Early Head Start Home Visits and Classrooms: Stability, Predictors, and Thresholds of Quality

June 2015

SHORT REPORT

Baby FACES 2009

OPRE Report 2015-34

This page is left blank for double-sided printing.

Early Head Start Home Visits and Classrooms: Stability, Predictors, and Thresholds of Quality

OPRE Report 2015-34

June 2015

Submitted to:
Amy Madigan, *Project Officer*
Office of Planning, Research and Evaluation
Administration for Children and Families
U.S. Department of Health and Human Services

Submitted by:
Nikki Aikens
Yange Xue
Eileen Bandel
Pia Caronongan
Cheri A. Vogel
Kimberly Boller
Mathematica Policy Research

Project Director:
Cheri A. Vogel Mathematica
Policy Research P.O. Box 2393
Princeton, NJ 08543-2393

Contract Number: HHSP23320072914YC
Mathematica Reference Number: 06432.140

Suggested citation:
N. Aikens, Y. Xue, E. Bandel, P. Caronongan, C.A. Vogel, and K. Boller (2015). Early Head Start Home Visits and Classrooms: Stability, Predictors, and Thresholds of Quality. OPRE Brief 2015-34. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

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

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 page is left blank for double-sided printing.

CONTENTS

I. INTRODUCTION 1

A. Early Head Start performance measures framework 2

B. Background and motivation of analysis 3

C. Research questions 5

II. BABY FACES METHODOLOGY AND ANALYTIC SAMPLE 7

A. Data sources and measurement in the current analyses 7

B. Analytic sample 12

III. WHAT IS THE QUALITY OF HOME VISITS AND CLASSROOMS IN EARLY HEAD START? DO CHILDREN EXPERIENCE VARYING LEVELS OF SERVICE QUALITY DURING THEIR ENROLLMENT? 13

IV. WHAT PREDICTS QUALITY IN EARLY HEAD START? 15

A. Predictors of home visit quality 15

B. Predictors of classroom quality 18

V. WHAT ARE THE ASSOCIATIONS BETWEEN EARLY HEAD START QUALITY AND CHILD AND PARENTING OUTCOMES? ARE THERE THRESHOLD EFFECTS? 21

A. Home visit quality 22

B. Classroom quality 24

VI. CONCLUSIONS AND IMPLICATIONS 27

REFERENCES 29

APPENDIX A. ASSOCIATIONS AND THRESHOLD EFFECTS FOR QUALITY MEASURES AND CHILD OUTCOMES FROM MULTIVARIATE HLM MODELS 35

This page is left blank for double-sided printing.

TABLES

II.1. Key measures for each analysis 8

III.1. Mean HOVRS-A—subscale ratings, by year 13

III.2. Descriptive statistics for the CLASS-T domains across time 13

IV.1. Home visit, child, family, staff and program predictors of home visit quality 17

IV.2. Staff, classroom and program predictors of classroom quality 19

V.1. Summary of GAM results for home visit quality 23

V.2. Summary of GAM results for the CLASS-T 25

A.1. Associations between Home Visitor Strategies and child outcomes at age 2 and tests for threshold effects 36

A.2. Associations between Home Visitor Strategies and child outcomes at age 3 and tests for threshold effects 37

A.3. Associations between classroom quality and child outcomes at age 2 and tests for threshold effects 38

A.4. Associations between classroom quality and child outcomes at age 3 and tests for threshold effects 39

FIGURES

I.1. Early Head Start conceptual framework 3

V.1. Age 2 HOVRS-A Home Visitor Strategies threshold with age 2 BITSEA Problem Behaviors 23

V.2. Age 2 CLASS-T Emotional and Behavioral Support threshold with age 2 BRS Orientation/Engagement 25

V.3. Age 3 CLASS-T Engaged Support for Learning threshold with age 3 ASQ-3 Problem Solving 26

This page is left blank for double-sided printing.

OVERVIEW

The Early Head Start Family and Child Experiences Survey (Baby FACES) is a descriptive study of Early Head Start programs designed to inform policy and practice at both national and local levels. This longitudinal study was conducted in 89 Early Head Start programs and follows two cohorts of children through their time in Early Head Start, starting in 2009, the first wave of data collection. The Newborn Cohort includes 194 pregnant mothers and newborn children, and the 1-year-old Cohort includes 782 children who were approximately 1 year old (ranging from 10 to 15 months) at the outset of the study.

An important aspect of the Early Head Start conceptual framework is the association between the provision of high quality services and children's growth and development. This report describes Early Head Start classroom and home visit quality, the stability of quality while children are enrolled, and the predictors of quality. The report also examines the association between quality and child and parenting outcomes, including whether there are threshold effects in these associations. Specifically, the report answers the following questions:

1. What is the quality of home visits and classrooms in Early Head Start? Do children experience varying levels of service quality during their enrollment?
    - Classroom and home visit quality scores are relatively stable over time, with one exception—engaged support for learning, a dimension of classroom quality, decreases over time.
2. What predicts quality in Early Head Start?
    - Several indicators of teacher wellbeing predict classroom quality. Higher job satisfaction, lower depressive symptoms, and better teacher-parent relationships are related to higher classroom quality. Additionally, higher levels of quality are found in classrooms with teachers with a bachelor's degree or higher.
    - None of the program characteristics predict classroom quality. The percentage of dual language learners in the classroom is the only classroom characteristics associated with quality.
    - Some characteristics of home visits are related to home visit quality. Higher quality is observed when more time is spent on parent-child activities (relative to child-focused activities), when the visit aligns with the visit plan, and when at least one other adult is present during the visit. Lower quality is observed when more time is spent on staff-family relationship building (relative to child-focused activities).
3. What are the associations between Early Head Start quality and child and parenting outcomes? Are there threshold effects?
    - There are both linear associations and threshold effects between quality and child and parenting outcomes and, in some cases, no association between quality and outcomes.

Taken together, the findings suggest there are areas related to quality in which teachers and home visitors may need more support (e.g., engaged support for learning). In addition, the results indicate that relationships between quality and child outcomes are complex. More research is needed before recommendations can be made about whether quality improvement efforts should aim to help lower quality settings achieve a particular quality cut point.

I. INTRODUCTION

The Early Head Start Family and Child Experiences Survey (Baby FACES) is a descriptive study of Early Head Start programs designed to inform policy and practice at both national and local levels. In 2007, the Office of Planning, Research & Evaluation (OPRE) in the Administration for Children and Families (ACF), U.S. Department of Health and Human Services (USDHHS), contracted with Mathematica Policy Research and its partners to implement this six-year longitudinal study in 89 Early Head Start programs around the country. It enrolled two cohorts of children in spring 2009 and followed them through their time in Early Head Start: (1) a Newborn Cohort comprising those families in which the mother was pregnant or the child was less than two months old, and (2) a 1-year-old Cohort comprising children who were approximately age 1 at the first wave of data collection. The study provides a descriptive snapshot of Early Head Start services, including their intensity and quality; the characteristics of the children and families served; and how children and families are faring in terms of key areas of development and well-being. Baby FACES gathers information from parent interviews on participating families; direct child assessments and staff (home visitor or teacher) reports of children's development; individual interviews with those teachers and home visitors about their characteristics, experience, and training; and observations of study children's classrooms and home visits.

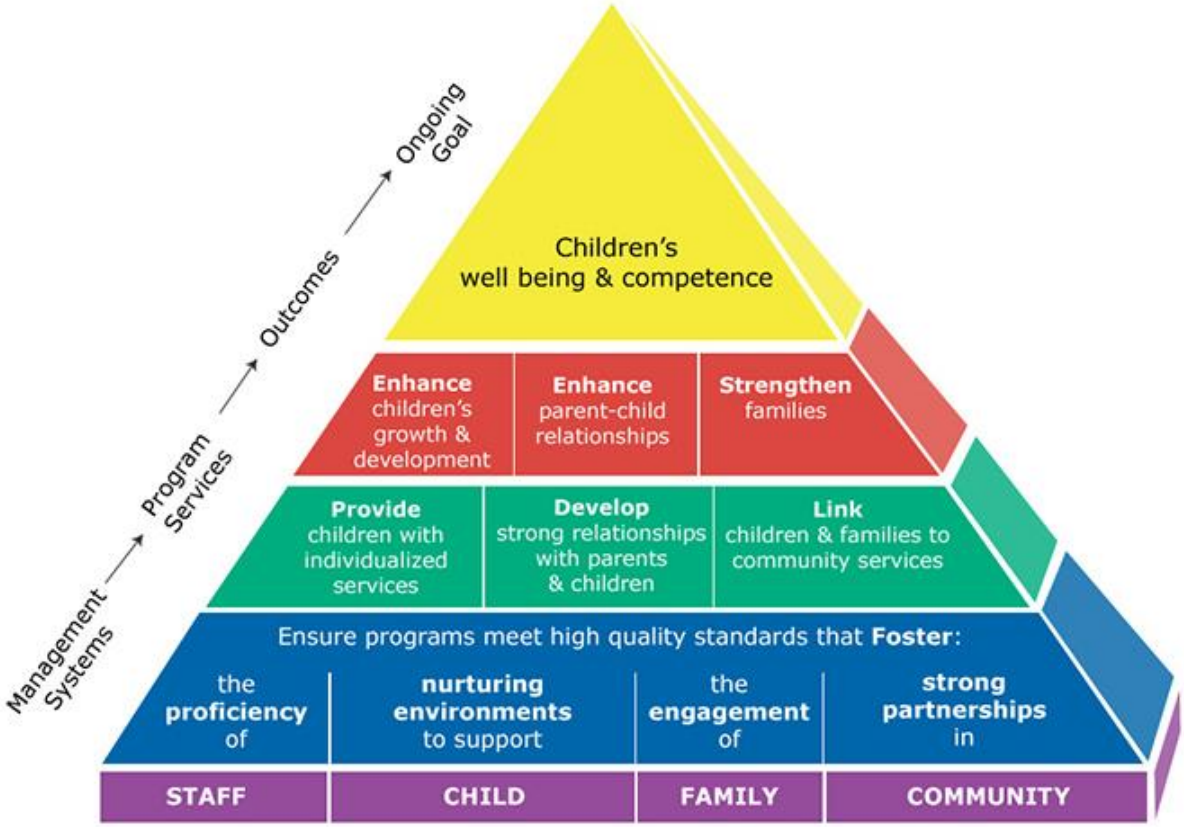
This report describes Early Head Start classroom and home visit quality, the stability of quality while children are enrolled, and the predictors of that quality. The report also examines associations between process (versus structural) quality and child and parenting outcomes, including whether there are threshold effects in these associations. Chapter I provides a brief overview of the Early Head Start program and the research questions and analytic motivations guiding this report. Chapter II provides an overview of the data and sample used for this report's analysis before turning to a description of findings (Chapters III through V) and implications of the findings (Chapter VI).

A. Early Head Start performance measures framework

Early Head Start is a two-generation program that began in 1995 as a federal initiative for low-income pregnant women and families with infants and toddlers 3 years old or younger. Early Head Start programs provide a wide range of services, including child development services, child care, parenting education, case management, health care and referrals, and family support. In addition to delivering many services directly, programs also form partnerships with other community service providers.

To ensure the quality of their offerings, Early Head Start programs adhere to two key institutional benchmarks: the Head Start Program Performance Standards (hereafter performance standards) and the Framework for Programs Serving Infants and Toddlers and Their Families. The performance standards are the rules and regulations that explicitly identify what programs must do to ensure high quality services (for example, they specify child-to-adult ratios in child care centers, educational requirements for staff, and the types of services that must be offered) (ACF 1996). The framework is a conceptual model that describes the mechanisms by which high quality programs are thought to affect children's outcomes. The framework is structured as a pyramid that rests on a foundation of four cornerstones (community, staff, family, and child development) that the Advisory Committee on Services for Infants and Toddlers deemed essential for quality Early Head Start programs. On this foundation, the framework builds four layers, with management systems forming the base that supports program services. These services bring about positive family and child outcomes and contribute to the goal of children's well-being and competence (Figure I.1).

Figure I.1. Framework for Programs Serving Infants and Toddlers and Their Families



Source: Framework for Programs Serving Infants and Toddlers and Their Families. Head Start Approach to School Readiness. HHS/ACF/OHS. 2012.

B. Background and motivation of analysis

An important aspect of the Early Head Start conceptual framework is the association between the provision of high quality services and children's growth and development. Early Head Start quality is important to understand because it has implications for training and technical assistance, professional development, staff hiring, and other efforts. The quality of early childhood services is multidimensional and encompasses not only characteristics of staff but also the quality of the interactions and relationships among staff members and the children and parents with whom they work. Accordingly, Baby FACES provides important information about the overall quality of classrooms and home-based services for families and children and about attributes of classrooms and home visits that relate to quality.

Research on quality typically considers two aspects of quality: structural and process factors (Love et al. 2005). Structural factors focus on aspects of the physical environment that may be more easily regulated—including, child-teacher ratios, group sizes, curriculum, and teacher education. Process factors focus on interactional aspects of quality—including teacher or home visitor behavior, teacher-child or home visitor-child interactions, and quality of instruction. In other words,

structural factors highlight *what* is provided and *by whom*, while process factors describe *how* services are provided. This report is especially focused on the home visit and classroom process factors of quality because few prior studies have examined process quality in infant and toddler settings, especially in Early Head Start.

