Tracking Quality in Head Start Classrooms:
FACES 2006 to FACES 2014
Technical Report

Tracking Quality in Head Start Classrooms: FACES 2006 to FACES 2014

Technical Report

OPRE Report 2016-95

November 2016

Submitted by:
Nikki Aikens
Charles Bush
Philip Gleason
Lizabeth Malone
Louisa Tarullo
Mathematica Policy Research

Project Director:
Lizabeth Malone, Mathematica Policy Research
1100 1st Street, NE, 12th Floor
Washington, DC 20002-4221

≡ OPRE   **MATHEMATICA**
Policy Research

This page has been left blank for double-sided copying.

CONTENTS

This page has been left blank for double-sided copying.

TABLES

I.   INTRODUCTION AND BACKGROUND

In any early care and education program, a central focus is ensuring effective classroom practice to promote learning. Research shows the quality and consistency of care are related to children's cognitive and social-emotional development (National Institute of Child Health and Development [NICHD] Early Child Care Research Network 1998, 2000; Shonkoff and Phillips 2000; Owen et al. 2008; Votruba-Drzal et al. 2004). Studies also show that better classroom quality is associated with a range of characteristics, including teachers' education and training; beliefs (for example, about developmentally appropriate practice); depressive symptoms; and job satisfaction (Aikens et al. 2010; Gerber et al. 2007; McLean and Connor 2015; Moiduddin et al. 2012; Resnick and Zill 2003). Group size, child/adult ratios, and child and family characteristics can also shape teachers' behaviors in the classroom (ACF 2003; Resnick and Zill 2003).

In this report, we highlight findings from cross-cohort analyses of data from the Head Start Family and Child Experiences Survey (FACES) 2006, 2009, and 2014. The analyses (1) provide a descriptive portrait of observed classroom quality and other relevant classroom, teacher,[1] and program characteristics at each time period; (2) determine the existence of trends or patterns in observed classroom quality and selected classroom, teacher, and program characteristics across cohorts; and (3) examine whether changes in such characteristics can explain the trends in observed classroom quality fully, partially, or not at all. We first describe the research questions the analyses are intended to address and then provide a brief overview of the FACES design across cohorts, including a description of any caveats related to the instrumentation and sampling across cohorts. Next, we describe our analytic approach and summarize findings from each analysis. We conclude this report with a discussion of the implications of the findings for future research. A companion policy brief (Aikens et al. 2016), also highlights a subset of the current findings.

A. Cross-cohort analysis research questions

Using data from FACES 2006, 2009, and 2014, the cross-cohort analyses address research questions[2] in the following areas:

Classroom characteristics: Observed classroom quality and instruction

1. Has the observed quality of Head Start classrooms changed from FACES 2006 to FACES 2014, or from FACES 2009 to FACES 2014?

Teacher characteristics: Teacher professional development, credentials, and background

2. Are there changes in mentoring in Head Start programs from FACES 2006 to 2014? Specifically, how many teachers report that they have a mentor? How frequently do they report meeting with mentors? Who provides mentoring?

---

[1] Throughout this report, we use the term teacher to refer to the lead teacher working in the classroom with children.

[2] As requested by the Administration for Children and Families (ACF), the analyses examine trends in observed classroom quality from FACES 2006 to 2014 and from FACES 2009 to 2014. We examine trends in all other characteristics from FACES 2006 to 2014, unless noted otherwise.

3. Has the number of trainings or workshops that Head Start teachers report attending changed from FACES 2009 to 2014?[3]

4. Have the curricular supports that teachers report, including curriculum and ongoing support for curriculum, changed from FACES 2006 to 2014?

5. Have the education levels and credentials of Head Start teachers changed from FACES 2006 to 2014?

6. Have the reported mental health status (depressive symptoms) and job satisfaction of Head Start teachers changed from FACES 2006 to 2014?

Program characteristics: Teacher staffing and turnover

7. Has the prevalence of teacher turnover in Head Start programs changed from FACES 2006 to 2014?

Mediators of observed classroom quality

8. Are changes in observed classroom quality across cohorts explained by selected teacher characteristics (prevalence of mentoring, who provides mentoring, and level of education)?

B. The FACES design across cohorts

The FACES sample provides information at the national level about Head Start programs, centers, and classrooms and the children and families they serve. Each cohort selected a sample of Head Start programs from the Head Start Program Information Report (PIR), with centers within each program and classrooms within each center selected for participation. Within each classroom, each study cohort randomly selected a sample of children. FACES 2006 and 2009 focused the sampling effort only on newly entering children and the classrooms and programs serving them; in FACES 2014, sampling represented all children—whether newly entering or returning—and their classrooms and programs. In Table I.1, we highlight the sample selection process and participating sample size for each FACES cohort in the current analyses.

---

[3] Information on training sessions and workshops is available only in FACES 2009 and 2014.

Table I.1. Sampling in the FACES 2006, 2009, and 2014 cohorts

Number expected	FACES 2006	FACES 2009	FACES 2014 With child-level data	Without child-level data
Programs	60	60	60	120
Centers	2 per program	2 per program	2 per program	2 per program
Classrooms[a]	3 per center	3 per center	2 per center	2 per center
Children	10 per classroom (newly entering)	10 per classroom (newly entering)	12 per classroom (newly entering and returning)	0
Actual sample size[b]				
Programs	60	60	60	116
Centers	134	129	119	228
Classrooms[a]	410	486	247	422
Children	3,315 (newly entering)	3,349 (newly entering)	2,462 (newly entering and returning)	0

[a]FACES 2006 and 2009 only selected classrooms serving at least one newly entering child.

[b]In some instances, the study selected more centers per program and more classrooms per center than expected, resulting in higher actual sample sizes.

In each cohort, Mathematica data collection teams assessed the children one-on-one at their Head Start centers and conducted observations of classroom quality. Children's parents completed interviews/surveys about family background and home experiences, and teachers were asked to complete a set of ratings for each sampled child in their classroom. Teachers also completed interviews/surveys about their classrooms and themselves; in addition, center and program directors completed interviews/surveys about their programs and themselves.[4]

In Table I.2, we highlight the mode, length, timing, and content of each instrument used in the current analyses. Information on teacher characteristics included in this report (including prevalence and frequency of mentoring and by whom, participation in training and workshops, and prevalence and nature of curriculum training and ongoing supports; education levels and credentials; depressive symptoms; and job satisfaction) draw from teacher interviews/surveys. We also use information from director interviews/surveys to describe the number of teachers employed and rates of teacher turnover, and parent interviews/surveys to describe the demographic characteristics of children served in classrooms. To measure the quality of Head Start classrooms, analyses use the Classroom Assessment Scoring System (CLASS; Pianta et al. 2008) and the Early Childhood Environment Rating Scale-Revised (ECERS-R; Harms et al. 1998), along with observed child-adult ratios and group sizes. The CLASS measures classroom quality in terms of both instructional and social-emotional aspects of the environment, across three domains of interaction: Instructional Support, Emotional Support, and Classroom Organization. The ECERS-R is a global rating of classroom quality based on structural features of the classroom. Two factor scores—Teaching and Interactions and Provisions for Learning— can be derived from both the full ECERS-R and the short form based on a subset of items (Clifford et al. 2005). Both the CLASS and the ECERS-R items are scored on a seven-point

---

[4] More details on the study methodology and response rates in each cohort appear in the User's Manual accompanying the public-use data files (West et al. 2010; Malone 2013; Kopack Klein et al. 2016).

scale, with higher scores reflecting better quality care. Table I.3 describes the factors and domains measured by the ECERS-R and the CLASS.

As noted, FACES 2014 differs from earlier cohorts in the available mode of the instruments and, in some instances, the timing of the data collection. For example, the FACES 2006 and 2009 cohorts conducted director interviews/surveys in fall, whereas FACES 2014 conducted these instruments in spring. Similarly, teacher interviews/surveys took place in both fall and spring in the FACES 2006 and 2009 cohorts[5] but only in spring in FACES 2014. In some instances, content also varies across cohorts. For example, only the Instructional Support domain of the CLASS is available in FACES 2006; however, all three domains of the CLASS (Instructional Support, Emotional Support, and Classroom Organization) are available in later cohorts. Similarly, whereas FACES 2006 used the full ECERS-R, FACES 2009 and 2014 used only the short form.[6,7] Each of these differences should be kept in mind when interpreting findings from the data.

---

[5] The FACES 2006 and 2009 cohorts followed newly entering children for two years, collecting interview data from Head Start parents and teachers and conducting classroom observations in a second spring (spring 2008 and 2011, respectively). We focus on only the first-year data in the current reporting and analyses.

[6] ECERS-R factor scores in FACES 2006 are derived from the full ECERS-R administration. Comparable scores in FACES 2009 and 2014 are based on only the short-form administration.

[7] In FACES 2009 and 2014, the ECERS-R short form and full CLASS observations were conducted at the same time by a single observer. Earlier analyses suggest that the ECERS-R short form is more reliable than the standard ECERS-R. A single observer may be trained to use the ECERS-R short form and CLASS and will obtain more and higher quality data than he or she would with the standard ECERS-R (Meagher and Verbitsky-Savitz 2011).

