pairs were in agreement at least 80 percent of the time on the ECERS-R, and 100 percent of the observer and gold standard scores were within one point of each other on the CLASS. To minimize observer drift, one quality assurance visit (that is, a paired observation) was conducted during the field period. If there was a discrepancy between the observer and the gold standard, the gold standard score was used as the final score.

⁵⁷ The average adult-child ratio and group size identified in the classroom observations differ from those reported by teachers. This is likely due to the fact that observers report the number of children and adults in a classroom on a particular day (averaged across four points during the observation), while teachers are asked to report how many children are *enrolled* in the class and how many adults are *usually* with the class.

⁵⁸ Although an alternative specification for classroom quality, these scores allow us to compare with other recent studies of classroom quality that have reported scores for the short form of ECERS-R (for example, the National Center for Early Development and Learning Study of Prekindergarten).

⁵⁹ Reported scores for all cohorts are weighted. All analyses are based on original calculations using FACES 2000, 2003, 2006, and 2009 data.

⁶⁰ The CLASS Instructional Support domain was used as part of the observation protocol for the first time in FACES 2006. The measure was added to provide information on the quality of the instructional environment. Comparative data with earlier cohorts is not available. In spring 2008, the classroom observation protocol was changed in order to study the feasibility of training, certification, field

administration, and quality assurance procedures on the full CLASS in a large sample of programs. All three domains from the CLASS were assessed, while the ECERS-R and Arnett were not used. A sample of 147 classes attended by FACES 2006 children (4-year-olds attending a second year of Head Start) was observed. The spring observation sample was not designed to support national estimates of Head Start classrooms. Instead, the goal was to learn as much as possible about what is required to prepare for and conduct the full CLASS in Head Start classrooms. Analyses of these data suggest higher average scores on the CLASS Instructional Support domain are seen when the domain is used with the full scale (as in spring 2008) as compared to when used in isolation in spring 2007. The findings for spring 2010 are consistent with this pattern.

⁶¹ Although the CLASS Instructional Support domain was used as part of the observation protocol for the first time in FACES 2006, the Emotional Support and Classroom Organization domains were not added until FACES 2009. Thus, comparative data with earlier cohorts is not available.

⁶² To calculate the Emotional Support domain score, the Negative Climate dimension is reverse coded.

⁶³ Of the 60 programs that participated in the study in fall 2009, data for 59 (98 percent) were included in the analyses. Of the 486 classrooms that participated in the study in fall 2009, data for 327 (67 percent) were included in the analyses.

⁶⁴ For each outcome, we estimated two models in the analysis. In Model 1, we included all teacher/classroom characteristics at level 1. In Model 2 we added program-level characteristics. We present findings regarding the association of teacher characteristics with classroom quality and teacher attitudes from Model 2.

⁶⁵ Teachers responded to questions regarding perceptions of management support on a 5-point scale ranging from 1 ("strongly disagree") to 5 ("strongly agree"). To aid in interpretation of this effect, we reran the analysis using a version of the management support variable that had been z-scored.

A one standard deviation increase in perceptions of management support is associated with a .19 standard deviation decline in CLASS Classroom Organization and a .23 standard deviation increase in teacher satisfaction with teaching as a career.

⁶⁶ The analysis of threshold effects is essentially another approach to determining whether there is a nonlinear relationship between quality and outcomes. If an association is linear, this indicates that a change in quality is associated with a change in child outcomes, and the magnitude of the change in child outcomes is the same regardless of the level of quality in the classroom. A nonlinear association would indicate that the change in child outcomes might differ depending on the level of quality in the classroom. In our first approach to assessing the relationship between quality and outcomes in FACES 2009, we included a quadratic term in the analyses to determine if the shape of the relationship between quality and outcomes differs across the range of quality. In our second approach, we specified cut-points to test whether the relationship between quality and outcomes differs in higher- and lower-quality classrooms.