Home visit quality. For families served in Early Head Start home-based settings, home visiting is intended to support children's development, parenting outcomes, and the parent-child relationship. In general, home visiting approaches vary in their success achieving these goals (Roggman et al. 2008b; Sweet and Appelbaum 2004). Although evidence in the field for the efficacy of home visiting strategies is mixed (Astuto and Allen 2009; Del Grosso et al. 2011; Paulsell et al. 2010), stronger effectiveness is likely when the quality of the home visit is high and when the relationship between the home visitor and the family is strong. There is evidence that parents are more engaged and retention is higher when visitors are matched to the family on characteristics such as ethnicity and language (Astuto and Allen 2009). The home visitor–parent relationship is also associated with Early Head Start parents' engagement and involvement in visits and with children's vocabulary at 36 months (Roggman et al. 2008a). In addition, qualitative research highlights the importance of home visitor conscientiousness (for example, honoring commitments to parents) and the match between home visitor and parent life experiences (Brookes et al. 2006). Other studies have found associations with Early Head Start children's cognitive and language development, Home Observation for Measurement of the Environment Inventory (HOME) scores, and parent support for language and learning at 36 months when home visits are child-focused rather than adult-focused (Raikes et al. 2006). These findings suggest that the content of home visits could be important for enhancing child and parenting outcomes.

Classroom quality. Classroom process quality is likely shaped by multiple factors. Central among these are perhaps the physical characteristics of the classroom (aspects of structural quality) and the background characteristics of the teachers. Furthermore, characteristics of the programs within which classrooms exist likely also have effects on classroom process quality. Specifically, the following program, classroom, and teacher characteristics have been associated with process quality: (1) the child–teacher ratio (the number of children per teacher), (2) group size (the total number of children assigned to a teacher or team of teachers), (3) teacher education and training, and (4) teacher salary (Lamb 1998; de Schipper et al. 2006; Vandell and Wolfe 2000). Related characteristics of families served also provide important contextual factors that may influence classroom process quality (Resnick and Zill 2003). Because of the interconnectedness of various program, classroom, teacher, and family factors, each of these might be examined as outcomes variables or as predictors of other desired outcomes. In this report, classroom structural quality as well as other classroom, teacher, and family characteristics will be analyzed as *predictors of process quality* and not as outcomes variables in their own right. We hypothesize that classroom structural quality as well as other classroom, teacher, and family characteristics help shape classroom process quality. Furthermore, we expect that program, home visitor, home visit, and family characteristics will predict home visit quality.

Classroom quality has also been linked to the developmental outcomes of infants, toddlers, and preschoolers served in these settings (Burchinal et al. 2010; Keys et al. 2013; Vandell et al. 2010). For example, modest associations between observed classroom quality (both process and structural) and cognitive, language, and social-emotional outcomes for low-income infants, toddlers, and preschoolers have been found in individual studies and meta-analyses (Burchinal et al. 1996, 2008a, 2009). Specifically, sensitive, responsive, and stimulating interactions with teachers—examples of process quality—have been linked with social-emotional and language development and pre-academic skills (Burchinal et al. 2008b; NICHD Early Child Care Research Network 1996, 1997,

1998, 2006; Whittaker and Harden 2010). Yet, the amount of evidence linking quality with child outcomes in infant and toddler classroom settings is limited, with process quality having the least existing research.

Threshold effects. Recent meta-analyses point to consistent, yet modest associations between classroom quality and child outcomes when studied using the assumption of a linear relationship (Belsky and Pluess 2012; Burchinal et al. under review; Burchinal et al. 2011a, 2011b; Burchinal et al. 2009; Keys et al. 2013). One possible explanation for these modest findings is that children need to experience quality at a high enough level (threshold) before programs can positively influence their outcomes. Recent analyses of data from large-scale and smaller studies have identified possible threshold effects of quality on preschool children's outcomes (Burchinal et al. 2011a, 2011b, 2010; Moiduddin et al. 2012; Torquati et al. 2011).[1] That is, analyses suggest that the association between quality and outcomes is stronger in higher-quality compared to lower-quality classrooms. For example, using data across a number of large child care studies, analyses in the Child Care and Early Education Quality Features, Thresholds and Dosage and Child Outcomes: (Q-DOT) Study (Zaslow et al. 2010)[2] identified threshold effects for two different measures of classroom quality and outcomes. First, there was a threshold effect in the association between Classroom Assessment Scoring System (CLASS) Instructional Support and both language and reading. Second, there was a threshold effect between the Early Childhood Environment Rating Scale-Revised (ECERS-R) Teaching and Interactions and language (Burchinal et al. under review; Burchinal et al. 2011a, 2011b; Xue et al. 2011). Few studies have investigated threshold effects in the associations between the quality of infant/toddler classrooms and child outcomes, and none has examined threshold effects for home visit quality.

C. Research questions

While there is some evidence from prior research that certain factors promote high quality services in early care and education settings, little is known about the predictors of quality in infant and toddler settings, specifically in Early Head Start. In order to fill gaps in the literature about quality and its predictors in Early Head Start, this report addresses the following research questions:

1. What is the quality of home visits and classrooms in Early Head Start? Do children experience varying levels of service quality during their enrollment? (RQ1)
2. What predicts quality in Early Head Start? (RQ2)

[1] The studies 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 (Burchinal et al. 2011).

[2] Q-DOT is a design project funded by the Office of Planning, Research and Evaluation (OPRE) at the Administration for Children and Families (ACF) that examines associations between the quality of early care and education 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. Products include a literature review, secondary data analyses, conceptual framework development, and development of alternative designs and methods to conduct research to systematically test which features, thresholds, and dosages of child care quality are linked to child outcomes.

a. Do characteristics of the home visitor, the home visit itself (such as what happens during visits and who participates), the child and family, or the program predict the quality of Early Head Start home visits?

b. Do teacher, classroom, or program characteristics predict the quality of Early Head Start classrooms?

3. What are the associations between Early Head Start quality and child and parenting outcomes? Are there threshold effects? (RQ3)

   a. What are the associations between Early Head Start home visit and classroom quality and child and parenting outcomes?

   b. Are there threshold effects in the associations between quality and outcomes? That is, does the magnitude of associations between quality and child and parenting outcomes differ in higher- and lower-quality classrooms and home visits?

   c. What is the minimum quality threshold needed to see significant associations between quality and outcomes?

The analyses focus on describing process quality in Early Head Start, the structural and other predictors of that quality, and the association of process quality with child and parenting outcomes. In the next chapter is a description of the data sources, measures, and sample used for the analyses.

II. BABY FACES METHODOLOGY AND ANALYTIC SAMPLE

This chapter describes the measures and sample included in the analyses, including a description of the approach to each set of analyses described in subsequent chapters. For additional details about data collection procedures, assessor training, and analytic issues and approaches to resolving them, see the first Baby FACES report and technical appendices (Vogel et al. 2011).

A. Data sources and measurement in the current analyses

The Baby FACES study collected data from multiple sources, including parent and staff interviews, staff reports of children's outcomes, observations of classroom and home visit quality, and direct assessment of children's outcomes. Data were collected in the spring of each year (spring 2009, 2010, 2011, 2012), and the specific measures used at each round depended on the age of participating children (age 1, 2, or 3).[3] Provided here is a brief overview of the measures used in each set of analyses described in this report: average quality and stability over time, predictors of quality, and linear relationships or threshold effects between quality and child and parenting outcomes (Table II.1). Given the focus of these analyses, additional information is provided about the study's home visit and classroom-quality measures in Box II.1. More detail is provided about the different sets of analyses in subsequent chapters.

As noted, to assess key aspects of the quality of both home visits and center-based classrooms, field staff conducted structured observational assessments of home visits and classrooms, depending on the study child's service option. The study team observed one home visit per home visitor (not one per child).[4] For home visits, the team used the Home Visitor Rating Scale-Adapted (HOVRS-A; Roggman et al. 2009) and its manual (Hallgren et al. 2009), an adaptation of the HOVRS (Roggman et al. 2006b). For classroom observations, the study used the Classroom Assessment Scoring System-Toddler (CLASS-T; Pianta et al. 2010a) for classrooms serving 2- and 3-year-old children. Box II.1 provides information on the scoring and use of these measures.

[3] The ages of children at each data collection round depended on their cohort. In spring 2009, participating children were either newborns or approximately age 1. Participating children were ages 1 and 2 in spring 2010, 2 and 3 in spring 2011, and 3 in spring 2012.

[4] The study team based the decision to observe one visit per home visitor rather than per child on two considerations. The first was the logistical difficulty of scheduling an observation of each child's home visit in the site visit week (five days). The second was that prior research suggested that home visitors tend to have low intra-visitor variability; in other words, home visitors tend to provide home visits of a consistently similar quality (personal communication with Lori Roggman, 2008).

Table II.1. Key measures for each analysis

		Average Quality and Stability (RQ1)[a]	Predictors of Quality (RQ2)[a]	Quality-Outcomes Associations (RQ3)[a]
Home visit and classroom quality				
Home Visit Rating Scale-Adapted (HOVRS-A; Roggman et al. 2009), modified from the HOVRS (Roggman et al. 2006)	Observations of home visits used the HOVRS-A, an adaptation of the HOVRS (Roggman et al. 2006). The HOVRS-A consists of seven items measuring the quality of home visitor strategies and effectiveness at involving and engaging the family during home visits.	X	X[b]	X[b]
Home visit characteristics and content (Boller et al. 2009)	During structured observations of home visits, field staff also collected data on the topics covered, activities, and structure of the home visit. They also collected information on the alignment of the visit with the home visitor's lesson plan and the extent of environmental distractions (television, phone calls, visitors, pets, other children, noise, and such) during the visit.		X	
Classroom Assessment Scoring System, Toddler Version (CLASS-T; Pianta et al. 2010a, 2010b)	Classroom observations used the CLASS-T (Pianta et al. 2010a), an adaptation of the Pre-K CLASS (Pianta et al. 2008), which focuses on teacher-child interaction quality in toddler child care classrooms. The CLASS-T measures process quality along two broad domains and eight dimensions.	X	X	X
Child-adult ratio and group size	Center-based classroom observations also included child-adult ratios and group sizes.		X[c]	
Classroom composition	Teachers reported on the percentage of dual language learners in the classroom.		X[c]	
Staff and program characteristics				
The Center for Epidemiologic Studies Depression Scale—Short Form (CESD-SF; Radloff 1977; Ross et al. 1983)	The CESD-SF is the short form of the full-version CESD, which is a self-administered screening tool used to identify symptoms of depression or psychological distress. The study used the tool to measure depressive symptoms in teachers and home visitors.		X	
Staff background characteristics	The teacher and home visitor interviews included sections with items that broadly covered: staff training and supervision, job satisfaction,[c] languages spoken (by the staff member and families in the classroom), racial/ethnic group membership, experience, and education.		X	
Program approach: program level	The program approach at the program level was based on director responses to questions regarding (a) the types of services their programs offer (center based, home based, or combination) and (b) separately for each service option, the frequency of services offered.		X	
Program implementation	Program directors reported on various aspects of program implementation across five cornerstones (child development, family development, staff development, community building, and management systems). The study created an aggregate measure of implementation by taking the average of ratings for each cornerstone from all three study years. Programs with ratings equal to or above 3 (out of a possible 4) on each of the implementation cornerstone ratings at each time point are designated as "fully implemented."		X	

		Average Quality and Stability (RQ1)[a]	Predictors of Quality (RQ2)[a]	Quality-Outcomes Associations (RQ3)[a]
Population served	Program directors reported the proportion of families enrolled in their programs who were facing socio-demographic and psychological risks. The analysis uses information from 2009 and 2010 to identify programs serving a high proportion (50 percent or more) of families facing these risks (more than 3 demographic risks, mental health or substance abuse problems, and living in an unsafe neighborhood or with exposure to family violence).		X	
Family characteristics, parenting, and the home environment				
Parent-Caregiver Relationship Scale (PCRS; Elicker et al. 1997)	The PCRS measures the perceived relationship between the parent and the teacher or home visitor. Items capture important dimensions of the parent-staff relationship, including trust and confidence, communication, respect and acceptance, caring, competence and knowledge, partnership and collaboration, and shared values.		X	
Maternal demographic risk index (ACF 2001)	The maternal demographic risk index captures the multiple dimensions of risk of poor developmental outcomes a child may face as a consequence of his or her mother's socioeconomic circumstances. The index comprises three risk groups (lower, medium, and highest). The index was constructed by summing the number of the following risk factors that the mother faced, based on the parent interview: (1) being a teenage mother, (2) having no high school credential, (3) receiving public assistance, (4) not being employed or in school or training, and (5) being a single mother.		X[d]	X
Maternal and child characteristics	The parent interview also included sections that broadly covered many different aspects of the family and home environment, including child age and birth weight, whether family enrolled during pregnancy, family racial/ethnic membership, and languages spoken in the home.		X[d]	X[e]
Parenting Synchronicity	The Parent-Child Interaction (PCI) Rating Scales for the Two-Bag Task consists of 12 scales that assess a range of child and parent behaviors. A total of eight scales address both positive and negative parenting behaviors. The study created a composite parenting score (synchronicity) by computing a mean of scores on parental sensitivity, positive regard, and relationship quality.			X
Parent support for child learning	This is a composite index that captures the degree of cognitive stimulation provided to children in the home setting by parents or other household members. Component measures include: PCI Cognitive Stimulation scores; the Parenting Interactions with Children: Checklist of Observations Linked to Outcomes (PICCOLO) Teaching scale scores; verbal responsiveness items from the HOME (whether the parent converses with the child at least twice during the visit, answers the child's questions/requests, responds to the child's talk verbally, and uses complex sentence structure in social exchanges); parent-reported frequency of book reading and storytelling with the child by any household member; and whether any household member helped the child learn shapes/sizes, the alphabet, colors or numbers. Scores were derived by averaging z-scored values for each of the components.			X