Table I.2. Mode, length, timing, and content of key instruments in the FACES 2006, 2009, and 2014 cohorts

	FACES 2006	FACES 2009	FACES 2014
Director interviews/surveys[a]			
Mode	Telephone and in-person	Telephone and in-person	Web and paper
Length	60 minutes	30 minutes	25-30 minutes
Timing	Fall 2006	Fall 2009	Spring 2015
Content included in analyses	Number of teachers employed and rates of teacher turnover	Number of teachers employed and rates of teacher turnover	Number of teachers employed and rates of teacher turnover
Teacher interviews/surveys[a, b]			
Mode	In-person	In-person	Web and paper
Length	60 minutes	30 minutes	30 minutes
Timing	Fall 2006, spring 2007	Fall 2009, spring 2010	Spring 2015
Content included in analyses	Teacher educational background, professional experience, credentials, job satisfaction, and depressive symptoms; curriculum training and ongoing supports; whether have a mentor and experiences with that mentor	Teacher educational background, professional experience, credentials, job satisfaction, and depressive symptoms; curriculum training and ongoing supports; whether have a mentor and experiences with that mentor; and involvement in training during program year	Teacher educational background, professional experience, credentials, job satisfaction, and depressive symptoms; curriculum training and ongoing supports; whether have a mentor and experiences with that mentor; and involvement in training during program year
Classroom observations[b]			
Mode	Paper	Paper	Paper
Length	Four hours	Four hours	Four hours
Timing	Spring 2007	Spring 2010	Spring 2015
Content included in analyses	Standard Early Childhood Environment Rating Scale-Revised (ECERS-R), Instructional Support domain of the Classroom Assessment Scoring System for prekindergarten (CLASS-PreK), child-adult ratios and group sizes	Early Childhood Environment Rating Scale-Revised short form (ECERS-R), the full Classroom Assessment Scoring System for prekindergarten (CLASS-PreK), child-adult ratios and group sizes	Early Childhood Environment Rating Scale-Revised short form (ECERS-R), the full Classroom Assessment Scoring System for prekindergarten (CLASS-PreK), child-adult ratios and group sizes
Parent interviews/surveys[a, b]			
Mode	Telephone and in-person	Telephone and in-person	Web and telephone
Length	60 minutes	50 minutes	20-25 minutes
Timing	Fall 2006, spring 2007	Fall 2009, spring 2010	Fall 2014, spring 2015
Content included in analyses	Child and family characteristics, including child age, dual language learner status, and household poverty status	Child and family characteristics, including child age, dual language learner status, and household poverty status	Child and family characteristics, including child age, dual language learner status, and household poverty status

Table I.2 (*continued*)

[a]In FACES 2006 and 2009, these instruments were referred to as interviews given that they were administered by another person. FACES 2014 refers to these instruments as surveys given the option of self-administration.

[b]The FACES 2006 and 2009 cohorts followed newly entering children for two years, collecting interview data from Head Start parents and teachers and conducting classroom observations in a second spring (spring 2008 and 2011, respectively). We focus on only the first spring data in the current reporting and analyses.

Table I.3. Description of ECERS-R factors and CLASS domains

Factors or domains	Characteristics assessed by factors or domains
ECERS-R Teaching and Interactions	Quality of teacher-child interactions
ECERS-R Provisions for Learning	Materials available in the classroom and the arrangement of classroom space
CLASS Instructional Support	Quality of instructional practices used in the classroom
CLASS Emotional Support	Social and emotional functioning in the classroom
CLASS Classroom Organization	Teacher's ability to organize the classroom to make efficient use of class time

C. Analytic approach

The cross-cohort analytic activities included producing descriptive estimates for each cohort of interest (FACES 2006, 2009, 2014) and conducting linear regressions to examine trends across cohorts (2006 to 2014 and 2009 to 2014, where applicable) and to determine if selected characteristics explain trends in classroom quality across cohorts. We provide descriptive estimates for all characteristics included in the analyses (Appendix Tables B.1 through B.7). Given the interest in child demographic characteristics as control variables,[8] we used data only from the 60 programs with child-level data in FACES 2014 in all analyses (descriptive and regression). Across time points, we aggregated the child-level characteristics to the classroom level. We drew demographic characteristics from the parent interviews/surveys given the availability of such information. As a result, the characteristics reflect the characteristics of sampled children rather than those of all children in classrooms. We also duplicated center-level reports of staffing and teacher turnover at the classroom level. We used sampling weights and the appropriate strata and primary sampling unit (PSU) variables in all analyses to account for the multistage sample design (Klein et al. 2016).[9]

---

[8] As noted in Appendix Table A.1, analyses included the following child demographic characteristics: average child age, percentage of dual language learners (DLLs), percentage of children with incomes below poverty. We define DLLs as those in households where a language other than English is primarily spoken to the child. This may differ from definitions used in other studies.

[9] As noted, there are differences in the population represented over time by the selected cohorts (that is, FACES 2006 and 2009 classrooms represent all Head Start classrooms serving newly entering children, whereas FACES 2014 classrooms represent all Head Start classrooms). Given these differences, we examined whether the descriptive statistics and trend analyses differed when including all FACES 2014 classrooms in the 60 programs (n=247) versus those classrooms in the 60 programs serving newly entering children (whether alone or with returning children; n=225). The latter analyses excluded 22 classrooms in which there were no newly entering children in the study's child sample. Findings were similar between the two versions of the analyses; therefore all reported findings include all classrooms in the 60 programs.

The analyses proceeded in stages. We first (1) estimated descriptive statistics (that is, means, standard deviations, response ranges, and percentages) before (2) examining trends in those characteristics across cohorts. We then examined the characteristics that might explain trends in the selected outcome of interest—observed classroom quality—including both (3) control variables (those unlikely to be influenced by policies implemented by Head Start in an effort to improve quality) and (4) mediating variables (likely to be influenced by Head Start program policies). We selected the control variables a priori and the mediators based on their performance in stage (2) (selecting variables that changed significantly across cohorts or changed in ways likely to explain change in observed quality) and for theoretical reasons (selecting variables that we hypothesized would explain improvements in classroom quality or would be influenced by Head Start policies). We did not include variables that did not meet these criteria as mediators. In these analyses, perhaps the best estimates of how Head Start quality changed across cohorts come from stage (3);[10] in fact, by comparing stages (3) and (4), we can develop some evidence about what policies or changes might contribute to any improvement or decline in quality. As an example, suppose the results in (3) suggest that ECERS-R factor scores increase from FACES 2006 to 2014, and when teacher characteristics are added in the model (4), there is no longer a significant trend in ECERS-R scores. This would suggest that the selected teacher characteristics mediate the trend in quality—that the change in the selected teacher characteristics from FACES 2006 to 2014 explain the increase in ECERS-R factor scores during this period. In addition, the individual regression coefficients for each teacher characteristic included in stage (4) provide information on the extent to which each helps explain the trend.

Because we were interested in explaining trends in quality, we focused on the quality outcomes that changed across cohorts and/or with adequate variability.[11] As noted above, to explain those trends, we focused on characteristics that also changed over the same period, changed in ways likely to explain trends in observed quality, we hypothesized would explain improvements in classroom quality, and were endogenous or likely to be influenced by Head Start policies. We excluded any variables that did not meet these criteria from the mediation analyses. Table B.8 lists the variables included in the trend analyses, findings from the trend analyses for each, and whether they were included as control or mediators in the mediation analyses.

For the mediation analyses, classroom quality outcomes included continuous ECERS-R factor and CLASS domain scores (means) and the distribution across publisher categories on the two ECERS-R factors and on CLASS Instructional Support and Emotional Support.[12] The analyses included control variables focused on child demographic characteristics (average child

---

[10] The analysis provides adjusted means, controlling for the demographic characteristics of interest among children in each classroom.

[11] We focused on quality outcomes that were relatively well distributed across the publisher categories rather than those having most or almost all classrooms falling into any of the categories. For example, across cohorts, no more than 2 percent of classrooms had CLASS Classroom Organization quality scores in the low or high range. As a result, we did not conduct mediation analyses with these quality outcome scores.

[12] We selected the latter outcome variables based on the descriptive and trend analyses, focusing primarily on those variables that showed change over time and/or with adequate variability.

age, percentage of DLLs, and percentage of children with family income below poverty)[13] and mediating variables focused on teacher characteristics (teacher has at least a BA and whether the teacher has a mentor and who serves as that mentor).[14]

Both the trend and mediation analyses involved regression analyses: linear regressions for continuous variables and simple logistic regressions for dichotomous variables. Simple logistic regressions examine change in a single category of a categorical or dichotomous variables (for example, percentage of classrooms with "low" Instructional Support scores over time or percentage of classrooms in which the teacher has a mentor over time). One can also conduct ordered logistic regressions if an examination of changes over time in the overall distribution across the categories of a categorical variable is of interest. Findings from such analyses can sometimes be challenging to interpret if patterns across categories differ; therefore, we focus in this report on a series of logistic regressions conducted for the different categories of each categorical variable. The models included a single variable of interest as the dependent variable (for example, CLASS Instructional Support mean scores or "low" Instructional Support scores [versus all others]) and dummy codes for cohort (for example, 2009, 2014) as the predictor variables.

Ultimately the analyses examined whether the control variables and mediators explain the trends in observed classroom quality fully, partially, or not at all. That is, we examined whether the increases in observed quality remained even after accounting for changes in the characteristics of Head Start children and whether the improvements in quality could be explained by changes in the types of staff providing mentoring, and the educational levels of teachers. To do so, the analyses focused on changes (that is, decreases, increases) in the regression coefficients for time/cohort with the addition of control and mediators to the models. If the time/cohort regression coefficient moved toward zero with the addition of the variables, we would say that the variables help explain the trends. If the coefficient did not change, we would say that the variables did not explain the trend. A mediating variable has the potential to explain a trend if (1) the mediator itself changes across cohorts and (2) the mediator is directly related to the outcome of interest. We entered the variables in blocks, first including only the set of controls and then including both the set of control variables and mediators.

Table I.4 highlights the continuous dependent variables (classroom quality mean scores) and dichotomous dependent variables (classroom quality publisher categories) examined in the trend and mediation analyses. For each of the analyses focused on categorical dependent variables, we

---

[13] We aggregated child demographic characteristics up to the classroom level from parent interview/survey data. As a result, the characteristics reflect the characteristics of sampled children rather than those of all children in classrooms. The child sample included 10 to 12 children per classroom.

[14] We excluded most teacher characteristics as mediators from the analyses because they did not change over time. A handful of other characteristics changed over time but were not included as mediators: ongoing supports for curriculum use, the staff providing teachers with ongoing supports for curriculum use, and mean level of teacher depressive symptoms. With regard to supports related to curriculum use, we found that there was a decrease in some types of support over time, and we would not expect that such a decrease would be associated with an increase in quality. On teacher depressive symptoms, while we saw a small significant decline in level of depressive symptoms over time, that shift was not accompanied by a change in the percentage of teachers who met criteria for depression. For all of these reasons, we excluded these variables from the mediation models.

recoded the variable to be dichotomous (equal to 1 or 0) and conducted analyses with each dichotomous variable. For the observed classroom quality variables, this included constructing and separately examining indicators based on publisher-reported categories. For example, analyses examining publisher-reported categories on CLASS Instructional Support included those with scores in the "low" range (versus all others), in the "mid" range (versus all others), and in the "high" range (versus all others) each serving as the dependent variable in separate analyses.