⁶⁷ Studies in which analyses identified threshold effects (Burchinal et al. 2011) include the National Institute of Child Health and Human Development (NICHD) Study of Early Child Care and Youth Development (NICHD Early Child Care Research Network and Duncan 2003); the National Center for Early Development and Learning (NCEDL) Prekindergarten Evaluation (Howes et al. 2008); and FACES 1997 (Administration for Children and Families 2003). Threshold effects were identified for language, mathematics, and reading.

⁶⁸ Of the 60 programs that participated in the study in fall 2009, data for 59 (98 percent) were included in the analyses. Of the 486 classrooms that participated in the study in fall 2009, data for between 318 (65 percent) and 327 (67 percent) were included in the analyses, depending on the outcome. The one exception is executive functioning as measured by the pencil-tapping task, which was administered only to 4-year-olds in fall 2009; for this measure, 261 classrooms (54 percent) were included in the analysis. Finally, of the 3,349 children who participated in the

study in fall 2009, data for 1,354 (40 percent) to 1,936 (54 percent) were included in the analysis; for the pencil-tapping task, data for 69 percent of the fall sample of 4-year-olds were included in the analyses.

⁶⁹ For Spanish-speaking children who changed language of assessment between fall and spring, we used their fall WM assessment scores when predicting corresponding spring WJ assessment scores.

⁷⁰ Unlike in the descriptive reporting, children who did not establish a basal on the PPVT-4 were included in the appropriate models. By including in the analyses children who did not establish a basal in the fall, we likely overestimate children's progress on this measure.

⁷¹ The PPVT-4 refers to W-scores as GSV scores. W-scores were not used for analyses focused on the EOWPVT; instead, we used raw scores. The PPVT and WJ/WM W-scores are on different scales, as are the EOWPVT raw scores.

⁷² Time interval between assessments was included only in models focused on the cognitive outcomes because the cognitive assessments were developed to account for child age in scoring procedures.

⁷³ We estimated a series of models in the analysis. In Model 1, we included child/family characteristics in level 1. In Model 2, we added classroom quality and teacher/classroom characteristics at level 2. In Model 3, we added program characteristics at level 3.

⁷⁴ We ran multiple models for each child outcome to test associations with multiple aspects of classroom quality. The magnitude of the association between teacher education and children's executive functioning changes depending on which aspect of classroom quality is also included in the model.

⁷⁵ Q-DOT is a two-and-a-half year design project funded by the Office of Planning, Research and Evaluation (OPRE) at the Administration for Children and Families (ACF) that is examining associations between the quality of early childhood settings and child outcomes by asking whether certain thresholds of quality or dosage need to be met, or particular aspects of quality need to be present, before linkages

are apparent. Q-DOT includes secondary analyses with specific data sets, including both program- and community-based settings that examine the presence of thresholds in the relationships between quality and outcomes.

⁷⁶ To test for threshold effects, we conducted spline regressions that included separate estimates of the slopes between quality and outcomes in classrooms with higher quality (quality scores above the cut-points) and lower quality (quality scores below the cut-points). We also tested whether the slopes in the lower and higher range of quality were significantly different from one another.

⁷⁷ We used multiple imputation to handle missing data.

⁷⁸ For Spanish-speaking children who changed language of assessment between fall and spring, we used their fall WM assessment scores when predicting to corresponding spring WJ assessment scores.

⁷⁹ As a reminder, standard scores provide information on children's performance relative to same-age peers, which is different from the information on absolute performance provided by IRT-based scores, such as the *W* scores used in the HLM analyses.

⁸⁰ For the Q-DOT threshold analyses, standard scores were not calculated for the EOWPVT, and raw scores were used. We followed the same procedures in this analysis.

⁸¹ Time interval between assessments was included only in models focused on the cognitive outcomes because the cognitive assessments were developed to account for child age in scoring procedures.

⁸² Note that in FACES 2006, analyses examined associations between children's cognitive outcomes (receptive vocabulary, letter-word knowledge, and mathematics ability) and ECERS-R Teaching and Interactions, CLASS Instructional Support, and CLASS Language Modeling. Those analyses identified consistent associations between classroom quality and receptive vocabulary (PPVT-4) but not with letter-word knowledge. There was one negative association with mathematics (Aikens et al. 2010).