		Average Quality and Stability (RQ1)[a]	Predictors of Quality (RQ2)[a]	Quality-Outcomes Associations (RQ3)[a]
Parent provision of learning materials	This was assessed through a combination of parent interview items and interviewer observations from the HOME scale. Component items include: the number of accessible children's books in the home; the availability of puzzles, toys that teach colors/sizes/shapes, and toys that teach numbers; the accessibility of toys, games, and books appropriate for preschoolers; and whether the child had access to an audio device and at least five children's tapes. Scores were derived by averaging z-scored values for each of the components.			X
	Child development			
Preschool Language Scale—Fourth Edition (PLS-4; Zimmerman et al. 2002a, 2002b).	The PLS-4 is a direct child assessment used to evaluate receptive and expressive language skills, as well as understanding and use of grammatical rules for children from birth to 6 years of age. It is composed of two subscales: Auditory Comprehension and Expressive Communication. The study used the Auditory Comprehension subscale for both of the English and Spanish editions of PLS-4 when children were 2.			X
Ages & Stages Questionnaires, Third Edition (ASQ-3; Squires et al. 2009)	The ASQ-3 is a parent-report tool for screening children aged 1 month through 5-1/2 years for developmental delays. These surveys focus on the assessment of five key developmental areas: (1) communication, (2) gross motor, (3) fine motor, (4) personal-social, and (5) problem solving.			X
Peabody Picture Vocabulary Test-4th Edition (PPVT-4;Dunn and Dunn 2007)	The PPVT-4 is a direct assessment measure of receptive vocabulary in which children are shown a plate with four pictures and asked to point to the one that indicates the target word that is stated by the assessor "point to [target word]." It is a norm-referenced standardized test and is suitable for a wide range of ages, from 2½ years through adulthood.			X
The Brief Infant Toddler Social Emotional Assessment (BITSEA; Briggs-Gowan and Carter 2006)	The BITSEA is the screener version of the longer ITSEA, which is designed to detect delays in the acquisition of social-emotional competencies as well as social-emotional and behavior problems in children 12 to 36 months old.			X
Bayley Behavioral Rating Scale (BRS; Bayley 1993)	The BRS measures the child's behavior during the child assessment. The BRS is one of the three component scales of the Bayley Scales of Infant Development—Second Edition (Bayley 1993). There are two subscales of the BRS used in Baby FACES: (1) Orientation/Engagement, measuring the child's cooperation with the assessor during the assessment, positive affect, and interest in the test materials, and (2) Emotional Regulation, measuring the child's ability to change tasks and test materials, negative affect, and frustration with tasks during the assessment.			X

[a] For each of these sets of analyses, separate analyses are run for home visit and classroom quality.
[b] As described in subsequent sections, these home visit analyses only include the Visitor Strategies subscale of the HOVRS-A.
[c] These variables are only included in the classroom quality analyses.
[d] These variables are only included in the home visit quality analyses.
[e] As noted in Box V.1, the child and family characteristics included in the linear and threshold analyses include cohort, age, gender, race/ethnicity, dual language learner status, family poverty ratio, and maternal demographic risk composite.

Box II.1 Measures of home visit and classroom quality

The HOVRS-A consists of seven items, which can be combined to form two subscale scores: Visitor Strategies (four items) and Visitor Effectiveness (three items). Visitor Strategies items include the home visitor's (1) Responsiveness to the Family, (2) Relationship with the Family, (3) Facilitation of Parent-Child Interaction, and (4) Nonintrusiveness. Visitor Effectiveness includes (1) Parent-Child Interaction during the visit, (2) Parent Engagement, and (3) Child Engagement. Items on HOVRS-A are rated from 1 to 5, based on indicators defined for each item and with anchor ratings of 1 (minimal), 3 (moderate), and 5 (good practice). A 3 rating indicates that the observer saw a sufficient level of indicators of the model of home visiting that aims to facilitate parent-child interaction. Because of the theoretical underpinnings of this instrument, home visiting programs that adhere to alternative models of home visiting and stress different types of behaviors (such as home visitor–child interactions) will not score as highly on the HOVRS-A.

During observations of home visits, field staff also collected data on the content and characteristics of the home visit, including topics (such as the child's health and development, parenting, the parent's health and well-being, parent employment and education, and community services); activities (including assessment, provision of information, goal setting, and crisis intervention); and structure (for example, participating children and adults and languages used) (Boller et al. 2009). The HOVRS-A was used for home visit observations when children were 1, 2, and 3 years old.

Description of subscales and items in HOVRS-A

Subscales/items	Description of subscales/items
Visitor Strategies Quality	Assesses the strategies used by the home visitor when working with families during home visits
Responsiveness to Family	Assesses the extent to which the home visitor is (1) prepared for the home visit, (2) observes and responds to the parent and child during the home visit, and (3) elicits input on the content and activities of the home visit from the parent
Relationship with Family	Examines the nature of the relationship between the home visitor and the family during the visit, including warmth, positive interactions, and respect
Facilitation of Parent-Child Interaction	Assesses the ability of the home visitor to facilitate and promote positive parent-child interactions during the home visit
Nonintrusiveness	Focuses on the lack of intrusiveness by the home visitor on parent behavior and parent-child interactions during the visit
Visitor Effectiveness Quality	Assesses how effective the home visit is at engaging the parent and child in the home visit activities and in interactions with each other
Parent-Child Interaction	Examines the nature of the parent-child relationship, as observed during the home visit, including parent-child warmth and physical closeness, parent attentiveness and responsiveness to the child, and parent-child joint attention
Parent Engagement	Examines the engagement of the parent in the activities of the home visit, including involvement and interest
Child Engagement	Focuses on the child's engagement in the activities of the home visit, including involvement and interest

Classroom observations of 2- and 3-year-olds were conducted using the CLASS-T (Pianta et al. 2010a), a downward extension of the Pre-K CLASS (Pianta et al. 2008), which focuses on teacher-child interaction quality in toddler child care classrooms. The CLASS-T measures process quality along eight dimensions (Positive Climate, Negative Climate, Teacher Sensitivity, Regard for Child Perspectives, Behavior Guidance, Facilitation of Learning and Development, Quality of Feedback, and Language Modeling). The dimensions exist within two broader domains: Emotional and Behavioral Support and Engaged Support for Learning. Dimensions are defined by observable indicators along a seven-point scale, with ratings reflecting scores in the low (1-2), mid (3-5), and high (6-7) ranges.

For both the HOVRS-A and CLASS-T, observers look for evidence of specific indicators as they rate each item.

Description of domains and dimensions in the CLASS-T

Domains/dimensions	Description of domain/dimension
Emotional and Behavioral Support	
Positive Climate	Focuses on the degree of warmth, respect, and mutual enjoyment communicated between the teacher and children, either verbally or non-verbally
Negative Climate	Examines the frequency and intensity of teacher and child expressions of negativity
Teacher Sensitivity	Assesses the teachers' responsiveness to and awareness of children's individual needs and emotional functioning, encompassing the extent to which the teacher is available to provide reassurance and encouragement
Regard for Child Perspectives	Examines the degree to which teacher-child interactions reflect children's interests and motivations, and encourage children's responsibility and independence
Behavior Guidance	Examines the teacher's ability to promote children's self-regulation by using proactive approaches, supporting positive behaviors, and guiding and curtailing problem behavior
Engaged Support for Learning	
Facilitation of Learning and Development	Assesses the manner by which the teacher facilitates activities that support children's learning and developmental opportunities
Quality of Feedback	Examines the degree to which the teacher provides feedback that promotes learning and understanding, and extends children's participation
Language Modeling	Focuses on the quality and quantity of the teacher's use of language to support and encourage children's language development

B. Analytic sample

Samples for analyses on the predictors of HOVRS-A and CLASS-T. For the HOVRS-A, the study team observed each home visitor in the spring of each year that she was linked to one or more study children (spring 2009, 2010, 2011, and 2012). Over four years, 322 different home visitors were observed at least once. Seven home visitors were observed four times and 74 home visitors three times. A total of 80 home visitors were observed twice and 161 home visitors only once. More than half (58 percent) of home visitors observed on multiple occasions were observed with the same family across assessments. The remaining home visitors were observed with two or three different families.

For the CLASS-T, each teacher was observed once in the spring of each year if he or she was linked to one or more study children (spring 2010, 2011, and 2012). Up to three observations were conducted for each of the 364 teachers that make up the sample (251 were observed once, 106 were observed twice, and 7 were observed three times).

Given the Baby FACES study was not designed to collect data on a representative sample of staff, the analytic sample is not representative of all Early Head Start teachers or home visitors. As such, the results of these analyses pertain only to staff who were linked to one of the sample children and who were observed at one or more time points.

Samples for HOVRS-A and CLASS-T linear and threshold analyses. The linear and threshold analyses for the HOVRS-A and CLASS-T include samples of children who were observed with a home visitor or whose classroom teachers were observed. The HOVRS-A analyses include 166 children and their home visitors at age 2 and 128 children and their home visitors at age 3. The CLASS-T analyses include 361 children in 243 classrooms at age 2 and 317 children in 226 classrooms at age 3. Analyses are weighted to be representative of children who were enrolled in Early Head Start in spring 2009 and continued to be enrolled at age 2 or age 3.

Next, the discussion turns to the analyses conducted to address each of the research questions.

III. WHAT IS THE QUALITY OF HOME VISITS AND CLASSROOMS IN EARLY HEAD START? DO CHILDREN EXPERIENCE VARYING LEVELS OF SERVICE QUALITY DURING THEIR ENROLLMENT?

Examined in this chapter are home visit and classroom quality in Early Head Start and the extent to which quality changes over the time children are enrolled.

Average home visit quality is in the moderate range and is stable over time. Descriptive statistics indicate that, across years, children and their parents have home visits scoring in the 3-point (out of a possible 5) or moderate range on average. There are no statistically significant differences in mean home visit quality ratings from year to year (Table III.1). Specifically, the average scores for the HOVRS-A Visitor Strategies are 3.2 in 2010 and 3.4 in 2012; scores for the Visitor Effectiveness ratings are 3.6 in 2010 and 3.9 in 2012. Results from fitting the multilevel model confirm that the year-to-year differences are not statistically significant.

Table III.1. Mean HOVRS-A—subscale ratings, by year

	2009			2010			2011			2012		
	N	Mean	SD	N	Mean	SD	N	Mean	SD	N	Mean	SD
Visitor Strategies	232	3.2	0.97	185	3.3	0.82	135	3.3	0.87	19	3.4	0.99
Visitor Effectiveness	231	3.6	0.97	185	3.6	0.90	135	3.7	0.83	19	3.9	0.99

Source: Spring 2009, 2010, 2011, 2012 home visit observations

Note: Home visit observations were conducted with staff of both the 1-year-old Cohort and Newborn Cohort children. Scores range from 1-5.

HOVRS-A = Home Visit Rating Scale-Adapted

SD = Standard deviation

Average classroom quality is the mid-range and only Engaged Support for Learning changes (declines) over time. Examining change in classroom quality over time reveals that the mean for Engaged Support for Learning decreases by about a third of a point between 2010 and 2012 [3.6 (SD = 1.21) in 2010 versus 3.3 in 2012 (SD = 0.67) (Table III.2)]. Results from multilevel models presented in Chapter IV confirm this trend (Table IV.2). Ratings in this domain are in the mid-range across years.

Table III.2. Descriptive statistics for the CLASS-T domains across time

	2010			2011			2012		
	N	Mean	SD	N	Mean	SD	N	Mean	SD
Engaged Support for Learning	218	3.56	1.21	226	3.19	1.24	40	3.25	0.67
Emotional and Behavioral Support	218	5.29	0.83	226	5.23	0.90	40	5.30	0.59

Source: Spring 2010 to spring 2012 classroom observation

Note: Classroom observations were conducted with staff of both the 1-year-old Cohort and Newborn Cohort children. Scores range from 1-7.

CLASS-T=Classroom Assessment Scoring System-Toddler

SD = Standard deviation

In contrast, the mean for Emotional and Behavioral Support is stable over time, with mean scores of 5.2 in 2010 and 5.3 in 2012 (Table III.2). Results from multilevel models presented in Chapter IV indicate no significant differences in Emotional and Behavior Support across years (Table IV.2). Ratings in this domain are in the mid-range across years. Scores compare favorably with those reported in the pilot study of the adapted toddler measure (Thomason and La Paro 2009). Also, lower instructional quality compared to emotional aspects of the classroom is consistent with findings in studies of older preschool children (for example, the Head Start Family and Child Experiences Survey, Aikens et al. 2010; and the Chicago Program Evaluation Project; Ross et al. 2008).

IV. WHAT PREDICTS QUALITY IN EARLY HEAD START?

Next, analyses examine whether staff and program characteristics are related to home visit and classroom quality in Early Head Start. For home visit quality, analyses examine whether: (1) home visit quality varies among home visitors within programs, (2) home visit quality varies between programs, and (3) specific visit content and characteristics and/or home visitor, program, or child/family characteristics contribute to the quality of home visits. For classroom quality, analyses use multilevel models to examine whether: (1) classroom quality varies between teachers within programs and across programs, and (2) teacher and program characteristics are associated with classroom quality.