Table I.4. Continuous and dichotomous classroom quality dependent variables used for trend and mediation analyses[a]

	FACES 2006-2014	FACES 2009-2014
ECERS-R short form factors		
Teaching and Interactions	Mean scores Inadequate (versus all others) Minimal (versus all others) Good/excellent (versus all others)	Mean scores Inadequate (versus all others) Minimal (versus all others) Good/excellent (versus all others)
Provisions for Learning	Mean scores Inadequate (versus all others) Minimal (versus all others) Good/excellent (versus all others)	Mean scores Inadequate (versus all others) Minimal (versus all others) Good/excellent (versus all others)
CLASS domains		
Instructional Support	Mean scores Low (versus all others) Mid (versus all others) High (versus all others)	Mean scores Low (versus all others) Mid (versus all others) High (versus all others)
Emotional Support	NA	Mean scores Low (versus all others) Mid (versus all others) High (versus all others)
Classroom Organization	NA	Mean scores

NA=not applicable. Data not available in cohort and analysis not conducted between cohorts.

[a]Across classroom quality outcomes, the trend and mediation analyses focused on the same dependent variables. The one exception is the CLASS Classroom Organization domain scores. The trend analyses examined low, mid, and high categories on CLASS Classroom Organization from FACES 2009 to 2014, whereas the mediation analyses did not. We omitted CLASS Classroom Organization categories from the mediation analyses because the scores on this domain were unevenly distributed.

In Appendix Tables A.1 through A.3, we list the variables in the analyses, indicating whether we explored each descriptively only or as an outcome, control, or mediating variable.

This page has been left blank for double-sided copying.

II. KEY FINDINGS FROM THE CROSS-COHORT ANALYSES

Next, we discuss the key findings from the analyses. Appendix Tables B.1 through B.7 provide findings from the analyses, highlighting descriptive estimates across cohorts and any trends in those estimates.[15] We also highlight whether other characteristics in the model mediate trends in observed classroom quality. The findings that follow do not indicate causation. In other words, we cannot attribute trends in classroom quality to specific policy and practice initiatives undertaken by OHS. Similarly, we are not able to conclude that selected teacher characteristics cause trends in observed classroom quality. Rather, the analyses identify whether selected teacher characteristics help to explain trends in observed quality by examining whether trends occurring across the same periods of time in observed quality and teacher characteristics are correlated with one another. Findings may help to inform future work on factors that may contribute to improvements in classroom quality.

A. Trends across cohorts

The trend analyses examined whether classroom quality and instruction, teacher professional development, and teacher background changed across cohorts. We describe findings from analyses in each of these areas. We identify statistically significant change across cohorts at the p<.05 level.

Classroom quality and instruction. The trend analyses first examined whether changes occurred in ECERS-R factor scores and CLASS domain scores—including average scores and the percentage meeting publisher-developed categories of quality (for example, "low," "mid," or "high" scores)[16]—from FACES 2006 to 2014 and from FACES 2009 to 2014. The analyses also examined trends in group sizes and child/adult ratios during the same periods. Key findings include the following:

- Average ECERS-R Teaching and Interactions factor scores improve across cohorts, including from FACES 2006 to FACES 2014 (4.0 to 5.1) and from FACES 2009 to FACES 2014 (4.7 to 5.1), moving from the "minimal" range to the "good" range on the 7-point scale, on average (Table B.1).

  - Based on publisher-developed cut-points, there are improvements in ECERS-R Teaching and Interactions factor scores from FACES 2006 to 2014, with fewer classrooms scoring

---

[15] As a reminder, we examine trends in observed classroom quality from FACES 2006 to 2014 and from FACES 2009 to 2014. We examine trends in all other characteristics from FACES 2006 to 2014, unless noted otherwise.

[16] Scores on the ECERS-R factors range from 1 to 7 with publisher-developed categories including inadequate (scores equal to 1 or 2), minimal (scores equal to 3 or 4), good (scores equal to 5 or 6), and excellent (scores equal to 7) quality. Scores for the CLASS domains range from 1 to 7 with publisher-developed categories including low (1 or 2), mid (3, 4, or 5), and high (6 or 7) quality. For the purpose of categorizing classrooms, the scores on the ECERS-R factors and CLASS domains were not rounded. For example, a classroom with a score of 5.9 on the CLASS Emotional Support domain would be categorized as falling in the mid range, rather than the high range; only scores of 6.0 or above would be included in the high range.

in the inadequate (12 to 1 percent) and minimal ranges (75 to 45 percent; Table B.1a) and more scoring in the good/excellent range[17] (13 to 54 percent) over time.

- From 2009 to 2014 on the ECERS-R Teaching and Interactions factor, fewer classrooms score in the inadequate range (5 to 1 percent; Table B.1a). There are no statistically significant changes in the other two categories of quality, however.

• Average ECERS-R Provisions for Learning factor scores improve across cohorts, including from FACES 2006 to FACES 2014 (3.6 to 4.4) and from FACES 2009 to FACES 2014 (4.0 to 4.4; Table B.1). However, on average, scores remain in the "minimal" range on the 7-point scale across cohorts.

- Based on publisher-developed cut-points, there are improvements in ECERS-R Provisions for Learning factor scores from FACES 2006 to 2014, with fewer classrooms scoring in the inadequate (19 to 9 percent) and minimal ranges (75 to 59 percent) and more scoring in the good/excellent range (6 to 32 percent; Table B.1a).

- Also from 2009 to 2014 on the Provisions for Learning factor, fewer classrooms score in the minimal range (79 to 59 percent), and more classrooms score in the good/excellent range (13 to 32 percent; Table B.1a).

• Average Instructional Support scores on the CLASS improve from FACES 2006 to 2014 (1.9 to 2.4), but do not change from FACES 2009 to 2014 (Table B.1). Classrooms score on the low end of the 7-point scale across cohorts, on average.

- Based on publisher-developed cut-points, there are improvements in CLASS Instructional Support scores from FACES 2006 to 2014, with fewer classrooms scoring in the low range (96 to 76 percent) and more scoring in the mid (4 to 24 percent) or high (0 to 1 percent) ranges across cohorts (Table B.1a).

- From 2009 to 2014 on the Instructional Support domain, fewer classrooms score in the low range (85 percent to 76 percent), and more classrooms in the high range (0 percent to 1 percent; Table B.1a).

• There are no statistically significant changes in CLASS Emotional Support and Classroom Organization scores from FACES 2009 to FACES 2014 (Table B.1). On both domains, classrooms score in the mid-range of the 7-point scale across cohorts, on average.

- On the Emotional Support domain, from 2009 to 2014, fewer classrooms score in the mid range, and more classrooms scored in the high range (Table B.1a). On the Classroom Organization domain, fewer classrooms score in the mid -range (99 to 96 percent) between 2009 to 2014 but no pattern of improvement was observed in the other categories of quality.

• Based on classroom observations, the child/adult ratio in classrooms decreases (6.3 to 5.9) from FACES 2006 to 2014 (Table B.1). Group size also decreases (14.6 to 13.5) from

---

[17] Given that no or fewer than 1 percent of classrooms scored in the "excellent" range, we collapsed the "good" and "excellent" response options for the analysis.

FACES 2006 to 2014. Group size and ratios do not change significantly in the period from FACES 2009 to 2014.

Teacher professional development. Next, we examined whether mentoring and training changed from FACES 2006 to 2014.[18] Key findings include the following:

- We found that the prevalence of mentoring remains stable from FACES 2006 to 2014, with about three-quarters of classrooms having teachers with a mentor, across cohorts. (Table B.2). Among those classrooms in which teachers have mentors, there are no statistically significant changes in mentoring frequency (that is, frequency of mentoring at least every two weeks or mentoring once a month or less)[19] during the same period.

- Among those classrooms in which teachers have mentors, there is an increase in teacher reports of mentoring by an educational coordinator or specialist (56 to 73 percent) from FACES 2006 to 2014 (Table B.2). Meanwhile, the percentage of classrooms in which teachers are mentored by a center or program director decreases (25 to 10 percent).

- From FACES 2009 to 2014, teachers report attending fewer training sessions or workshops that last for less than one day (5.0 to 3.2; Table B.2). The number of training sessions or workshops attended by classroom teachers that lasted for a day or more than one day remains stable during this period.

- From FACES 2006 to 2014, there is a decrease in the percentage of classrooms that have teachers who receive help in implementing the curriculum (59 to 45 percent) and who receive feedback on implementing the curriculum (61 to 44 percent; Table B.3).

- From FACES 2006 to 2014, there is an increase in teacher report of receiving support from a mentor or master teacher (15 to 32 percent) or other Head Start teachers in the program (24 to 37 percent; Table B.3).

Teacher credentials and background. We examined whether teacher education, experience, job satisfaction, or mental health changed during the period from FACES 2006 to 2014. Key findings include the following:

- The percentage of classrooms that have a teacher with a bachelor's degree or higher increases from FACES 2006 to FACES 2014 (40 to 70 percent; Table B.4). Meanwhile, fewer classrooms have a teacher with an associate's degree (AA) or less (60 to 31 percent).

- There is no statistically significant change in teacher experience or job satisfaction from FACES 2006 to 2014 (Table B.5).

- The average level of depressive symptoms (mean score) reported by teachers decreases (4.7 to 3.6) from FACES 2006 to 2014, but the percentage of teachers classified as not depressed remains stable during this period (Table B.5).

---

[18] Given the timing of available data, the trend analyses that focused on training sessions and workshops examined change from FACES 2009 to 2014.

[19] For easier interpretation of findings, we collapsed the response options for the analysis (from once a week or more, once every two weeks, once a month, and less than once a month).

Teacher staffing and turnover. We examined whether teacher turnover changed during the period from FACES 2006 to 2014. Key findings include:

- According to center director reports of staffing and teacher turnover, the number of employed teachers and prevalence of teacher turnover in Head Start classrooms do not change significantly across cohorts (Table B.6).

Child demographic characteristics. Finally, we examined whether average child age, percentage of DLLs served, and percentage of children with family income below poverty served changed during the period from FACES 2006 to 2014. Key findings include the following:

- The average age in months of children served in Head Start classrooms increases from FACES 2006 to 2014 (45.5 to 47.9; Table B.7).[20]

- The percentage of children with family income below poverty served in classrooms increases from FACES 2006 to 2014 (58 to 68 percent; Table B.7). Meanwhile, the percentage of DLLs served in classrooms remains stable during the same period.