FACES 2009 COPYRIGHT PERMISSIONS

Peabody Picture Vocabulary Test, Fourth Edition (PPVT-4). Copyright © 2007, Wascana Limited Partnership. Reproduced with permission of the publisher NCS Pearson, Inc. All rights reserved.

Test de Vocabulario en Imágenes Peabody (TVIP). Copyright © 1987, Dunn Educational Services, Inc. Reproduced with permission of the publisher NCS Pearson, Inc. All rights reserved.

Social Skills Rating System (SSRS). Copyright © 1990, NCS Pearson, Inc. This adaptation Copyright © 2006. Reproduced with permission of the publisher. All rights reserved.

Woodcock-Johnson® III (WJ III®), WJ III® Tests of Achievement. Copyright © 2001, 2007, The Riverside Publishing Company. Reproduced with permission of the publisher. All rights reserved.

No part of this work may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying and recording or by any information storage or retrieval system without the proper written permission of The Riverside Publishing Company unless such copying is expressly permitted by federal copyright law. Address inquiries to Contracts and Permissions Department, The Riverside Publishing Company, 3800 Golf Road, Rolling Meadows, Illinois 60008-4015.

Batería III Woodcock-Muñoz ®. Copyright © 2004, 2007, The Riverside Publishing Company. Reproduced with permission of the publisher. All rights reserved.

No part of this work may be reproduced or transmitted in any form or by any means, electronic or mechanical, including photocopying and recording or by any information storage or retrieval system without the proper written permission of The Riverside Publishing Company unless such copying is expressly permitted by federal copyright law. Address inquiries to Contracts and Permissions Department, The Riverside Publishing Company, 425 Spring Lake Drive, Itasca, Illinois 60143-2079.

preLAS 2000, by Sharon E. Duncan, Ph.D., and Edward A. DeAvila, Ph.D. Copyright © 1998 CTB/McGraw-Hill LLC, a subsidiary of The McGraw-Hill Companies, Inc. Reproduced by permission of CTB/McGraw-Hill LLC.

Leiter International Performance Scale-Revised Examiner Ratings. Copyright © 1997, 2002 Stoelting Co., 620 Wheat Lane, Wood Dale, IL 60191. All rights reserved.

Classroom Assessment Scoring System™ (CLASS™) by Robert C. Pianta, Karen M La Paro, and Bridget K. Hamre. Copyright © 2008 by Paul H. Brooks Publishing Co. Used with permission of publisher.

Early Childhood Environment Rating Scale, Revised Edition by Thelma Harms, Richard M. Clifford, and Debby Cryer. Copyright © 2005. New York: Teacher College Press. Reproduced with permission from the authors and the publisher. This copyrighted material may not be sold, copied, or distributed for any reason. All rights reserved.

Expressive One-Word Picture Vocabulary Test (EOWPVT). Copyright © 2000, Academic Therapy Publications, 20 Commercial Boulevard, Novato, CA, 94949-6191. All rights reserved. Reproduced by permission of Academic Therapy Publications.

Expressive One-Word Picture Vocabulary Test—Spanish-Bilingual Edition. Copyright © 2001, Academic Therapy Publications, 20 Commercial Boulevard, Novato, CA, 94949-6191. All rights reserved. Reproduced by permission of Academic Therapy Publications.

Adaptation of the Diamond and Taylor (1996) *Peg-Tapping Executive Functioning Task*. Copyright © 1996; Blair 2002; Smith-Donald, Raver, Hayes, and Richardson, 2007.

Selected items from the *Early Childhood Longitudinal Study, Kindergarten Class of 1998-99 (ECLS-K)*, National Center for Education Statistics. To include items reproduced from the *Test of Early Mathematics Ability, 3rd Ed. (TEMA-3)*, by H.P. Ginsburg, and A.J. Baroody. Copyright © 2003, Pro Ed, Inc. Used with permission.

Selected items from the *Early Childhood Longitudinal Study, Birth Cohort (ECLS-B)*, National Center for Education Statistics. Used with permission.