Correlations between characteristics of programs and staff with quality ratings are reported elsewhere (Vogel et al. 2011). Analyses focused on a relatively narrow set of variables as predictors in order to maintain model parsimony. The approach to selecting predictors for the models was guided by a number of factors, including theory, practicality, and prior empirical findings (that is, bivariate associations). Wherever possible, analyses included variables with less missing data and more variability. Analyses also sought to minimize redundancy/collinearity and maintain parallelism with models across settings (that is, home- and center-based models). The final set of variables was selected to represent key domains and levels (program, staff, and family) of interest.

A. Predictors of home visit quality

Analyses used three-level Hierarchical Linear Modeling (HLM) to describe predictors of quality of home visits (Visitor Strategies subscale of the HOVRS-A) and classrooms (Engaged Support for Learning and Emotional and Behavioral Support domains from the CLASS-T) using child, family, staff, classroom or home visit, and program characteristics. Because the Visitor Effectiveness subscale of the HOVRS-A measures a similar construct to one of our key outcomes variables—the engagement of the family in the visit— it was excluded as a predictor from this set of analyses.

In each model, level 1 represents how quality varies over time within individual staff members; level 2 describes how quality varies between different staff members within the same program; and level 3 describes how quality varies between staff members in different programs. These models account for the fact that multiple assessments of the same staff member over time are not independent observations. Analyses also adjusted for the shared experiences of staff members within the same program whose observed quality may be more similar than staff from different programs. Box IV.1 describes the covariates included at each level of the models. To examine factors that account for variability in home visit quality, analyses included staff and program characteristics as predictors in the models. The models also include characteristics of the visit and the child/family observed with the home visitor at each visit.

Box IV.1 HLM analyses predicting HOVRS-A and CLASS-T scores

The level-1 model estimates the variation in observed quality across years. Analyses included dummy variables for each year to capture any differences by year, with 2009 serving as the reference year in the HOVRS-A models and 2010 as the reference year in the CLASS-T models. As noted elsewhere, 2009 and 2010 are the first measurement points for the HOVRS-A and CLASS-T, respectively. Mean quality in the reference year serves as the intercept of the level-1 model.

The level-1 models for the HOVRS-A and CLASS-T include year-specific (age 1) depressive symptoms and job satisfaction for each staff member. In addition, the level-1 HOVRS-A models include time-varying home visit characteristics consisting of (1) the percentage of time spent on different activities (child-focused, family-focused, and parent-child activities, staff-family relationship building, and crisis management activities), (2) alignment with visit plan, (3) interference from environmental distractions, (4) languages used during the visit, and (5) presence of other children and/or adults during the visit. Child and family characteristics, used only in the HOVRS-A analyses, include whether the family enrolled at pregnancy, child age, birth weight, dual language learner (DLL) status, race/ethnicity, and maternal demographic risks (all measured at age 1). The level-1 model for the CLASS-T includes time-varying classroom characteristics, that is, staff-parent relationship, adult/child ratio in the classroom, group size, and percentage of DLLs in the classroom at each time point.

Additional year-specific (age 1) staff characteristics are examined in the level-2 models, including teacher's or home visitor's race/ethnicity; years of experience in Early Head Start; and whether the teacher or home visitor has a Bachelor's degree or higher, has a degree in early childhood education, has a CDA credential, ever received support from a mentor or coach, or speaks a language other than English.

The level-3 models include year-specific (age 1) program characteristics. Both the HOVRS-A and CLASS-T models include program approach (center-based versus multiple service options), program implementation status, and indicators of whether the program has a high percentage (more than 50 percent) of families who experience multiple demographic risks, who have mental health or substance abuse problems, or who live in unsafe neighborhoods or experience family violence.

Home visit content and characteristics are related to home visit quality (Visitor Strategies)

Several characteristics of the home visit are related to home visit quality as measured using the Visitor Strategies component of the HOVRS-A. For example, the percentage of visit time spent on different activities is related to Visitor Strategies scores (Table IV.1). Compared to a visit spent entirely on child-focused activities, the amount of time spent on staff-family relationship building is negatively associated with home visit quality. That is, when holding other characteristics constant, visits in which more time is spent on staff-family relationship building tend to be rated at lower quality on Visitor Strategies.[5] In contrast, spending more time on parent-child activities is positively associated with Visitor Strategies. The observed degree of alignment of the actual home visit with the home visitor's lesson plan is associated with higher Visitor Strategies ratings. The presence of at least one other adult (in addition to the child's parent) during the home visit is associated with higher Visitor Strategies ratings. The use of language other than English during the visit is not related to quality.

[5] The HOVRS-A was specifically designed to assess quality in a way that prioritizes the home visitor as facilitator of parent-child interaction. By definition, more time spent on family issues and interactions reduces the time and emphasis placed on the Visitor Strategies activities.

Table IV.1. Home visit, child, family, staff and program predictors of home visit quality

	HOVRS-A Visitor Strategies	
	Coefficient	Standard error
Intercept (mean quality in 2009)	1.86***	0.32
Home visit characteristics (vs. child-focused activities)		
Percentage of time spent on family-focused activities	-0.00	0.00
Percentage of time spent on parent-child activities	0.01***	0.00
Percentage of time spent on staff-family relationship building	-0.01**	0.00
Percentage of time spent on crisis management activities	0.01	0.00
Alignment with visit plan[a]	0.15***	0.04
Less interference from environmental distractions[b]	-0.01	0.02
Other children present	-0.09	0.07
Other adult(s) present	0.20*	0.09
Conducted in another language (vs. English only)	-0.06	0.14
Conducted in multiple languages (vs. English only)	0.23	0.18
Child and family characteristics		
Enrolled at pregnancy	0.01	0.08
Child age in months	-0.02	0.01
Child born with low or very low birth weight	0.23	0.12
Dual language learner	0.07	0.11
Race/ethnicity (vs. white)		
African American	-0.04	0.18
Hispanic	0.15	0.13
Other race	-0.07	0.13
Moderate maternal demographic risks[c]	0.16	0.09
High maternal demographic risks[c]	0.04	0.10
One psychological risk factor[d]	-0.04	0.09
Two or more psychological risk factors[d]	-0.02	0.13
Staff time-varying characteristics		
Staff job satisfaction	0.32**	0.10
Has any depressive symptoms	0.00	0.01
Staff characteristics		
Race/ethnicity (vs. white)		
African American	-0.24	0.16
Hispanic	-0.08	0.12
Other race	0.08	0.19
Speaks language other than English	0.02	0.1
Has a BA degree or higher	0.08	0.09
Years of experience in Early Head Start	0.01	0.01
Has a degree in early childhood	0.10	0.08
Has a CDA credential[e]	-0.01	0.09
Ever assigned a mentor or coach	0.10	0.08
Program characteristics		
Multiple approach[f]	0.15	0.13
Population served		
Over 50% of families with more than three demographic risks	-0.01	0.13
Over 50% of families with mental health or substance abuse problems	-0.08	0.17
Over 50% of families in unsafe neighborhoods or experiencing family violence	0.08	0.11
Fully implemented[g]	-0.10	0.11
Year effects		
Year 2010	0.04	0.08
Year 2011	0.07	0.09
Year 2012	-0.02	0.27
Sample size (home visitors)	**322**	

Source: Spring 2009, 2010, 2011 and 2012 home visit observations; parent interview; program director interview.

Note: Estimated using a three-level model with random effects for program and staff. Coefficients indicate differences in the home visit rating associated with each characteristic. For example, holding all other characteristics constant, a one percentage-point difference in amount of time spent on parent-child activities is associated with a 0.01 difference in Visitor Strategies ratings. The year effects capture any differences in quality by year with 2009 serving as the reference year.

[a] Assessed on a scale of 1 (not well aligned) to 5 (very well aligned).

[b] Original scale recoded so that scores range from 1 (not interfering) to 5 (very interfering). A score of 0 indicates there were no distractions during the visit.
[c] The maternal demographic risk is constructed by summing the number of the following risk factors that the mother faced: (1) being a teenage mother; (2) having no high school credential; (3) receiving public assistance; (4) not being employed or in school or training; and (5) being a single mother.
[d] Family psychological risk is based on moderate or severe depressive symptoms, high parenting stress, and current or past substance use problems.
[e] CDA = Child Development Associate
[f] Programs offering both center- and home-based services (as opposed to only one of the two) are designated as multiple approach.
[g] Programs with ratings equal to or above 3 on each of the implementation cornerstones are designated as "fully implemented."
*$p < .05$; **$p < .01$; ***$p < .001$

Only staff job satisfaction is related to home visit quality; child and program characteristics are unrelated to quality

Home visitors who are more satisfied with their jobs tend to have higher Visitor Strategies scores. No other differences were found in quality, including among home visitors with varying levels of education and experience, depressive symptoms, or experience working with a mentor or coach. Also, no evidence of differences was found in home visit quality by staff race and ethnicity.

None of the child characteristics examined is associated with home visit quality. Significant differences in home visit quality ratings between home visitors in programs offering both center- and home-based service options and home visitors in programs offering only a home-based approach did not emerge. We also found similar levels of home visit quality among programs that are rated as fully implemented and those that are not.

B. Predictors of classroom quality

Analyses examined the predictors of classroom quality, specifically, the domains of Emotional and Behavioral Support and Engaged Support for Learning of the CLASS-T. The results indicate that both domains of the CLASS-T vary significantly between teacher/classrooms and between programs, suggesting that the variation in classroom quality is attributable to both staff and classroom characteristics and program characteristics.

Several teacher characteristics are associated with classroom quality; program characteristics are not

Teacher education level is associated with classroom instructional quality. Having a BA degree or higher is significantly associated with higher scores in Engaged Support for Learning; however, it is not associated with Emotional and Behavioral Support.

Teacher depressive symptoms and the percentage of DLLs in the classroom are associated with the social-emotional quality of the classroom. Greater teacher depressive symptoms are significantly associated with lower scores in Emotional and Behavioral Support. Conversely, classrooms with greater concentrations of DLLs have higher scores in Emotional and Behavioral Support. Neither variable, however, predicts Engaged Support for Learning. Adult/child ratio and class size are not associated with either domain of classroom quality.

Table IV.2. Staff, classroom, and program predictors of classroom quality

	Engaged Support for Learning		Emotional and Behavioral Support	
	Coefficient	Standard error	Coefficient	Standard error
Intercept (mean classroom quality in 2010)	3.21***	0.33	4.93***	0.22
Staff characteristics				
Race/ethnicity (vs. white)				
African American	-0.08	0.19	-0.06	0.13
Hispanic	-0.16	0.21	0.04	0.12
Other race	-0.09	0.26	-0.09	0.18
Speaks language other than English	-0.08	0.16	0.13	0.09
Years of experience in Early Head Start	0.01	0.02	0.01	0.01
Has a BA degree or higher	0.27*	0.13	0.10	0.09
Has a degree in early childhood	0.05	0.12	0.02	0.08
Has a CDA credential	0.01	0.11	0.08	0.07
Ever assigned a mentor or coach	-0.08	0.14	-0.04	0.07
Staff/classroom time-varying characteristics				
Staff depressive symptoms	-0.01	0.01	-0.03**	0.01
Staff-parent relationship[a]	0.16*	0.07	0.12**	0.05
Staff job satisfaction	0.31*	0.15	0.24*	0.12
Adult/child ratio	0.10	0.08	-0.01	0.05
Class size	-0.04	0.04	-0.03	0.03
Percentage DLLs in the classroom	-0.00	0.00	0.003*	0.001
Program characteristics				
Multiple approach[b]	-0.06	0.22	0.06	0.12
Population Served				
Over 50% of families with more than three demographic risks	0.07	0.21	0.15	0.13
Over 50% of families with mental health or substance abuse problems	0.21	0.21	0.09	0.16
Over 50% of families in unsafe neighborhoods or experiencing family violence	-0.03	0.16	-0.17	0.12
Fully implemented	0.14	0.17	0.04	0.11
Year effects				
Year 2011	-0.34**	0.13	-0.05	0.08
Year 2012	-0.32*	0.15	0.02	0.13
Sample size (teachers)	**364**			

Source: Spring 2010 to spring 2012 classroom observation, staff interview, and program director interview

Note: Classroom observations were conducted with staff of both the 1-year-old cohort and newborn cohort children. The year effects capture any differences in quality by year with 2010 serving as the reference year.

[a] This represents the average staff-reported staff-parent relationship calculated across sample children in the classroom.
[b] Programs offering both center- and home-based services (as opposed to only one of the two) are designated as multiple approach.
*p < .05; **p < .01; ***p< .001
CDA = Child Development Associate

Teacher job satisfaction and teacher-parent relationships are associated with both instructional and social-emotional quality of classrooms. Staff members who report that they are very likely to continue working in early childhood education have higher scores in both Emotional and Behavioral Support and Engaged Support for Learning than those who do not. As rated by staff, the teacher-parent relationship is significantly and positively associated with higher scores in both Emotional and Behavioral Support and Engaged Support for Learning. That is, the higher the quality of teacher-parent relationships, the higher the scores in each of these areas. None of the program characteristics included in the models relates to classroom quality.

V. WHAT ARE THE ASSOCIATIONS BETWEEN EARLY HEAD START QUALITY AND CHILD AND PARENTING OUTCOMES? ARE THERE THRESHOLD EFFECTS?