B. Explaining trends across cohorts

The final set of analyses examined whether selected child and teacher characteristics help explain trends in classroom quality outcomes. As a reminder, based on the trend analyses, we found evidence of increases in average classroom quality from FACES 2006 to 2014—on both ECERS-R factors and on CLASS Instructional Support. There was less evidence of change in observed classroom quality during the period from FACES 2009 to 2014, with only average ECERS-R factor scores changing significantly during the period.

Findings from the mediation analyses appear in Appendix Tables B.9 and B.9a.[21] The tables include columns that report the coefficient for (1) the trend in observed quality, (2) the trend when the control variables are included in the model, and (3) the trend when we include the control variables and mediators in the model. The coefficient in the main trend column (1) reflects the change in the quality outcome from FACES 2006 to 2014 (or from FACES 2009 to 2014). The coefficient in (2) represents the change in the quality outcome, holding constant the demographic characteristics of the children, and (3) represents the change in quality, holding constant both the child and teacher characteristics. The difference between coefficients (1) and (2) and between (1) and (3) provides an estimate of how much of the trend is explained by the controls alone and the controls and mediators together, respectively. By dividing these differences by (1), we derive an estimate of the proportion of the trend that is explained by the controls alone and the controls and mediators together. We highlight only those instances in which the controls/mediators explain a substantial proportion of the trend—in these instances, we use a threshold of 10 percent (that is, the time/cohort regression coefficient changes by more than 10 percent). A change in the regression coefficient of less than 10 percent would suggest

---

[20] Age in months is based on age as of September 1.

[21] Tables B.10, B.10a, B.10b, B.11, B.11a, and B.11b include the regression coefficients for each of the variables in the final mediation models.

that most of the trend in quality (90 percent) is driven by factors other than the controls and mediators included in the models.

Child demographic characteristics (controls). First, we examined whether changes across cohorts in the characteristics of children served by Head Start changed trends in average ECERS-R factor scores from FACES 2006 to 2014 and from FACES 2009 to 2014 and in CLASS domain scores from FACES 2006 to 2014. We also examined whether they explain changes in publisher-reported categories from FACES 2006 to 2014 and from FACES 2009 to 2014.[22] Key findings include the following:

- Accounting for child characteristics does not change trends in average ECERS-R Teaching and Interactions or Provisions for Learning scores or publisher-reported categories on these scores from FACES 2006 to 2014 or from FACES 2009 to 2014.

- Accounting for child characteristics also does not change trends in average CLASS domain scores or in publisher-reported categories on Instructional Support from FACES 2006 to 2014 or from FACES 2009 to 2014 or Emotional Support from FACES 2009 to 2014.

Teacher characteristics (mediators). Finally, we examined whether the set of mediators—teacher education and mentoring—along with the control variables (child demographic characteristics including child age, percentage of DLLs served, and percentage of children with family income below poverty served) help to explain changes in average ECERS-R factor scores from FACES 2006 to 2014 and from FACES 2009 to 2014 and in CLASS domain scores from FACES 2006 to 2014. We also examined whether the mediators explain changes in publisher-reported categories from FACES 2006 to 2014 and from FACES 2009 to 2014.[23] Key findings include the following:

- Teacher characteristics included in the analyses do not explain changes in average ECERS-R Teaching and Interactions or Provisions for Learning scores or the distribution across publisher-reported categories on these scores from FACES 2006 to 2014 or from FACES 2009 to 2014.

- The selected teacher characteristics partially explain changes in average CLASS Instructional Support scores from FACES 2006 to 2014. In fact, they explain 15 percent of the change in these scores during this period. Specifically, whether the teacher has at least a bachelor's degree explains approximately 12 percent of the total classroom quality trend (that is, the increase in CLASS Instructional Support scores), while other characteristics in the model contribute more modestly. In short, the increase in teachers with at least a bachelor's degree accounts for a small portion of the improvement in CLASS Instructional Support from FACES 2006 to 2014.

- The selected teacher characteristics do not explain changes in average Emotional Support or Classroom Organization scores. In fact, trends in the distribution across publisher-reported

---

[22] The latter analyses that focused only on publisher-reported categories examined the two ECERS-R factor scores and CLASS Instructional Support and Emotional Support scores.

[23] The latter analyses that focused only on publisher-reported categories examined the two ECERS-R factor scores and CLASS Instructional Support and Emotional Support scores.

categories on Emotional Support are stronger with the inclusion of these variables in the models. In other words, the trend would have been stronger (a larger increase in quality) if the selected teacher's characteristics had not changed during this period.

III. SUMMARY AND IMPLICATIONS OF FINDINGS

The goal of this report was to examine trends in observed classroom quality and selected classroom, teacher, and program characteristics across FACES cohorts. We also examined whether changes in classroom, teacher, and program characteristics are related to changes in observed classroom quality.

With respect to trends, some aspects of classroom quality change across cohorts, most commonly from FACES 2006 to 2014. For example, we see increases in average classroom quality from FACES 2006 to 2014 on both ECERS-R factors and on CLASS Instructional Support. From FACES 2009 to 2014, there are increases on average ECERS-R factor scores, but not on CLASS domain scores. In examining publisher categories, we found more classrooms score in the good or excellent range on both ECERS-R factors from 2006 to 2014, but from 2009 to 2014 only on ECERS-R Provisions for Learning scores. Fewer classrooms score in the low range on CLASS Instructional Support from 2006 to 2014. From 2009 to 2014, fewer score in the low range on Instructional Support, and more score in the high range on Emotional Support. Observed group sizes and child/adult ratios decrease from FACES 2006 to 2014.

As for teacher professional development, some aspects of mentoring and training change across cohorts. Although the prevalence and frequency of mentoring do not change from FACES 2006 to 2014, the staff who provide mentoring does change (with education coordinators/specialists more common across cohorts). We also observe changes in ongoing support for curriculum implementation (with teachers receiving less help and feedback in implementing curriculum) as well as changes in who provides those supports during the same period (with teachers receiving more support from master/mentor teachers or other teachers in the program). We also see changes in the number of brief trainings in which teachers participate, with teachers attending fewer trainings that last for less than one day, but no increase in attendance of trainings lasting longer than a day. Finally, from FACES 2006 to 2014, Head Start teachers are increasingly likely to have at least a bachelor's degree; meanwhile, teacher experience, job satisfaction, or turnover remain stable during the same period.

Accounting for child characteristics does not change the trends in the two ECERS-R factor scores or CLASS domain scores. The teacher characteristics that we examined also do not explain trends in the two ECERS-R factor average scores or the CLASS Emotional Support and Classroom Organization average scores. They also do not explain the trends in publisher categories on ECERS-R factors or CLASS domains. However, these characteristics partially explain trends in only CLASS Instructional Support scores (accounting for 15 percent of the change in scores). Of the variables we examined, whether the teacher has at least a bachelor's degree explains about 12 percent of the trend, and other characteristics in the model contribute more modestly to the trend. In short, the increase in teachers with at least a BA most helps to explain the improvement in CLASS Instructional Support from FACES 2006 to 2014.[24] The change in mentors from directors to education coordinators does not appear to account for

---

[24] Note, however, that having teachers with at least a BA is not significantly associated with CLASS Instructional Support scores. That is to say, the percentage of teachers with at least a BA is not significantly related to the CLASS Instruction Support score in a given year, based on data from all three years. However, the patterns between teacher education and CLASS scores are in expected directions.

changes in observed quality. Other factors might help to explain changes in observed quality, including those that are not measured in FACES. For example, although FACES captures information on the prevalence of mentoring and the staff who provide mentoring to teachers, the quality, intensity, and intentionality of that mentoring may be more critical for explaining improvements in classroom quality over time.

It is possible that, within the group of selected mediators, some mediators make the trends less strong while others truly mediate the trend, and perhaps cancelling out each other. Examining the mediators individually (that is, adding them individually to the models) would help determine if they cancel each other out. In addition, it is possible that a different set of characteristics may better explain the trends. As noted, a mediating variable has the potential to explain a trend if (1) the mediator itself changes across cohorts and (2) the mediator is directly related to the outcome of interest. Although some of the selected variables were associated with the quality scores in some of the models, they did not all change significantly across cohorts. Finally, it is possible that a selected mediator changes across cohorts but it is not related to the outcome (for example, fewer teachers are mentored by directors but that mentoring is not related to observed quality). In such an instance, the mediator cannot explain the trend in quality. Future work might examine additional factors that could be associated with observed classroom quality, such as program management quality or the use of data for quality improvement. Identifying processes that can explain the recent increases in Head Start classroom quality may help target resources for quality improvement efforts in the future. Additionally, further research would be needed to establish causal relationships between quality improvement efforts, changes in teacher characteristics, and changes in observed classroom quality.

REFERENCES

Administration for Children and Families. "Head Start FACES 2000: A Whole-Child Perspective on Program Performance, Fourth Progress Report." Washington, DC: U.S. Department of Health and Human Services, 2003.

Aikens, Nikki, Charles Bush, Philip Gleason, Lizabeth Malone, and Louisa Tarullo. (2016). *Tracking Quality in Head Start Classrooms: FACES 2006 to FACES 2014.* OPRE Brief 2016-82. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

Aikens, Nikki, Louisa Tarullo, Lara Hulsey, Christine Ross, Jerry West, and Yange 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.

Clifford, Richard, Oscar Barbarin, Florence Chang, Diane M. Early, Donna Bryant, Carollee Howes, Margaret Burchinal, and Robert Pianta. "What Is Pre-Kindergarten? Characteristics of Public Pre-Kindergarten Programs." *Applied Developmental Science*, vol. 9, no. 3, 2005, pp. 126–143.

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.

Harms, Thelma, Richard Clifford, and Debby Cryer. "Early Childhood Environment Rating Scale–Revised (ECERS-R)." New York, NY: Teachers College Press, 1998.

Kopack Klein, Ashley, Barbara Lepidus Carlson, Nikki Aikens, Anne Bloomenthal, Jerry West, Lizabeth Malone, Emily Moiduddin, Melissa Hepburn, Sara Skidmore, Sara Bernstein, Annalee Kelly, Felicia Hurwitz, and Grace Lim. "Head Start Family and Child Experiences Survey (FACES 2014) User's Manual." Draft 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 2016.