Recent analyses have identified possible threshold effects of quality on preschool children's outcomes (Burchinal et al. 2011; 2010; under review). While these studies have focused on center-based preschool classrooms, none has examined threshold effects for home visit quality. Previous studies using the assumption of a linear relationship found modest associations between classroom quality and child outcomes. It is possible that the associations between quality and child outcomes are different above a certain threshold of quality. The generalized additive modeling (GAM) approach was used to identify empirically whether or not there were thresholds in home visit and classroom quality. Using HLM we also replicated analyses conducted as part of the Q-DOT Study that examined thresholds at or near cut points set by measure developers as denoting high quality. (Score distributions sometimes required the cut points to be shifted lower; Burchinal et al. under review; Zaslow et al. 2010). (See Box V.1 for descriptions of GAM and HLM analyses). For the classroom data, these analyses answered the question of whether quality at a relatively high level was associated with better outcomes. Given the interest in understanding the *minimum* quality threshold, significant associations between quality and child and parenting outcomes needed to be observed. Findings from the GAM analyses are summarized below. (Appendix tables summarize the findings from the analyses that use approaches similar to those used in the Q-DOT study.)

Box V.1. GAM and HLM analyses

GAM allows for empirical identification of thresholds without making any assumptions about whether the associations between quality and outcomes are linear or nonlinear. So, rather than focusing on the developers' definitions to identify thresholds, the GAM approach helps to decide where the thresholds should be set empirically. The GAM plots visually depict where the inflection points might be located. Analyses used thresholds based on the GAM plots, and then validated the selected thresholds using piecewise regression, which is the same approach used in Q-DOT.

In addition to the GAM analyses, the team also conducted three-level HLM using the Q-DOT approaches and cut points, adjusting for child and family characteristics (models controlled for cohort, age as of spring 2009, gender, race/ethnicity, DLL status, family poverty ratio, maternal demographic risk composite) as well as clustering of children within classrooms/home visitors and classrooms/home visitors within programs. Quadratic regressions were conducted that included the linear and quadratic terms of quality as predictors. The quadratic term was dropped from the model if it was not statistically significant. Although the focus is on the GAM results in the report, when the HLM models differ from the GAM findings, it is highlighted. Linear relationships are reported when no threshold was identified.

The focus was on child and parenting outcomes at ages 2 and 3 in relation to quality observed at the same age. The child outcomes cover language and social-emotional domains: the PPVT-4 and PLS-4 from direct child assessments; parent- and staff-reported English CDI; the Emotional Regulation and Orientation/Engagement subscales of the BRS, rated by assessors; and the BITSEA Problem and Competence domains reported by parents and staff (see Table II.1). Analyses of home visit quality also include Parenting Synchronicity, parent support for child learning, and parent provision of learning materials as family or parenting outcomes. Models of child and parenting

outcomes at each age were estimated, controlling for their status in the prior year.[6] A *p*-level less than 0.05 is considered to indicate that the association between quality and outcomes is statistically significant. The results for the home visit and classroom observation analyses are summarized in Tables V.1 and V.2, respectively.

A. Home visit quality

Looking first at one component of home visit quality (the Visitor Strategies subscale of the HOVRS-A),[7] analyses identified several threshold effects. For example, analyses identify threshold effects for staff-reported BITSEA Problem Behavior scores at age 2. The Visitor Strategies threshold is at 3 based on the plot (Figure V.1). A validation test confirms this significant association. This suggests that the association between Visitor Strategies and children's problem behaviors is stronger for high quality (above a 3) rather than for low quality home visiting. Increases in Visitor Strategies scores are associated with lower problem behaviors for high quality home visits. The association is reversed for low quality home visiting.

Evidence was found of linear associations between Visitor Strategies and ASQ-3 Problem Solving and Parenting Synchronicity at age 2, but in unexpected directions. Better Visitor Strategies are associated with lower child Problem Solving and less Parenting Synchronicity. No other thresholds or linear relationships were found for the other outcomes (see Table V.1 and Appendix Table A.1).

[6] For age 2 child outcomes, the study team controlled for the corresponding age 1 scores. For age 2 parenting outcomes, the team controlled for Parenting Stress Index scores at age 1. For age 3 child and parenting outcomes, the team controlled for the corresponding age 2 scores.

[7] Because the Visitor Effectiveness subscale of the HOVRS-A can be interpreted as an outcome measure rather than a quality indicator, it was excluded from this set of analyses.

Figure V.1. Age 2 HOVRS-A Home Visitor Strategies threshold with age 2 BITSEA Problem Behaviors

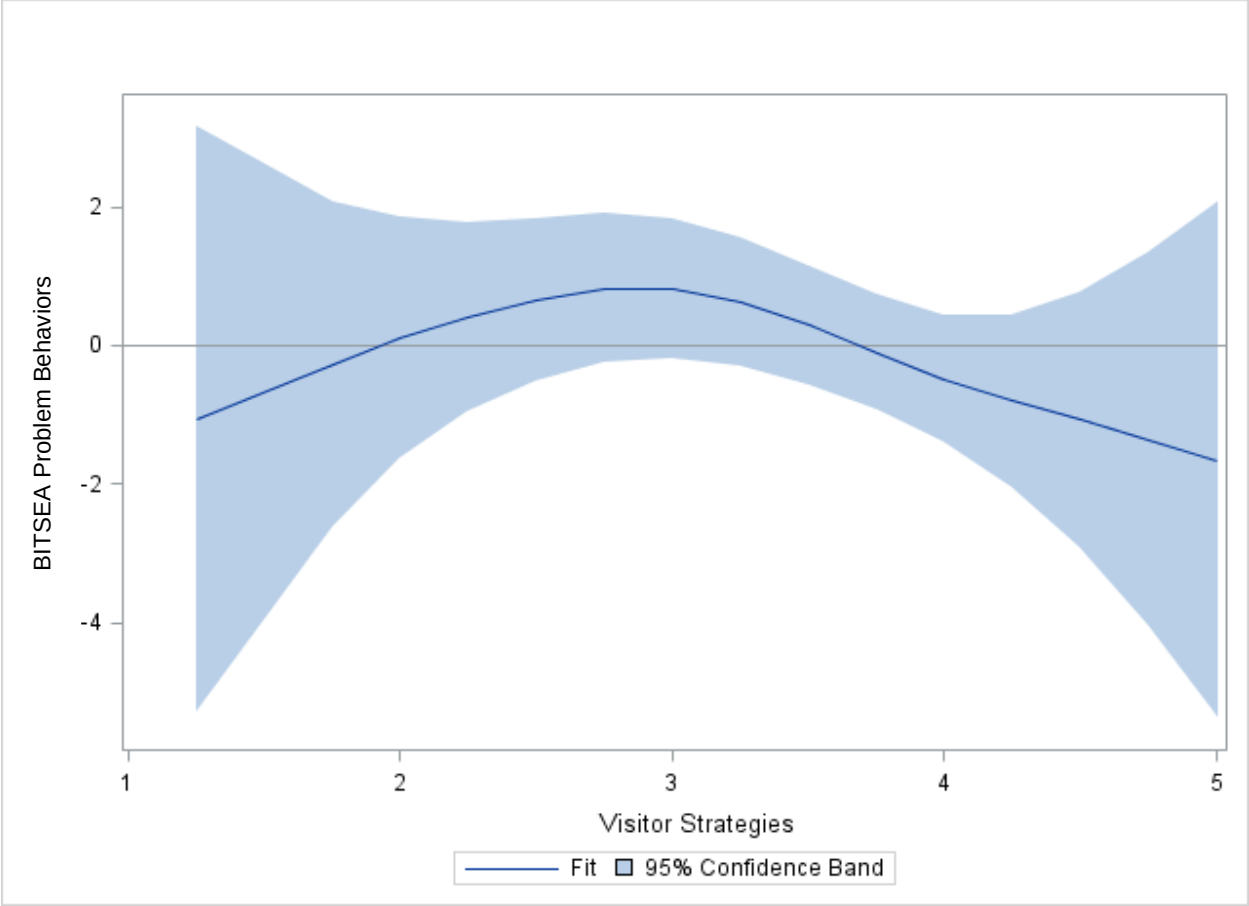


Table V.1. Summary of GAM results for home visit quality

	HOVRS-A Visitor Strategies	
	Age 2	Age 3
PLS-4	ns	ns
PPVT-4	NA	ns
Parent-reported ASQ-3 Problem Solving	Linear (-)	ns
BRS Emotional Regulation	ns	ns
BRS Orientation/Engagement	ns	ns
Staff-Reported BITSEA		
Competence	ns	ns
Problem	Cut point 3 Validation test sig.	ns
Parenting Outcomes		
Synchronicity	Linear (-)	ns
Parent support for child learning	NA	ns
Parent provision of learning materials	NA	ns

Source: Baby FACES direct child assessment, Parent Interview, Staff Child Report from age 1 to age 3, and Home Visit Observation

Notes: Analyses are weighted to be representative of children who were enrolled in Early Head Start in spring 2009 and continued to be enrolled at age 2 or age 3.

NA = Not applicable
ns = non-significant

B. Classroom quality

Looking next at classroom quality, evidence of threshold effects are found again. The results indicate there are threshold effects for the Emotional and Behavioral Support domain with BRS Orientation/Engagement scores at age 2 and staff-reported BITSEA Problem Behavior scores at age 3. The Emotional and Behavioral Support threshold is at 5 in the association with BRS Orientation/Engagement scores at age 2 (Figure V.2) and with the BITSEA Problem Behavior scores at age 3. Validation tests are also significant. The GAM plot also suggests thresholds for Emotional and Behavioral Support in the association with BITSEA Problem Behavior scores at age 2; however, the validation test is not significant. These findings are somewhat consistent with the results from the Q-DOT study, suggesting that Emotional and Behavioral Support quality above a 5 is more strongly associated with social-emotional child outcomes and is the minimum level at which significant associations may be found. However, regarding children's language skills, only a positive, linear relationship between Emotional and Behavioral Support and PLS-4 scores at age 2 is found. Children in classrooms with higher social-emotional quality have stronger language outcomes (see Table V.2 and Appendix Table A.3).

Turning to the Engaged Support for Learning domain, analyses identified threshold effects at age 2 for the PLS-4. At age 3, there are threshold effects for the PLS-4, parent-reported ASQ-3 Problem Solving, and BRS Emotional Regulation and Orientation/Engagement subscale, all based on GAM plots.[8] Validation tests are significant. Figure V.3 shows the association with ASQ-3 Problem Solving scores at age 3. The GAM plot also indicates a threshold for Engaged Support for Learning in the association with BITSEA Problem Behavior scores at age 2; however, the validation tests are not significant. Aside from the associations with staff-reported BITSEA scores, threshold findings for Engaged Support for Learning are largely consistent with the results from the Q-DOT approach where the cut point set and tested was at 4.0. These findings suggest that Engaged Support for Learning quality above a 4 is more strongly associated with child outcomes, but quality as low as a 3 may also be associated with child outcomes.

Analyses also found evidence of linear associations. The results show a positive linear association of Engaged Support for Learning with teacher-reported BITSEA Competence scores at age 2. These findings suggest that higher levels of Engaged Support for Learning are associated with better child behavior outcomes.

[8] At age 2, the threshold is at 3.5 or 4 for relationship with the PLS-4. At age 3, it is 3.5 or 4 for relationships with the ASQ-3 Problem Solving and BRS Emotional Regulation, and 3 for relationships with the PLS-4 and the BRS Orientation/Engagement scores.

Figure V.2. Age 2 CLASS-T Emotional and Behavioral Support threshold with age 2 BRS Orientation/Engagement

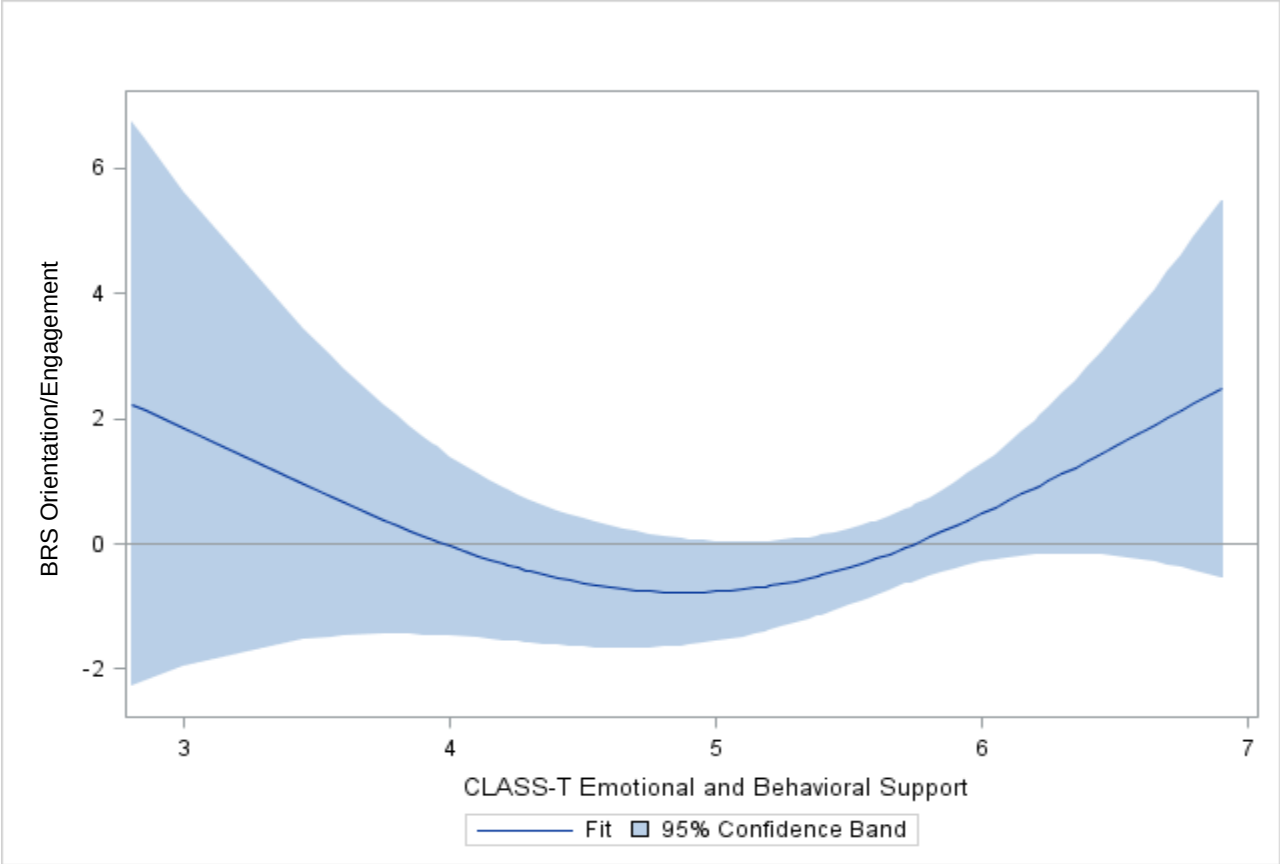


Table V.2. Summary of GAM results for the CLASS-T

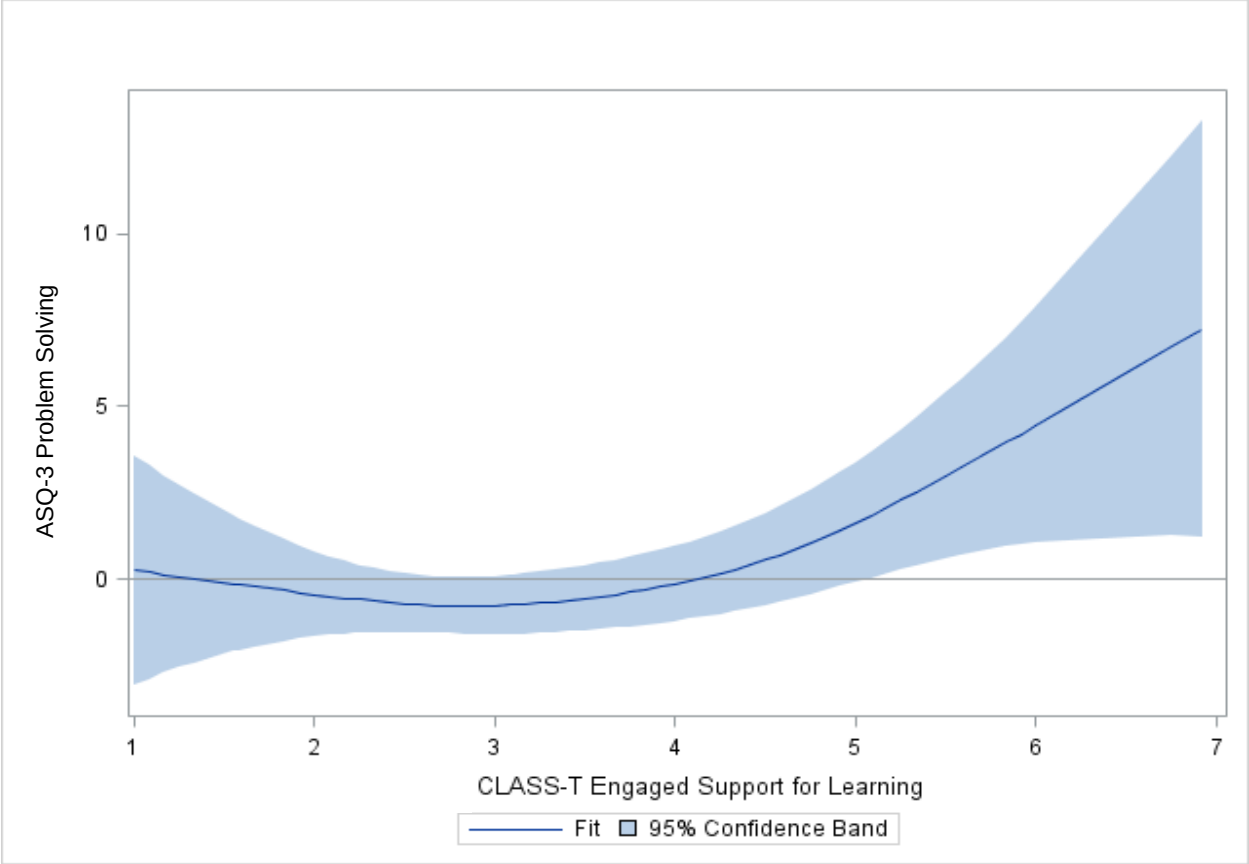
	Emotional and Behavioral Support		Engaged Support for Learning	
	Age 2	Age 3	Age 2	Age 3
PLS-4	Linear (+)	ns	Cut point at 3.5 or 4 Validation test sig.	Cut point 3 Validation test sig.
PPVT-4	NA	ns	NA	ns
Parent-reported ASQ-3 Problem Solving	ns	ns	ns	Cut point 4 Validation test sig.
BRS Emotional Regulation	ns	ns	ns	Cut point 3.5 or 4 Validation test sig.
BRS Orientation/ Engagement	Cut point 5 Validation test sig.	ns	ns	Cut point 3 Validation test sig.
Staff-Reported BITSEA Competence	ns	ns	Linear (+)	ns
Problem	Cut point 4.5 or 5 Validation test ns	5 Validation test sig.	Cut point 3.5 or 4 Validation test ns	ns

Source: Baby FACES Direct Child Assessment, Parent Interview, Staff Child Report from age 1 to age 3, and Classroom Observation

Notes: Analyses are weighted to be representative of children who were enrolled in Early Head Start in spring 2009 and continued to be enrolled at age 2 or age 3.

NA = Not applicable
ns = non-significant

Figure V.3. Age 3 CLASS-T Engaged Support for Learning threshold with age 3 ASQ-3 Problem Solving



VI. CONCLUSIONS AND IMPLICATIONS

The provision of high quality services is an important component of Early Head Start's conceptual framework. However, very limited information is available on the stability, predictors, and thresholds of home visit and classroom quality. The present findings demonstrate that that Early Head Start home visit and classroom quality is in the moderate or mid-range, with patterns similar to those found in the broader literature using similar measures (Thomason and LaParo 2009). Quality scores are relatively stable over time, with one exception—the quality of Engaged Support for Learning decreases over time. The findings suggest there is room for improvement efforts focused on (1) improving lower- quality home visits and classrooms, as well as (2) encouraging continuous improvement within the higher-quality range. In particular, it appears that teachers are strongest in interactions that provide emotional and behavior support, but their instructional support to children can be bolstered. Teachers may benefit from ongoing support to maintain and improve instructional quality.

The findings also show that there is significant variability in home visit and classroom quality across teacher and home visitors that is captured by several characteristics. Indeed, teacher education, job satisfaction, depressive symptoms, and relationships with parents are all related to classroom quality in the expected direction. The percentage of DLL children in the classroom is also positively associated with the social-emotional aspects of classroom quality. Program characteristics are not associated with classroom quality. These findings suggest that program leadership/supervisors may want to consider providing training on supporting teacher-parent relationships and searching for ways to identify and support teachers who may be struggling with depressive symptoms or job satisfaction.

Unlike findings for classroom quality, few characteristics predict Early Head Start home visit quality. Home visit quality does not differ by staff education, experience, mental health status, or race and ethnicity. However, home visitors who are more satisfied with their jobs tend to have higher Visitor Strategies scores. Child and program characteristics are unrelated to home visit quality. This is not to conclude that program and staff characteristics are not important factors related to home visit quality. It is quite possible that unmeasured features of programs and staff are more predictive of service quality. Aspects of home visits themselves appear to be particularly important to home visit quality as measured by the HOVRS-A. Higher quality was observed when more time was spent on parent-child activities (relative to child-focused activities), when visit activities aligned with the visit plan, and when more time was spent on staff-family relationship building. Findings indicate that home visitors may benefit from use of a curriculum to plan and execute home visits. Supervisors may find it useful to ask about home visitor practices for planning the content of visits and examine the content and characteristics of visits during observations.

It is important to note that the associations explored in the current analyses are not causal and thus do not suggest that classroom quality will improve as a result of modifying teacher and home visitor, classroom, or program characteristics. However, coupled with the growing body of research in this area, the findings may provide one source of information to help programs decide where to target efforts to potentially sustain and enhance home visit and classroom quality. Furthermore, the findings suggest there may be different factors to consider when tackling approaches to improve home visit versus classroom quality.

Lastly, these results demonstrate that the relationship between quality and child and parenting outcomes is complex. The analyses find both linear associations and threshold effects between quality and child and parenting outcomes and, in some instances, no association between quality and outcomes. Though there is no consistent pattern of threshold effects, there appear to be more significant associations for classroom quality than for home visit quality. In fact, some of the home visit associations are in counterintuitive directions (including relationships with parenting outcomes). However, findings generally suggest that low quality classroom and home visit interactions are not necessarily linked to children's development, positively or negatively. For quality in the higher range, there are stronger associations between quality and better child outcomes than for quality in the lower range but only for certain aspects of quality and certain outcomes. This is particularly true for the association between Engaged Support for Learning and some of the language and social emotional outcomes primarily at age 3. For most of the results, the identified thresholds are stable across two analytic approaches but not stable across age. In general, these findings are somewhat consistent with the research literature that examines the threshold effects in the relationships between classroom quality and preschool children's outcomes. However, more research is needed before recommendations can be made about whether quality improvement efforts should aim to help lower quality settings achieve a particular quality cut point. In addition, more work is needed to help ensure the early childhood community more fully understands this area of childhood education.

REFERENCES

Administration for Children and Families. "Head Start Program Performance Standards." Final rule cited in *Federal Register,* November 5, 1996. Available at www.acf.hhs.gov/programs/hsb/performance. Accessed January 18, 2010.

Administration on Children, Youth, and Families. "Building Their Futures: How Early Head Start Programs Are Enhancing the Lives of Infants and Toddlers in Low-Income Families." Washington, DC: U.S. Department of Health and Human Services, June, 2001.

Aikens, N., L. Tarullo, L. Hulsey, C. Ross, J. West, and Y. Xue. "A Year in Head Start: Children, Families and Programs." Report submitted to Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services. Washington, DC: Mathematica Policy Research, October 2010.

Astuto, J., and L. Allen. "Home Visitation and Young Children: An Approach Worth Investing In?" *Social Policy Report*, vol. 23, no. 4, 2009.

Bayley, N. *Bayley Scales of Infant Development, Second Edition, Manual*. New York: The Psychological Corporation, Harcourt Brace and Company, 1993.

Belsky, J., and M. Pluess. "Differential Susceptibility to Long-Term Effects of Quality of Child Care on Externalizing Behavior in Adolescence?" *International Journal of Behavioral Development,* vol. 36, no. 1*,* 2012, pp. 2–10.

Bogard, K., F. Traylor, and R. Takanishi. "Teacher Education and PK Outcomes: Are We Asking the Right Questions?" *Early Childhood Research Quarterly*, vol. 23, no. 1, 2008, pp. 1–6.

Boller, K., C. Vogel, R. Cohen, N. Aikens, and K. Hallgren. "Home Visit Characteristics and Content Form." Princeton, NJ: Mathematica Policy Research, 2009.

Briggs-Gowan, M.J. *BITSEA Brief Infant-Toddler Social and Emotional Assessment Examiner's Manual*. San Antonio, TX: Harcourt Assessment, 2006.

Brookes, S., J. Ispa, J. Summers, K. Thornburg, and V. Lane. "Building Successful Home Visitor-Mother Relationships and Reaching Program Goals: A Qualitative Look at Contributing Factors." *Early Childhood Research Quarterly,* vol. 21, no. 1, 2006, pp. 25–45.

Burchinal, M.R., J.E. Roberts, L.A. Nabors, and D.M. Bryant. "Quality of Center Child Care and Infant Cognitive and Language Development." *Child Development,* vol. 67, 1996, pp. 606–620.

Burchinal, M., J. Roberts, R. Riggins, S. Zeisel, E. Neebe, and D. Bryant. "Relating Quality of Center-Based Child Care to Early Cognitive and Language Development Longitudinally." *Child Development*, vol. 71, 2000, pp. 339–357.

Burchinal, M., K. Kainz, K. Cai, K. Tout, M. Zaslow, I. Martinez-Beck, and C. Rathgeb. "Child Care Quality and Child Outcomes: Multiple-Studies Analyses." Presentation at the Working Meeting of Developing the Next Wave of Quality Measures for Early Childhood and School-Age Programs, January 2008a.

Burchinal, M., C. Howes, R. Pianta, D. Bryant, D. Early, R. Clifford, and O. Barbarin. "Predicting Child Outcomes at the End of Kindergarten from the Quality of Pre-Kindergarten Teacher-Child Interactions and Instruction." *Applied Developmental Science*, vol. 12, no. 3, 2008b, pp. 140–153.