Malone, Lizabeth, Barbara Lepidus Carlson, Nikki Aikens, Emily Moiduddin, Ashley Kopack Klein, Jerry West, Annalee Kelly, Cassandra Meagher, Anne Bloomenthal, Lara Hulsey, and Kristina Rall. "Head Start Family and Child Experiences Survey: 2009 User's Manual." 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, April 2013.

McLean, Leigh, and Carol McDonald Connor. "Depressive Symptoms in Third-Grade Teachers: Relations to Classroom Quality and Student Achievement." *Child Development,* vol. 86, no. 3, May/June 2015, pp. 945–954.

Meagher, Cassandra, and Natalya Verbitsky-Savitz. "Classroom Quality in Numbers: Using Multiple Measures of Classroom Quality and the Effect on Reliability and Validity." Presentation to the Society for Research in Child Development Biennial Meeting, Montreal, Canada, April 2, 2011.

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.

NICHD Early Child Care Research Network. "Early Child Care and Self-Control, Compliance, and Problem Behavior at 24 and 36 Months." *Child Development,* vol. 69, no. 4, 1998, pp. 1145–1170.

NICHD Early Child Care Research Network. "The Relation of Child Care to Cognitive and Language Development." *Child Development,* vol. 71, no. 4, 2000, pp. 960–980.

Owen, M.T., J.F. Klauski, A. Mata-Otero, and M.O. Caughy. "Relationship-Focused Child Care Practices: Quality of Care and Child Outcomes for Children in Poverty." *Early Education and Development,* vol. 19, no. 2, 2008, pp. 302–329.

Pianta, R., K. LaParo, and B. Hamre. "The Classroom Assessment Scoring System Pre-K Manual." Charlottesville, VA: University of Virginia, 2008.

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.

Shonkoff, J.P., and D.A. Phillips. *From Neurons to Neighborhoods: The Science of Early Childhood Development*. Washington, DC: National Academies Press, 2000.

Votruba-Drzal, E., R.L. Coley, and P.L. Chase-Lansdale. "Child Care and Low-Income Children's Development: Direct and Moderated Effects." *Child Development,* vol. 71, no. 1, 2004, pp. 296–312.

West, Jerry, Nikki Aikens, Barbara Carlson, Cassandra Meagher, Lizabeth Malone, Anne Bloomenthal, Annalee Kelly, Kristina Rall, and Rita Zota. "Head Start Family and Child Experiences Survey: 2006 User's Manual." Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research and Evaluation, August 2010.

APPENDIX A

CROSS-COHORT ANALYSIS VARIABLES

This page has been left blank for double-sided copying.

Table A.1. Classroom environment constructs: FACES 2006, FACES 2009, and FACES 2014

Construct	Instrument	Used in FACES 2014?	Used in FACES 2009?	Used in FACES 2006?	Data reported	Analytic use of data[a]
Group size						
Class size and child/adult ratios	Classroom observation	Yes	Yes	Yes	Spring	Descriptive only
Child characteristics						
Classroom demographic characteristics: Child age distribution, proportion of DLLs, poverty status[b]	Parent survey	Yes	Yes	Yes	Spring	Control
Observed classroom quality						
Early Childhood Environment Rating Scale (ECERS-R; abbreviated form; 21 items)	Classroom observation	Yes	Yes	Full ECERS-R used[c]	Spring	Outcome of interest
Classroom Assessment Scoring System (CLASS)	Classroom observation	Yes	Yes	CLASS Instructional Support domain used[d]	Spring	Outcome of interest

[a]In this column, we identify whether we use the data descriptively, as a mediator or control variable, or as an outcome of interest in the mediation models.

[b]Given that parent-reported data on children's age, primary home language, and poverty status are available across cohorts, we constructed these variables by aggregating up from the child-level data.

[c]Despite use of the full ECERS-R across classrooms in FACES 2006, we derived short form scores for the cohort. We use the short form scores in the analyses.

[d]FACES 2006, used only the CLASS Instructional Support domain in the first Head Start spring wave and piloted the full CLASS in the second Head Start spring wave in a subset of classrooms.

Table A.2. Teacher background, credentials, and professional development constructs: FACES 2006, FACES 2009, and FACES 2014

Construct	Instrument	Used in FACES 2014?	Used in FACES 2009?	Used in FACES 2006?	Data reported	Analytic use of data[a]
Teacher training and mentoring						
How many training sessions or workshops attended (and length)	Teacher survey	Yes	Yes	No	Spring	Descriptive only
Whether have mentor	Teacher survey	Yes	Yes	Yes	Spring	Descriptive only
Who is mentor	Teacher survey	Yes	Yes	Yes	Spring	Mediator
Mentoring frequency	Teacher survey	Yes	Yes	Yes	Spring	Descriptive only
Curricular supports						
Whether teacher has been trained in curriculum in past year	Teacher survey	Yes	Yes	Yes	Spring	Descriptive only
Who provided curriculum training	Teacher survey	Yes	Yes	Yes	Spring	Descriptive only
Type of ongoing support for curriculum and who provides	Teacher survey	Yes	Yes	Yes	Spring	Descriptive only
Teacher education and experience						
Teacher education	Teacher survey	Yes	Yes	Yes	Fall (2006, 2009), spring (2015)	Mediator
Teacher credentials	Teacher survey	Yes	Yes	Yes	Fall (2006, 2009), spring (2015)	Descriptive only
Teacher years of experience	Teacher survey	Yes	Yes	Yes	Fall (2006, 2009), spring (2015)	Descriptive only
Teacher mental health and beliefs						
Teacher depression	Teacher survey	Yes	Yes	Yes	Spring	Descriptive only
Teacher job satisfaction	Teacher survey	Yes	Yes	Yes	Spring	Descriptive only

[a]In this column, we identify whether we use the data descriptively, as a mediator or control variable, or as an outcome of interest in the mediation models.

Table A.3. Program constructs: FACES 2006, FACES 2009, and FACES 2014

Construct	Instrument	Used in FACES 2014?	Used in FACES 2009?	Used in FACES 2006?	Data reported	Analytic use of data[a]
Teacher staffing and turnover in Head Start centers	Center director survey	Yes	Yes	Yes	Fall (2006, 2009), spring (2015)	Descriptive only

[a]In this column, we identify whether we use the data descriptively, as a mediator or control variable, or as an outcome of interest in the mediation models.

This page has been left blank for double-sided copying.

APPENDIX B

CROSS-COHORT ANALYSIS FINDINGS

This page has been left blank for double-sided copying.

Table B.1. Summary statistics for observed classroom quality: FACES 2006 - 2014

Scales	Mean (SD)			p
	FACES 2006 (Spring 2007)	FACES 2009 (Spring 2010)	FACES 2014 (Spring 2015)	
ECERS-R short form factors				
Teaching and Interactions	4.0 (0.9)	4.7 (1.0)	5.1 (1.0)	†, ‡
Provisions for Learning	3.6 (0.8)	4.0 (0.8)	4.4 (1.0)	†, ‡
CLASS domains				
Instructional Support	1.9 (0.6)	2.3 (0.6)	2.4 (0.9)	†, n.s.
Emotional Support	NA	5.3 (0.5)	5.4 (0.5)	n.a., n.s.
Classroom Organization	NA	4.7 (0.6)	4.8 (0.7)	n.a., n.s.
Child/adult ratio	6.3 (2.0)	6.2 (1.9)	5.9 (1.8)	†, n.s.
Group size	14.6 (2.8)	14.2 (2.7)	13.5 (3.0)	†, n.s.

Source: Spring 2007, 2010, and 2015 FACES Classroom Observation.

Note: Statistics are weighted to represent all Head Start classrooms in the cohort year.

In the table column labeled "p" we identify statistically significant change across cohorts at the $p \leq .05$ level. Statistically significant change from FACES 2006 to 2014 is represented by '†' and from FACES 2009 to 2014 by '‡'.

The FACES 2006, 2009, and 2014 columns in this table present actual means or percentages in each cohort; the significance of trends in the column labeled "p" is based on predicted values.

Researchers in other large scale studies have derived alternative dimensions of quality using a subset of items from the ECERS-R. Two factors reported in the Multi-State Study of Prekindergarten represent the key dimensions of quality tapped by the full ECERS-R: Provisions for Learning and Teaching and Interactions. These factors represent the key dimensions of quality tapped by the full ECERS-R. The short form score reported here is calculated by taking the mean of this subset of items.

NA = not available. Data not available in cohort.

n.a. = not applicable. Analysis not conducted between cohorts due to data availability, analytic priority, or limited variability.

n.s. = not significant. Change between cohorts is not statistically significant at the $p \leq .05$ level.

Table B.1a. Distribution of observed classroom quality by publisher categories: FACES 2006 - 2014

	Percentage of classrooms			
	FACES 2006 (Spring 2007)	FACES 2009 (Spring 2010)	FACES 2014 (Spring 2015)	p
ECERS-R short form factors[a]				
Teaching and Interactions				
Inadequate	11.7	4.6	0.7	†, ‡
Minimal	75.1	48.5	45.2	†, n.s.
Good or Excellent[b]	13.2	46.9	54.1	†, n.s.
Provisions for Learning				
Inadequate	19.5	8.4	8.8	†, n.s.
Minimal	74.7	79.0	59.1	†, ‡
Good or Excellent[b]	5.8	12.7	32.1	†, ‡
CLASS domains[c]				
Instructional Support				
Low	96.1	85.4	75.7	†, ‡
Mid	3.9	14.6	23.7	†, n.s.
High	0.0	0.0	0.6	†, ‡
Emotional Support				
Low	NA	0.2	0.1	n.a., n.s.
Mid	NA	95.3	86.8	n.a., ‡
High	NA	4.6	13.1	n.a., ‡
Classroom Organization				
Low	NA	0.4	1.4	n.a., n.s.
Mid	NA	98.7	96.1	n.a., ‡
High	NA	0.9	2.5	n.a., n.s.

Source: Spring 2007, 2010, and 2015 FACES Classroom Observation.

Note: Statistics are weighted to represent all Head Start classrooms in the cohort year.

In the table column labeled "p" we identify statistically significant change across cohorts at the $p \leq .05$ level. Statistically significant change from FACES 2006 to 2014 is represented by '†' and from FACES 2009 to 2014 by '‡'.

The FACES 2006, 2009, and 2014 columns in this table present actual means or percentages in each cohort; the significance of trends in the column labeled "p" is based on predicted values.

NA = not available. Data not available in cohort.

n.a. = not applicable. Analysis not conducted between cohorts due to data availability, analytic priority, or limited variability

n.s. = not significant. Change between cohorts is not statistically significant at the $p \leq .05$ level.

[a]Scores on the ECERS-R range from 1 to 7 with 1 or 2 = inadequate, 3 or 4 = minimal, 5 or 6 = good, and 7 = excellent quality.

[b]Given that no or very few classrooms scored in the "excellent" range, we collapsed the "good" and "excellent" response options.

[c]Scores for the CLASS range from 1 to 7 with 1 or 2 = low; 3, 4, or 5 = mid, and 6 or 7 = high.

Table B.2. Training and mentoring reported by teachers: FACES 2006 – 2014

Teacher training and mentoring	Percentage of classrooms/mean (SD)			
	FACES 2006 (Spring 2007)	FACES 2009 (Spring 2010)	FACES 2014 (Spring 2015)	p
Teacher has mentor	76.7	72.6	75.7	n.s., n.a.
If teacher has mentor, mentoring usually conducted by				
Another teacher	9.0	8.3	5.9	n.s., n.a.
Education coordinator, specialist	56.4	47.3	73.4	†, n.a.
Center/program director	24.6	26.8	9.8	†, n.a.
Other, including someone from outside the program	10.0	17.6	10.9	n.s., n.a.
If teacher has mentor, frequency mentor visits classroom[a]				
At least every two weeks	47.8	54.5	36.9	n.s., n.a.
Once a month or less	52.2	45.5	63.1	n.s., n.a.
How many trainings or workshops have you attended that were				
Less than one day	NA	5.0 (6.4)	3.2 (4.2)	n.a., ‡
One day	NA	3.4 (3.6)	3.0 (3.7)	n.a., n.s.
More than one day	NA	1.1 (2.0)	2.1 (7.1)	n.a., n.s.

Source: Spring 2007, 2010, and 2015 FACES Teacher Interview/Survey.

Note: Statistics are weighted to represent all Head Start classrooms in the cohort year.

In the table column labeled "p" we identify statistically significant change across cohorts at the $p \leq .05$ level. Statistically significant change from FACES 2006 to 2014 is represented by '†' and from FACES 2009 to 2014 by '‡'.

The FACES 2006, 2009, and 2014 columns in this table present actual means or percentages in each cohort; the significance of trends in the column labeled "p" is based on predicted values.

NA = not available. Data not available in cohort.

n.a. = not applicable. Analysis not conducted between cohorts due to data availability, analytic priority, or limited variability.

n.s. = not significant. Change between cohorts is not statistically significant at the $p \leq .05$ level.

[a]For easier interpretation of findings, we collapsed the response options for this variable (from once a week or more, once every two weeks, once a month, and less than once a month).

Table B.3. Teacher curriculum supports: FACES 2006 – 2014

Curriculum supports	Percentage of classrooms			
	FACES 2006 (Spring 2007)	FACES 2009 (Spring 2010)	FACES 2014 (Spring 2015)	p
Training on main curriculum in last 12 months				
Teacher received training	88.6	80.4	82.7	n.s., n.a.
Mean hours (SD) of training received	17.1 (22.7)	14.6 (16.7)	19.6 (30.3)	n.s., n.a.
Reported response range	1-205	1-120	1-192	
If teacher received curriculum training, who conducted training				
Staff from this Head Start program	58.4	60.5	70.1	n.s., n.a.
Other staff[a]	41.6	39.5	29.9	n.s., n.a.
Staff from another Head Start program	*2.6*	*3.5*	*2.0*	
Staff or consultant(s) from curriculum developers	*19.2*	*24.0*	*21.4*	
Faculty from school of education	*5.5*	*5.9*	*4.9*	
Head Start regional office training and technical assistance staff	*11.4*	*4.7*	*1.5*	
Other	*2.9*	*1.5*	*0.0*	
Ongoing supports received by teacher for curriculum use				
Help understanding the curriculum	65.1	75.2	68.9	n.s., n.a.
Help implementing the curriculum	58.8	61.1	44.6	†, n.a.
Feedback on implementing the curriculum	61.0	54.0	44.5	†, n.a.
Provide opportunities to observe someone implementing curriculum	*44.7*	*44.4*	*27.7*	
Refresher training on the curriculum	*64.6*	*63.7*	*55.6*	
Help planning curriculum-based activities	*63.2*	*69.7*	*54.1*	
Help individualizing the curriculum for children	*67.1*	*63.4*	*42.9*	
Help identifying and/or receiving additional resources to expand the scope of the curriculum and activities	*53.2*	*53.9*	*36.8*	
Other	*3.2*	*0.4*	*3.3*	

Table B.3. *(continued)*

Curriculum supports	Percentage of classrooms			
	FACES 2006 (Spring 2007)	FACES 2009 (Spring 2010)	FACES 2014 (Spring 2015)	p
Who provides ongoing supports for curriculum use				
Mentor or master teacher	15.3	20.2	31.5	†, n.a.
Other Head Start teachers in program	24.0	32.6	36.7	†, n.a.
Supervisor/education coordinator	73.3	73.8	74.4	n.s., n.a.
Staff from other Head Start Program	*8.5*	*11.1*	*8.1*	
Staff or consultant(s) from curriculum developers	*24.1*	*25.3*	*26.7*	
Faculty from school of education	*6.7*	*8.4*	*5.2*	
Head Start regional office training and technical assistance staff	*17.6*	*9.9*	*11.2*	
Other	*4.3*	*2.2*	*2.0*	

Source: Spring 2007, 2010, and 2015 FACES Teacher Interview/Survey.

Note: Statistics are weighted to represent all Head Start classrooms in the cohort year.

In the table column labeled "p" we identify statistically significant change across cohorts at the $p \leq .05$ level. Statistically significant change from FACES 2006 to 2014 is represented by '†' and from FACES 2009 to 2014 by '‡'.

The FACES 2006, 2009, and 2014 columns in this table present actual means or percentages in each cohort; the significance of trends in the column labeled "*p*" is based on predicted values.

Given their distribution and for easier interpretation of findings, the rows in italics were not tested for trends.

n.a. = not applicable. Analysis not conducted between cohorts due to data availability, analytic priority, or limited variability

n.s. = not significant. Change between cohorts is not statistically significant at the $p \leq .05$ level.

[a]For easier interpretation of findings, we collapsed the response options for this variable (from staff from another Head Start program, staff or consultant(s) from curriculum developers, faculty from School of Education, Head Start regional office training and technical assistance staff, and Other). We include the descriptive statistics for these categories for comprehensiveness.

Table B.4. Teacher education and credentials: FACES 2006 – 2014

Teacher education and credentials	Percentage of classrooms			
	FACES 2006 (Fall 2006)	**FACES 2009 (Fall 2009)**	**FACES 2014 (Spring 2015)**	*p*
Years teaching in Head Start[a]				
2 years or less	17.1	21.8	15.5	n.s., n.a.
3+ years	82.9	78.2	84.5	n.s., n.a.
Highest level of education				
Associate's degree (AA) or less	59.9	51.2	30.5	†, n.a.
Bachelor's degree (BA) or higher	40.1	48.8	69.5	†, n.a.
State-sponsored credentials				
Has a child development associate (CDA)	54.0	45.9	44.6	n.s., n.a.
Has a teaching certificate or license	42.2	41.1	51.4	n.s., n.a.
Has a state-awarded certificate	*31.2*	*29.3*	*29.9*	
Mean years teaching in Head Start	8.7	8.9	9.3	n.s., n.a.

Source: Fall 2006, Fall 2009, and Spring 2015 FACES Teacher Interview/Survey.

Note: Statistics are weighted to represent all Head Start classrooms in the cohort year.

In the table column labeled "p" we identify statistically significant change across cohorts at the $p \leq .05$ level. Statistically significant change from FACES 2006 to 2014 is represented by '†' and from FACES 2009 to 2014 by '‡'.

The FACES 2006, 2009, and 2014 columns in this table present actual means or percentages in each cohort; the significance of trends in the column labeled "*p*" is based on predicted values.

Given the distribution and for easier interpretation of findings, the row in italics was not tested for trends.

n.a. = not applicable. Analysis not conducted between cohorts due to data availability, analytic priority, or limited variability.

n.s. = not significant. Change between cohorts is not statistically significant at the $p \leq .05$ level.

[a]For easier interpretation of findings, we collapsed the response options for this variable (from two years or less, three to four years, five to nine years, and ten or more years).

Table B.5. Teacher mental health and job satisfaction: FACES 2006 – 2014

Teacher beliefs	Percentage of classrooms/mean scores			
	FACES 2006 (Spring 2007)	FACES 2009 (Spring 2010)	FACES 2014 (Spring 2015)	p
Degree of depressive symptoms[a, b]				
Not depressed	63.4	66.7	70.8	n.s., n.a.
Depressed	36.6	33.3	29.2	n.s., n.a.
Mean level of depressive symptoms (SD)	4.7 (5.5)	3.9 (4.6)	3.6 (4.1)	†, n.a.
Teacher satisfaction				
Enjoys present teaching job[c]	89.7	95.0	90.7	n.s., n.a.
Is making a difference in the lives of children s/he teaches[c]	97.8	98.6	96.4	n.s., n.a.
Would choose teaching again as a career[c]	86.8	88.5	84.8	n.s., n.a.

Source: Spring 2007, 2010, and 2015 FACES Teacher Interview/Survey.

Note: Statistics are weighted to represent all Head Start classrooms in the cohort year.

In the table column labeled "p" we identify statistically significant change across cohorts at the $p \leq .05$ level. Statistically significant change from FACES 2006 to 2014 is represented by '†' and from FACES 2009 to 2014 by '‡'.

The FACES 2006, 2009, and 2014 columns in this table present actual means or percentages in each cohort; the significance of trends in the column labeled "p" is based on predicted values.

n.a. = not applicable. Analysis not conducted between cohorts due to data availability, analytic priority, or limited variability.

n.s. = not significant. Change between cohorts is not statistically significant at the $p \leq .05$ level.

[a]Degree and level of depressive symptoms are based on the total score on the Center for Epidemiological Studies Depression Scale (CES-D) short form (12 items on a 4-point scale for frequency in the past week). Total scores range from 0 to 36. Scores ranging from 0 to 4 are coded as not depressed; from 5 to 9 as mildly depressed; from 10 to 14 as moderately depressed; and 15 and above as severely depressed. The CES-D is a screening tool and not a diagnostic tool, but scores have been correlated with clinical diagnosis.