Burchinal, P., K. Kainz, K. Cai, K. Tout, M. Zaslow, I. Martinez-Beck, and C. Rathgeb. "Early Care and Education Quality and Child Outcomes." Research-to-Policy Research-to-Practice Brief: OPRE Research-to-Policy Brief #1. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, HHS, and Child Trends, 2009, and Publication 2009-15. Washington, DC: Administration for Children and Families, Office of Planning, Research and Evaluation, May 2009.

Burchinal, M., N. Vandergrift, R. Pianta, and A. Mashburn. "Threshold Analysis of Association Between Child Care Quality and Child Outcomes for Low-Income Children in Pre-Kindergarten Programs." *Early Childhood Research Quarterly*, vol. 25, no. 2, pp.166–176, 2010.

Burchinal, M., K. Kainz, and Y. Cai. "How Well Do Our Measures of Quality Predict Child Outcomes? A Meta-Analysis and Coordinated Analysis -of Data from Large-Scale Studies of Early Childhood Settings." edited by M. Zaslow. *Reasons to Take Stock and Strengthen Our Measures of Quality*. Baltimore, MD: Brooks Publishing, 2011a.

Burchinal, Margaret, Yange Xue, and Andrew 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.

Burchinal, Margaret, Yange Xue, Louisa Tarullo, Marty Zaslow, Hsiao Tien, Anamarie Auger, Elizabeth Cavadel, Andrew Mashburn, and Ivelisse Martinez-Beck. "Child Care Quality and Dosage Thresholds in Association with Child Outcomes: Secondary Data Analyses." *Monographs of the Society for Research in Child Development,* under review.

Clarke-Stewart, K.A. "Infant Day Care: Maligned or Malignant?" *American Psychologist,* vol. 44, no. 2, 1989, pp. 266–273.

Del Grosso, P., R. Kleinman, A. Esposito, E. Martin, and D. Paulsell. "Assessing the Evidence of Effectiveness of Home Visiting Program Models Implemented in Tribal Communities." Princeton, NJ: Mathematica Policy Research, February 2011.

de Schipper, E. J., Riksen-Walraven, J., and S. Geurts. "Effects of Child–Caregiver Ratio on the Interactions Between Caregivers and Children in Child-Care Centers: An Experimental Study." *Child Development*, vol. 77, 2006, pp. 861–874.

Dunn, L. and D. Dunn. *Peabody Picture Vocabulary Test—Fourth edition.* Bloomington, MN: NCS Pearson, Inc., 2007.

Early, D., D. Bryant, R. Pianta, R. Clifford, M. Burchinal, S. Ritchie, C. Howes, and O. Barbarin. "Are Teachers' Education, Major, and Credentials Related to Classroom Quality and Children's Academic Gains in Pre-Kindergarten?" *Early Childhood Research Quarterly*, vol. 21, 2006, pp. 174–195.

Elicker, J., I. Noppe, L. Noppe, and C. Fortner-Wood. "The Parent-Caregiver Relationship Scale: Rounding Out the Relationship System in Infant Care." *Early Education and Development,* vol. 8, no. 1, 1997, pp. 83–100.

Gerber, E., M. Whitebook, and R. Weinstein. "At the Heart of Child Care: Predictors of Teacher Sensitivity in Center-Based Child Care." *Early Childhood Research Quarterly*, vol. 22, no. 3, 2007, pp. 327–346.

Hallgren, K., N. Aikens, K. Boller, J. Stein, L. Roggman, G. Cook, and M. Innocenti. "Home Visit Rating Scales-Adapted (HOVRS-A) Manual." Princeton, NJ: Mathematica Policy Research, 2009.

Hayes, C.D., J.L. Palmer, and M.J. Zaslow. "Who Cares for America's Children? Child Care Policy for the 1990's." Washington, DC: American Family Press, 1990.

Howes, C., M. Burchinal, R. Pianta, D. Bryant, D. Early, R. Clifford, and O. Barbarin. "Ready to Learn? Children's Pre-Academic Achievement in Pre-Kindergarten Programs." *Early Childhood Research Quarterly,* vol. 23, 2008, pp. 27-50.

Hulsey, Lara, Nikki Aikens, Yange Xue, Louisa Tarullo, and Jerry West. "Data Tables for FACES 2006: A Year in Head Start Report." 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.

Keys, T., G. Farkas, M. Burchinal, G. Duncan, D. Vandell, W. Li, E. Ruzek, and C. Howes. "Preschool Center Quality and School Readiness: Quality Effects and Variation by Demographic and Child Characteristics." *Child Development*, vol. 84, no. 4, 2013, pp. 1171–1190.

Kisker, E.E., K. Boller, C.T. Nagatoshi, C. Sciarrino, V. Jethwani, T. Zavitsky, M. Ford, and J.M. Love. "Resources for Measuring Services and Outcomes in Head Start Programs Serving Infants and Toddlers." Princeton, NJ: Mathematica Policy Research, 2003.

Lamb, M. E.. Nonparental Child Care: Context, Quality, Correlates, and Consequences. Edited by W. Damon (series ed.) and I. Sigel and K. Renninger (vol. eds.), Handbook of Child Psychology, Child Psychology in Practice, vol. 4, 5th edition, pp. 73–133. New York: Wiley, 1998.

Love, J., E. Kisker, C. Ross, H. Raikes, J. Constantine, K. Boller, J. Brooks-Gunn, R. Chazan-Cohen, L. Banks Tarullo, C. Brady-Smith, A. Fuligni, P. Schochet, D. Paulsell, and C. Vogel. "The Effectiveness of Early Head Start for 3-Year-Old Children and Their Parents: Lessons for Policy and Programs." *Developmental Psychology,* vol. 41, no. 6, 2005, pp. 885–901.

Moiduddin, E., N. Aikens, L. Tarullo, J. West, and Y. Xue. "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, 2012.

National Institute of Child Health and Human Development Early Child Care Research Network. "Early Child Care and Children's Development Prior to School Entry: Results from the NICHD Study of Early Child Care. Rockville, MD: NICHD, 1996.

National Institute of Child Health and Human Development Early Child Care Research Network. "Early Child Care and Children's Development Prior to School Entry: Results from the NICHD Study of Early Child Care. Rockville, MD: NICHD, 1997.

National Institute of Child Health and Human Development Early Child Care Research Network. "Early Child Care and Children's Development Prior to School Entry: Results from the NICHD Study of Early Child Care. Rockville, MD: NICHD, 1998.

National Institute of Child Health and Human Development Early Child Care Research Network. "Early Child Care and Children's Development Prior to School Entry: Results from the NICHD Study of Early Child Care. Rockville, MD: NICHD, 2006.

National Institute of Child Health and Human Development Early Child Care Research Network and G. J. Duncan. "Modeling the Impacts of Child Care Quality on Children's Preschool Cognitive Development." *Child Development,* vol. 74, no. 5, 2003, pp. 1454–1475.

Pianta, R.C., K.M. La Paro, and B.K. Hamre. *Classroom Assessment Scoring System (CLASS). Pre-K Version.* Baltimore: Brookes, 2008.

Pianta, R.C., K.M. La Paro, and B.K. Hamre. *Classroom Assessment Scoring System Pilot Toddler (CLASS-T) Manual.* Charlottesville, VA: Teachstone, Inc., 2010a.

Pianta, R.C., K.M. La Paro, and B.K. Hamre. *Classroom Assessment Scoring System Toddler (CLASS-T) Manual.* Charlottesville, VA: Teachstone, Inc., 2010b.

Paulsell, Diane, Kimberly Boller, Kristin Hallgren, and Andrea Mraz Esposito. "Assessing Home Visiting Quality: Dosage, Content, and Relationships." *ZERO TO THREE,* vol. 30, no. 2, July 2010, pp. 16-21.

Phillips, D., D. Mekos, S. Scarr, K. McCartney, and M. Abbott-Shim. "Within and Beyond the Classroom Door: Assessing Quality in Child Care Centers." *Early Childhood Research Quarterly*, vol. 15, 2000, pp. 475–496.

Phillipsen, L., M. Burchinal, C. Howes, and D. Cryer. "The Prediction of Process Quality from Structural Features of Child Care." *Early Childhood Research Quarterly*, vol. 12, 1997, pp. 281–303.

Radloff, L.S. "The CES-D Scale: A Self-Report Depression Scale for Research in the General Population." *Applied Psychological Measurement,* vol. 1, 1977, pp. 385–401.

Raikes, H., B. Green, J. Atwater, E. Kisker, J. Constantine, and R. Chazan-Cohen. "Involvement in Early Head Start Home Visiting Services: Demographic Predictors and Relations to Child and Parent Outcomes." *Early Childhood Research Quarterly,* vol. 21, no. 1, 2006, pp. 2–24.

Resnick, Gary, and Nicholas Zill. "Understanding Quality in Head Start Classrooms: The Role of Teacher and Program-Level Factors." Presented at the Biennial Meeting of the Society for Research in Child Development, Tampa, FL, April 2003.

Roggman, L., G. Cook, C. Peterson, and H. Raikes. "Who Drops Out of Early Head Start Home Visiting Programs?" *Early Education and Development*, vol. 19, no. 4, 2008a, pp. 574–599.

Roggman, L., G. Boyce, and M. Innocenti. *Developmental Parenting: A Guide for Early Childhood Practitioners*. Baltimore, MD: Paul H. Brooks, 2008b.

Roggman, L., G. Cook, V. Jump, M. Innocenti, K. Christiansen, L. Boyce, , N. Aikens, K. Boller, D. Paulsell, and K. Hallgren. *Home Visit Rating Scales-Adapted (HOVRS-A)*. Unpublished document. Princeton, NJ: Mathematica Policy Research, 2009.

Roggman, L., K. Christiansen, G. Cook, V. Jump, L. Boyce, and C. Peterson. "Home Visits: Measuring How They Work." Paper presented at the Early Intervention Research Institute Mini-Conference, Logan, UT, 2006.

Ross, C., J. Mirowsky, and J. Huber. "Dividing Work, Sharing Work, and in Between: Marriage Patterns and Depression." *American Sociological Review,* vol. 48, 1983, pp. 809–823.

Ross, C., E. Moiduddin, C. Meagher, and B. Carlson. "The Chicago Program Evaluation Project: A Picture of Early Childhood Programs, Teachers, and Preschool-Age Children in Chicago." Final external report submitted to the Erickson Institute, Chicago Department of Children and Youth Services, and Chicago Public Schools. Princeton, NJ: Mathematica Policy Research, December 2008.

Ruopp, R., J. Travers, F. Glantz, and C. Coelen. "Children at the Center." Cambridge, MA: Abt Associates, 1979.

Scarr, S., M. Eisenberg, and K. Deater-Deckard. "Measurement of Quality in Child Care Centers." *Early Childhood Research Quarterly,* vol. 9, 1994, pp. 131–151.

Squires, J., E. Twombly, D. Bricker, and L. Potter. *ASQ-3 User's Guide*. Baltimore, MD: Paul H. Brookes, 2009.

Sweet, Monica A., and Mark I. Appelbaum. "Is Home Visiting an Effective Strategy? A Meta-Analytic Review of Home Visiting Programs for Families with Young Children." *Child Development,* vol. 75, no. 5, 2004, pp. 1435–1456.

Thomason, A.C., and K.M. La Paro. "Measuring the Quality of Teacher-Child Interactions in Toddler Child Care." *Early Education and Development*, vol. 20, no. 2, 2009, pp. 285–304.

Torquati, J., H. Raikes, G. Welch, J. Ryoo, and X. Tu. *Testing Thresholds of Quality on Child Outcomes Globally and in Subgroup*s. Presented at State and Territory CCDF Administrators Meeting (STAM 2011)—Annual Meeting of the Child Care Policy Research Consortium (CCPRC 2011), Bethesda, Maryland, November, 2011.

Vandell, D., and B. Wolfe. "Child Care Quality: Does It Matter and Does It Need To Be Improved?" Washington, DC: Department of Health and Human Services, 2000.

Vandell, D., J. Belsky, M. Burchinal, L. Steinberg, N. Vandergrift, and NICHD Early Child Care Research Network. "Do Effects of Early Child Care Extend to Age 15 Years?" *Child Development,* vol. 81*,* 2010, pp. 737–756.

Vogel, C., K. Boller, Y. Xue, R. Blair, N. Aikens, A. Burwick, Y. Shrago, B. Carlson, L. Kalb, L. Mendenko, J. Cannon, S. Harrington, and J. Stein. "Learning As We Go: A First Snapshot of Early Head Start Programs, Staff, Families, and Children." OPRE Report #2011-7, Washington, DC: Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services, 2011.

Whitebook, M., C. Howes, and D. Phillips. "Who Cares? Child Care Teachers and the Quality of Care in America." Final Report of the National Child Care Staffing Study. Oakland, CA: Child Care Employee Project, 1989.

Whitaker, J.E., and B.J. Harden. "Beyond ABCs and 123s: Enhancing Teacher-Child Relationship Quality to Promote Children's Behavioral Development." *National Head Start Association Dialog*, vol. 13, no. 3, 2010, pp. 185–191.

Xue, Yange, Margaret Burchinal, Anamarie Auger, Hsiao-Chuan Tien, and Louisa Banks Tarullo. "Dosage Effects in Early Intervention and Child Care Quality: Evidence from Secondary Data Analysis." Paper presented at the Society for Research in Child Development Biennial Meeting, Montreal, April 2011.

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, OPRE 2011-5." Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research and Evaluation, 2010.

Zaslow, Martha J. "Variation in Child Care Quality and Its Implications for Children." *Journal of Social Issues,* vol. 47, no. 2, 1991, pp. 125–138.