[b]For easier interpretation of findings, we collapsed the response options for this variable (from not depressed, mildly depressed, moderately depressed, and severely depressed).

[c]Percentages reflect teachers who agree or strongly agree with this item.

Table B.6. Teacher staffing and turnover in Head Start centers: FACES 2006 – 2014

Staffing and recruitment	Classroom mean and range			
	FACES 2006 (Fall 2006)	**FACES 2009 (Fall 2009)**	**FACES 2014 (Spring 2015)**	*p*
Number of lead teachers employed in center classrooms				
Mean (SD)	5.0 (4.0)	5.6 (4.1)	5.5 (3.9)	n.s., n.a.
Range	1-15	1-15	1-15	
Lead teacher turnover in center classrooms[a]				
Mean percentage (SD)	21.2 (35.3)	14.4 (22.8)	13.7 (26.0)	n.s., n.a.
Range	0-200	0-150	0-200	

Source: Fall 2006, Fall 2009, and Spring 2015 FACES Center Director Interview/Survey.

Note: Statistics are weighted to represent all Head Start classrooms in the cohort year.

In the table column labeled "p" we identify statistically significant change across cohorts at the $p \leq .05$ level. Statistically significant change from FACES 2006 to 2014 is represented by '†' and from FACES 2009 to 2014 by '‡'.

The FACES 2006, 2009, and 2014 columns in this table present actual means or percentages in each cohort; the significance of trends in the column labeled "*p*" is based on predicted values.

n.a. = not applicable. Analysis not conducted between cohorts due to data availability, analytic priority, or limited variability.

n.s. = not significant. Change between cohorts is not statistically significant at the $p \leq .05$ level.

[a]Turnover is defined as the number of teachers that left and had to be replaced in the last 12 months divided by the total number of teachers currently employed at the center.

Table B.7. Classroom demographic characteristics: FACES 2006 – 2014

Characteristics	Percentage of classrooms/mean (SD)			
	FACES 2006 (Spring 2007)	FACES 2009 (Spring 2010)	FACES 2014 (Spring 2015)	p
Age of children as of September 1				
Mean (SD)	45.5 (4.5)	46.22 (4.7)	47.9 (4.7)	†, n.a.
Range	33-60	35-58	35-58	
DLL children served				
Mean percentage of DLLs served (SD)	25.3 (31.8)	25.2 (31.2)	23.8 (28.0)	n.s., n.a.
Range	0-100	0-100	0-100	
Children with income below poverty served				
Mean percentage of children living below poverty (SD)	58.3 (22.3)	62.1 (23.6)	68.2 (22.3)	†, n.a.
Range	0-100	0-100	0-100	

Source: Fall 2006, 2009, and 2014 FACES Parent Interview/Survey.

Note: Statistics are weighted to represent all Head Start classrooms in the cohort year.

In the table column labeled "p" we identify statistically significant change across cohorts at the $p \leq .05$ level. Statistically significant change from FACES 2006 to 2014 is represented by '†' and from FACES 2009 to 2014 by '‡'.

The FACES 2006, 2009, and 2014 columns in this table present actual means or percentages in each cohort; the significance of trends in the column labeled "p" is based on predicted values.

We define DLLs as those in households where a language other than English is primarily spoken to the child.

n.a. = not applicable. Analysis not conducted between cohorts due to data availability, analytic priority, or limited variability.

n.s. = not significant. Change between cohorts is not statistically significant at the $p \leq .05$ level.

Table B.8. Potential control and mediating variables included in the analyses

Variable	Trend analysis findings	In mediation analyses?
Teacher mentoring		
Teacher has mentor	n.s., n.a.	Yes, mediator variable (no mentor as referent)
Teacher receives mentoring from education coordinator	†, n.a.	Yes, mediator variable
Teacher receives mentoring from center/program director	†, n.a.	Yes, mediator variable
Teacher receives mentoring from teacher or someone outside the program	n.s., n.a.	Yes, mediator variable
Teacher receives mentoring at least every two weeks	n.s., n.a.	No
Teacher receive mentoring once a month or less	n.s., n.a.	No
Teacher participation in training and workshops		
Number of trainings teacher attended lasting less than one day	n.a., ‡	No
Number of trainings teacher attended lasting one day	n.a., n.s.	No
Number of trainings teacher attended lasting more than one day	n.a., n.s.	No
Teacher curricular supports		
Teacher receives curriculum training	n.s., n.a.	No
Mean hours of curriculum training received	n.s., n.a.	No
Staff from program provide curriculum training	n.s., n.a.	No
Others provide curriculum training	n.s., n.a.	No
Teacher receives help understanding the curriculum	n.s., n.a.	No
Teacher receives help implementing the curriculum	†, n.a.	No
Teacher receives feedback on implementing the curriculum	†, n.a.	No
Mentor or master teacher provides ongoing support for curriculum	†, n.a.	No
Other teachers in program provide ongoing support for curriculum	†, n.a.	No
Supervisor/education coordinator provides ongoing support for curriculum	n.s., n.a.	No
Teacher education and credentials		
Teaching in Head Start 2 years or less	n.s., n.a.	No
Teaching in Head Start 3 or more years	n.s., n.a.	No
Mean years teaching in Head Start	n.s., n.a.	No
Teacher has associate's degree or less	†, n.a.	No
Teacher has bachelor's degree or higher	†, n.a.	Yes, mediator variable
Teacher has CDA	n.s., n.a.	No
Teacher has a teaching certificate or license	n.s., n.a.	No
Teacher mental health and job satisfaction		
Teacher is not depressed	n.s., n.a.	No
Teacher is depressed	n.s., n.a.	No
Teacher mean level of depressive symptoms	†, n.a.	No

Table B.8. *(continued)*

Variable	Trend analysis findings	In mediation analyses?
Teacher enjoys present teaching job	n.s., n.a.	No
Teacher is making a difference in the lives of children s/he teaches	n.s., n.a.	No
Teacher would choose teaching again as a career	n.s., n.a.	No
Staffing and recruitment		
Mean number of lead teachers employed in center classrooms	n.s., n.a.	No
Lead teacher turnover in center classrooms	n.s., n.a.	No
Classroom demographic characteristics		
Mean age of children	†, n.a	Yes, control variable
Mean percentage of DLLs served	n.s., n.a.	Yes, control variable
Mean percentage of children living below poverty	†, n.a	Yes, control variable

Note: Statistically significant change from FACES 2006 to 2014 is represented by '†' and from FACES 2009 to 2014 by '‡'.

n.a. = not applicable. Analysis not conducted between cohorts due to data availability, analytic priority, or limited variability.

n.s. = not significant. Change between cohorts is not statistically significant at the $p \leq .05$ level.

Table B.9. Associations between time and observed classroom quality, independently and controlling for selected characteristics: FACES 2006 – 2014

| | FACES 2006-2014 | | | FACES 2009-2014 | | |
| | Coefficient (standard error) | | | Coefficient (standard error) | | |
Scales	Trend	Controls	Controls + Mediators	Trend	Controls	Controls + Mediators
ECERS-R short form factors						
ECERS-R Teaching and Interactions	1.08 (0.12)*	1.10 (0.12)*	1.11 (0.13)*	0.39 (0.13)*	0.40 (0.13)*	0.41 (0.14)*
ECERS-R Provisions for Learning	0.86 (0.13)*	0.88 (0.14)*	0.91 (0.14)*	0.47 (0.14)*	0.48 (0.15)*	0.50 (0.16)*
CLASS domains						
Instructional Support	0.48 (0.11)*	0.46 (0.11)*	0.41 (0.11)*	0.13 (0.11)	0.12 (0.11)	0.07 (0.12)
Emotional Support	NA	NA	NA	0.11 (0.08)	0.13 (0.08)	0.13 (0.09)
Classroom Organization	NA	NA	NA	0.12 (0.11)	0.11 (0.11)	0.12 (0.12)

Source: Spring 2007, 2010, and 2015 FACES Classroom Observation.

Note: Statistics are weighted to represent all Head Start classrooms in the cohort year.

*Asterisk indicates that the association between time (cohort) and observed quality is statistically significant at the $p \leq .05$ level.

Researchers in other large scale studies have derived alternative dimensions of quality using a subset of items from the ECERS-R. Two factors reported in the Multi-State Study of Prekindergarten represent the key dimensions of quality tapped by the full ECERS-R: Provisions for Learning and Teaching and Interactions. These factors represent the key dimensions of quality tapped by the full ECERS-R. The short form score reported here is calculated by taking the mean of this subset of items.

NA = not available. Data not available in cohort.

Table B.9a. Associations between time and observed classroom quality, independently and controlling for selected characteristics: FACES 2006 – 2014

	FACES 2006-2014			FACES 2009-2014		
	Estimate (standard error)			Estimate (standard error)		
	Trend	Controls	Controls + Mediators	Trend	Controls	Controls + Mediators
ECERS-R short form factors[a]						
Teaching and Interactions						
Inadequate	2.93 (0.69)*	3.67 (0.84)*	3.49 (0.85)*	1.91 (0.73)*	2.67 (0.86)*	2.37 (0.90)*
Minimal	1.39 (0.26)*	1.40 (0.26)*	1.49 (0.28)*	0.13 (0.24)	0.16 (0.25)	0.28 (0.27)
Good or Excellent	-2.05 (0.31)*	-2.15 (0.31)*	-2.23 (0.33)*	-0.29 (0.25)	-0.32 (0.26)	-0.41 (0.28)
Provisions for Learning						
Inadequate	0.92 (0.45)*	0.99 (0.46)*	0.95 (0.46)*	-0.05 (0.45)	-0.01 (0.49)	0.05 (0.47)
Minimal	0.71 (0.33)*	0.66 (0.34)*	0.74 (0.34)*	0.95 (0.31)*	0.95 (0.32)*	0.97 (0.33)*
Good or Excellent	-2.04 (0.43)*	-2.02 (0.44)	-2.15 (0.46)*	-1.18 (0.36)*	-1.20 (0.36)*	-1.28 (0.40)*
CLASS domains[b]						
Instructional Support						
Low	2.06 (0.39)*	2.02 (0.41)*	1.96 (0.44)*	0.63 (0.32)*	0.63 (0.33)*	0.58 (0.36)
Mid	-2.03 (0.39)*	-1.99 (0.41)*	-1.94 (0.44)*	-0.60 (0.32)	-0.60 (0.33)	-0.56 (0.36)
High	-19.10 (1.12)*	-18.70 (0.72)*	-31.42 (0.37)*	-19.11 (0.91)*	-17.80 (0.72)*	-77.40 (0.38)*
Emotional Support						
Low	NA	NA	NA	0.88 (1.44)	0.53 (1.21)	0.98 (1.25)
Mid	NA	NA	NA	1.15 (0.40)*	1.13 (0.41)*	1.44 (0.47)*
High	NA	NA	NA	-1.18 (0.41)*	-1.15 (0.41)*	-1.52 (0.48)*
Classroom Organization						
Low	NA	NA	NA	n.a.	n.a.	n.a.
Mid	NA	NA	NA	n.a.	n.a.	n.a.
High	NA	NA	NA	n.a.	n.a.	n.a.