Zimmerman, I.L., V.G. Steiner, and R.E. Pond. *Preschool Language Scale, Fourth Edition (PLS-4)*. San Antonio, TX: Pearson, 2002a.

Zimmerman, I.L., V.G. Steiner, and R.E. Pond. *Preschool Language Scale, Fourth Edition (PLS-4) Spanish Edition*. San Antonio, TX: Pearson, 2002b.

APPENDIX A.

ASSOCIATIONS AND THRESHOLD EFFECTS FOR QUALITY MEASURES AND CHILD OUTCOMES FROM MULTIVARIATE HLM MODELS

When investigating the relationship between quality and child outcomes using Baby FACES data, we first explored whether the relationship is linear or non-linear. We conducted three-level HLMs adjusting for child and family characteristics as well as clustering of children within classrooms and classrooms within programs. We conducted quadratic regressions that included the linear and quadratic terms of quality as predictors. The quadratic term was dropped from the model if it was not statistically significant. We then conducted spline regression analyses to examine whether there are thresholds in the quality-outcome associations, that is, whether the associations between quality and outcomes are different in higher- than in lower-quality classrooms.

The column of "Test of Associations" in the tables presents the results from quadratic regressions. For each outcome, the first row shows the coefficients for the linear and quadratic associations, respectively; if the quadratic term was not statistically significant, it was dropped from the model and the result of the linear association was presented in the second row for that outcome measure. The column of "Test for Threshold" shows the results from spline regressions and the regression coefficients for the lower and higher range of quality were bolded if they were significantly different. The cut-points for quality measures in these analyses are 4 for HOVRS-A Visitor Strategies, 5 for Emotional and Behavioral Support, and 4 for Engaged Support for Learning.

Table A.1. Associations between Home Visitor Strategies and child outcomes at age 2 and tests for threshold effects

	Test of associations		Test for threshold[b]	
	Linear	Quadratic[a]	Low quality	High quality
PLS-4	10.89 (13.12)	-1.81 (2.09)	1.23 (1.39)	-5.16 (11.74)
	-0.37 (1.81)			
ASQ-3 Problem Solving	-7.37 (5.53)	0.91 (0.87)	0.23 (0.55)	-1.59 (4.80)
	-1.63* (0.76)			
BRS Emotional Regulation	4.87 (6.75)	-0.86 (1.07)	0.51 (0.61)	-3.08 (5.94)
	-0.51 (0.86)			
BRS Orientation/Engagement	4.20 (6.05)	-0.69 (0.96)	0.03 (0.56)	-3.86 (5.39)
	-0.15 (0.77)			
Staff-reported BITSEA Competence	-2.73 (2.73)	0.41 (0.43)	-0.34 (0.26)	-2.26 (2.47)
	-0.16 (0.37)			
Staff-reported BITSEA Problem	4.51 (3.93)	-0.77 (0.62)	0.58 (0.38)	0.48 (3.42)
	-0.39 (0.53)			
Parenting Synchronicity	-0.89 (0.69)	0.10 (0.11)	0.02 (0.07)	-0.30 (0.67)
	-0.26** (0.09)			

Source: Baby FACES Direct Child Assessment, Staff Interview, and Parent Interview and Home Visit Observation.

Note: Analyses are weighted to be representative of children who were enrolled in Early Head Start in spring 2009 and continued to be enrolled at age 3. Estimates from mixed models in SAS. We used multiple imputation (N=10) to handle missing data. Covariates include cohort, child age, gender, race, home language, maternal demographic risk, poverty ratio, and child's age 1 scores.

[a] Quadratic term was dropped from the model if not significant, and results without the quadratic term are presented in the second row for each outcome measure.

[b] Two separate variables representing quality in the low and high range are included in the model. The coefficients are bolded if the associations between quality and child outcome are different for quality in the higher versus lower range ($p < .10$). Cut-point for quality measures in these analyses is 4 for Visitor Strategies.

†$p < .10$; *$p < .05$; **$p < .01$; ***$p < .001$.

Table A.2. Associations between Home Visitor Strategies and child outcomes at age 3 and tests for threshold effects

	Test of associations		Test for threshold[b]	
	Linear	Quadratic[a]	Low quality	High quality
PLS-4	-13.21 (12.12)	2.30 (2.02)	-1.13 (1.68)	4.27 (14.16)
	0.44 (1.64)			
PPVT-4	9.37 (12.72)	-1.10 (2.10)	-2.81† (1.60)	-15.55 (14.25)
	2.80 (1.77)			
ASQ-3 Problem Solving	-12.24* (6.03)	1.79† (0.98)	-0.80 (0.76)	-7.14 (6.40)
BRS Emotional Regulation	-2.46 (5.93)	0.31 (0.97)	-0.47 (0.77)	-9.66 (6.63)
	-0.50 (0.89)			
BRS Orientation/Engagement	-1.03 (4.86)	0.07 (0.80)	-0.41 (0.63)	-6.44 (5.38)
	-0.61 (0.72)			
Staff-reported BITSEA Competence	0.79 (1.76)	-0.19 (0.28)	-0.01 (0.22)	-2.82† (1.70)
	-0.39 (0.27)			
Staff-reported BITSEA Problem	-2.26 (2.82)	0.33 (0.46)	-0.36 (0.37)	-0.81 (2.92)
	-0.22 (0.43)			
Parenting Synchronicity	-0.16 (0.60)	0.03 (0.10)	-0.09 (0.08)	-0.38 (0.67)
	0.01 (0.09)			
Parent support for child learning	-0.58† (0.34)	0.10† (0.05)	0.00 (0.04)	0.40 (0.34)
Parent provision of learning materials	-0.09 (0.36)	0.01 (0.06)	0.06 (0.05)	0.34 (0.42)
	-0.05 (0.05)			

Source: Baby FACES Direct Child Assessment, Staff Interview, and Parent Interview and Classroom Observation.

Note: Analyses are weighted to be representative of children who were enrolled in Early Head Start in spring 2009 and continued to be enrolled at age 3. Estimates from mixed models in SAS. We used multiple imputation (N=10) to handle missing data. Covariates include cohort, child age, gender, race, home language, maternal demographic risk, poverty ratio, and child's age 2 scores.

[a] Quadratic term was dropped from the model if not significant, and results without the quadratic term are presented in the second row for each outcome measure.

[b] Two separate variables representing quality in the low and high range are included in the model. The coefficients are bolded if the associations between quality and child outcome are different for quality in the higher versus lower range ($p < .10$). Cut-point for quality measures in these analyses is 4 for Visitor Strategies.

†$p < .10$; *$p < .05$; **$p < .01$; ***$p < .001$.

Table A.3. Associations between classroom quality and child outcomes at age 2 and tests for threshold effects

	Test of associations		Test for threshold[b]	
	Linear	Quadratic[a]	Low quality	High quality
Emotional and Behavioral Support				
PLS-4	15.03 (13.90)	-1.12 (1.38)	0.05 (0.71)	4.69 (3.21)
	3.77 (1.38) **			
ASQ-3 Problem Solving	-4.12 (5.80)	0.41 (0.57)	-0.17 (0.30)	-0.13 (1.17)
	0.05 (0.57)			
BRS Emotional Regulation	-8.57 (7.18)	0.87 (0.71)	0.20 (0.35)	1.59 (1.46)
	0.24 (0.70)			
BRS Orientation/Engagement	-13.65 (5.61)*	1.37 (0.55)*	**0.38 (0.28)**	**2.32* (1.12)**
Teacher-reported BITSEA Competence	2.06 (2.88)	-0.15 (0.28)	-0.02 (0.14)	0.52 (0.58)
	0.49† (0.28)			
Teacher-reported BITSEA Problem	-8.36† (4.56)	0.81† (0.45)	-0.05 (0.22)	0.11 (0.91)
Engaged Support for Learning				
PLS-4	-1.32 (5.17)	0.52 (.70)	**0.70 (1.00)**	**7.06* (2.75)**
	2.44 (.91)**			
ASQ-3 Problem Solving	0.01 (2.27)	-0.01 (0.30)	0.23 (0.44)	0.49 (1.10)
	-0.04 (0.37)			
BRS Emotional Regulation	1.94 (2.57)	-0.24 (.35)	-0.08 (0.51)	0.08 (1.39)
	.22 (.45)			
BRS Orientation/Engagement	-1.55 (2.00)	0.25 (.27)	0.22 (0.41)	1.58 (1.07)
	.30 (.35)			
Teacher-reported BITSEA Competence	.95 (1.04)	-0.07 (.14)	**-0.10 (0.22)**	**0.87 (0.58)**
	.43 (.18)*			
Teacher-reported BITSEA Problem	-2.99† (1.66)	0.38† (.22)	-0.08 (0.33)	-0.16 (0.87)

Source: Baby FACES Direct Child Assessment, Staff Interview, and Parent Interview and Classroom Observation.

Note: Analyses are weighted to be representative of children who were enrolled in Early Head Start in spring 2009 and continued to be enrolled at age 3. Estimates from mixed models in SAS. We used multiple imputation (N=10) to handle missing data. Covariates include cohort, child age, gender, race, home language, maternal demographic risk, poverty ratio, and child's age 1 scores.

[a] Quadratic term was dropped from the model if not significant, and results without the quadratic term are presented in the second row for each outcome measure.

[b] Two separate variables representing quality in the low and high range are included in the model. The coefficients are bolded if the associations between quality and child outcome are different for quality in the higher versus lower range ($p < .10$). Cut-points for quality measures in these analyses are 5 for Emotional and Behavioral Support and 4 for Engaged Support for Learning.

†$p < .10$; *$p < .05$; **$p < .01$; ***$p < .001$.

Table A.4. Associations between classroom quality and child outcomes at age 3 and tests for threshold effects

	Test of associations		Test for threshold[b]	
	Linear	Quadratic[a]	Low quality	High quality
Emotional and Behavioral Support				
PLS-4	0.33 (7.73)	-0.08 (0.84)	-0.40 (0.72)	-2.11 (3.40)
	-0.35 (1.12)			
PPVT-4	0.80 (7.51)	-0.08 (0.80)	0.03 (0.61)	0.30 (2.70)
	0.04 (1.02)			
ASQ-3 Problem Solving	4.20 (4.08)	-0.39 (0.44)	0.18 (0.39)	1.11 (1.57)
	0.60 (0.60)			
BRS Emotional Regulation	-0.11 (3.39)	0.09 (0.35)	-0.09 (0.29)	1.00 (1.19)
	0.74 (0.47)			
BRS Orientation/Engagement	-2.85 (3.43)	0.35 (0.36)	-0.06 (0.30)	0.75 (1.26)
	0.39 (0.50)			
Teacher-reported BITSEA Competence	0.60 (1.87)	-0.05 (0.20)	0.23 (0.16)	0.92 (0.69)
	0.18 (0.28)			
Teacher-reported BITSEA Problem	2.76 (2.95)	-0.36 (0.31)	**0.03 (0.26)**	**-1.62 (1.12)**
	-0.63 (0.44)			
Engaged Support for Learning				
PLS-4	-1.57 (4.16)	0.26 (.57)	0.92 (1.10)	3.02 (2.54)
	0.32 (.82)			
PPVT-4	-4.67 (4.20)	0.68 (.57)	-0.80 (0.90)	0.86 (2.24)
	0.21 (.73)			
ASQ-3 Problem Solving	-1.68 (2.18)	0.33 (0.30)	**1.38** (0.53)**	**3.99** (1.36)**
	0.61 (0.43)			
BRS Emotional Regulation	0.10 (1.79)	0.15 (.25)	**0.13 (0.42)**	**2.69* (1.13)**
	1.19 (.34)***			
BRS Orientation/Engagement	-0.77 (1.83)	0.25 (.25)	**0.81† (0.43)**	**3.65** (1.17)**
	0.98 (.36)**			
Teacher-reported BITSEA Competence	-0.36 (1.01)	0.08 (.14)	-0.28 (0.23)	0.18 (0.62)
	0.23 (.19)			
Teacher-reported BITSEA Problem	0.01 (1.58)	-0.04 (.22)	0.17 (0.38)	-0.36 (0.99)
	-0.24 (.31)			

Source: Baby FACES Direct Child Assessment, Staff Interview, and Parent Interview and Classroom Observation.

Note: Analyses are weighted to be representative of children who were enrolled in Early Head Start in spring 2009 and continued to be enrolled at age 3. Estimates from mixed models in SAS. We used multiple imputation (N=10) to handle missing data. Covariates include cohort, child age, gender, race, home language, maternal demographic risk, poverty ratio, and child's age 2 scores.

[a] Quadratic term was dropped from the model if not significant, and results without the quadratic term are presented in the second row for each outcome measure.

[b] Two separate variables representing quality in the low and high range are included in the model. The coefficients are bolded if the associations between quality and child outcome are different for quality in the higher versus lower range ($p < .10$). Cut-points for quality measures in these analyses are 5 for Emotional and Behavioral Support and 4 for Engaged Support for Learning.

†$p < .10$; *$p < .05$; **$p < .01$; ***$p < .001$.

This page is left blank for double-sided printing.

This page is left blank for double-sided printing.

www.mathematica-mpr.com

Improving public well-being by conducting high quality, objective research and data collection

PRINCETON, NJ ▪ ANN ARBOR, MI ▪ CAMBRIDGE, MA ▪ CHICAGO, IL ▪ OAKLAND, CA ▪ WASHINGTON, DC

MATHEMATICA
Policy Research

Mathematica® is a registered trademark
of Mathematica Policy Research, Inc.