Source: Spring 2007, 2010, and 2015 FACES Classroom Observation.
Notes: Statistics are weighted to represent all Head Start classrooms in the cohort year.
          This table includes regression coefficients and associated standard errors.
*Asterisk indicates that the association between time (cohort) and observed quality is statistically significant at the $p \leq .05$ level.
NA = not available. Data not available in cohort.
n.a. = not applicable. Analysis not conducted between cohorts due to data availability, analytic priority, or limited variability.
[a]Scores on the ECERS-R range from 1 to 7 with 1 or 2 = inadequate, 3 or 4 = minimal, 5 or 6 = good, and 7 = excellent quality.
[b]Scores for the CLASS range from 1 to 7 with 1 or 2 = low; 3, 4, or 5 = mid, and 6 or 7 = high.

Table B.10. Associations between selected characteristics and average observed classroom quality: FACES 2006 – 2014

Characteristics	ECERS-R short form factors Estimate (standard error)			CLASS domains Estimate (standard error)	
	Teaching and Interactions	Provisions for Learning	Instructional Support	Emotional Support	Classroom Organization
Controls					
Age of children in classroom	0.010 (0.009)	0.003 (0.010)	0.010 (0.005)*	NA	NA
Proportion of DLLs in classroom	0.001 (0.002)	-0.001 (0.001)	-0.001 (0.001)	NA	NA
Proportion of children below poverty in classroom	-0.004 (0.002)	-0.002 (0.002)	-0.000 (0.001)	NA	NA
Mediators					
Teacher receives no mentoring (referent)					
Teacher receives mentoring from education coordinator	-0.056 (0.087)	0.040 (0.095)	-0.055 (0.070)	NA	NA
Teacher receives mentoring from center/program director	-0.092 (0.109)	0.183 (0.104)	-0.097 (0.074)	NA	NA
Teacher receives mentoring from another teacher or someone else	-0.118 (0.121)	-0.120 (0.102)	-0.170 (0.085)*	NA	NA
Teacher has BA+	-0.029 (0.080)	-0.082 (0.084)	0.043 (0.052)	NA	NA
Model R-square	0.16	0.12	0.09	NA	NA

Source: Spring 2007, 2010, and 2015 FACES Classroom Observation.

Notes: Statistics are weighted to represent all Head Start classrooms in the cohort year.

This table includes regression coefficients and associated standard errors. We present the model R-square for the analyses that included both control and mediating variables.

*Asterisk indicates that the association between the characteristic and observed quality is statistically significant at the $p \leq .05$ level.

Researchers in other large scale studies have derived alternative dimensions of quality using a subset of items from the ECERS-R. Two factors reported in the Multi-State Study of Prekindergarten represent the key dimensions of quality tapped by the full ECERS-R: Provisions for Learning and Teaching and Interactions. These factors represent the key dimensions of quality tapped by the full ECERS-R. The short form score used here is calculated by taking the mean of this subset of items.

NA = not available. Data not available in cohort and analysis not conducted between cohorts.

Table B.10a. Associations between selected characteristics and observed ECERS-R short form categories: FACES 2006 – 2014

Characteristics	Teaching and Interactions Estimate (standard error)			Provisions for Learning Estimate (standard error)		
	Minimal	Inadequate	Good or Excellent	Minimal	Inadequate	Good or Excellent
Controls						
Age of children in classroom	0.007 (0.037)	0.013 (0.020)	-0.018 (0.022)	-0.000 (0.034)	0.009 (0.022)	-0.017 (0.026)
Proportion of DLLs in classroom	-0.001 (0.005)	0.004 (0.003)	-0.004 (0.004)	-0.004 (0.005)	-0.002 (0.004)	0.010 (0.005)*
Proportion of children below poverty in classroom	0.009 (0.008)	-0.013 (0.004)*	0.012 (0.004)*	-0.010 (0.006)	0.541 (0.004)	0.000 (0.005)
Mediators						
Teacher receives no mentoring (referent)						
Teacher receives mentoring from education coordinator	-0.352 (0.372)	-0.046 (0.206)	0.157 (0.216)	-0.218 (0.291)	-0.019 (0.275)	0.253 (0.416)
Teacher receives mentoring from center/program director	-0.236 (0.504)	-0.084 (0.260)	0.159 (0.266)	0.975 (0.425)	0.238 (0.309)	-0.482 (0.380)
Teacher receives mentoring from another teacher or someone else	-0.600 (0.471)	0.095 (0.262)	0.097 (0.281)	-0.811 (0.371)*	0.403 (0.333)	0.204 (0.395)
Teacher has BA+	0.409 (0.328)	-0.377 (0.198)	0.304 (0.220)	-0.009 (0.247)	-0.021 (0.202)	0.048 (0.289)

Source: Spring 2007, 2010, and 2015 FACES Classroom Observation.
Notes: Statistics are weighted to represent all Head Start classrooms in the cohort year.
            This table includes regression coefficients and associated standard errors.
*Asterisk indicates that the association between the characteristic and observed quality is statistically significant at the $p \leq .05$ level.

Researchers in other large scale studies have derived alternative dimensions of quality using a subset of items from the ECERS-R. Two factors reported in the Multi-State Study of Prekindergarten represent the key dimensions of quality tapped by the full ECERS-R: Provisions for Learning and Teaching and Interactions. These factors represent the key dimensions of quality tapped by the full ECERS-R. The short form score used here is calculated by taking the mean of this subset of items.

NA = not available. Data not available in cohort and analysis not conducted between cohorts.

Table B.10b. Associations between selected characteristics and observed CLASS domain categories: FACES 2006 – 2014

Characteristics	Instructional Support Estimate (standard error)			Emotional Support Estimate (standard error)			Classroom Organization Estimate (standard error)		
	Low	Mid	High	Low	Mid	High	Low	Mid	High
Controls									
Age of children in classroom	0.015 (0.024)	-0.016 (0.024)	14.232 (0.085)*	NA	NA	NA	NA	NA	NA
Proportion of DLLs in classroom	-0.007 (0.005)	0.007 (0.005)	3.207 (0.004)*	NA	NA	NA	NA	NA	NA
Proportion of children below poverty in classroom	-0.000 (0.006)	0.001 (0.006)	-9.305 (0.008)*	NA	NA	NA	NA	NA	NA
Mediators Teacher receives no mentoring (referent)									
Teacher receives mentoring from education coordinator	-0.174 (0.292)	0.182 (0.292)	-75.720 (0.198)*	NA	NA	NA	NA	NA	NA
Teacher receives mentoring from center/program director	-0.208 (0.405)	0.209 (0.404)	-92.150 (0.183)*	NA	NA	NA	NA	NA	NA
Teacher receives mentoring from another teacher or someone else	-1.020 (0.426)*	1.020 (0.427)*	-19.840 (0.191)*	NA	NA	NA	NA	NA	NA
Teacher has BA+	0.677 (0.306)	-0.062 (0.306)	34.870 (0.152)*	NA	NA	NA	NA	NA	NA

Source: Spring 2007, 2010, and 2015 FACES Classroom Observation.
Notes: Statistics are weighted to represent all Head Start classrooms in the cohort year.
This table includes regression coefficients and associated standard errors.
*Asterisk indicates that the association between the characteristic and observed quality is statistically significant at the $p \leq .05$ level.
NA = not available. Data not available in cohort and analysis not conducted between cohorts.

Table B.11. Associations between selected characteristics and average observed classroom quality: FACES 2009 – 2014

Characteristics	ECERS-R short form factors Estimate (standard error)		CLASS domains Estimate (standard error)		
	Teaching and Interactions	Provisions for Learning	Instructional Support	Emotional Support	Classroom Organization
Controls					
Mean age of children in classroom	0.011 (0.011)	0.002 (0.012)	0.010 (0.006)	0.000 (0.006)	0.010 (0.007)
Proportion of DLLs in classroom	0.003 (0.002)	-0.001 (0.002)	-0.000 (0.001)	0.001 (0.001)	0.002 (0.001)
Proportion of children below poverty in classroom	-0.004 (0.003)	-0.001 (0.002)	-0.001 (0.002)	-0.002 (0.001)	0.001 (0.001)
Mediators					
Teacher receives no mentoring (referent)					
Teacher receives mentoring from education coordinator	-0.119 (0.114)	0.056 (0.127)	-0.065 (0.091)	-0.098 (0.067)	-0.228 (0.093)*
Teacher receives mentoring from center/program director	-0.245 (0.137)	0.168 (0.137)	-0.107 (0.100)	-0.111 (0.062)	-0.214 (0.098)*
Teacher receives mentoring from another teacher or someone else	-0.105 (0.152)	0.052 (0.118)	-0.249 (0.107)*	-0.107 (0.081)	-0.313 (0.076)*
Teacher has BA+	-0.003 (0.112)	-0.122 (0.114)	0.066 (0.076)	-0.034 (0.051)	-0.061 (0.061)
Model R-square	0.04	0.05	0.02	0.03	0.02

Source: Spring 2007, 2010, and 2015 FACES Classroom Observation.
Notes: Statistics are weighted to represent all Head Start classrooms in the cohort year.
             This table includes regression coefficients and associated standard errors. We present the model R-square for the analyses that included both control and mediating variables.
*Asterisk indicates that the association between the characteristic and observed quality is statistically significant at the $p \leq .05$ level.

Researchers in other large scale studies have derived alternative dimensions of quality using a subset of items from the ECERS-R. Two factors reported in the Multi-State Study of Prekindergarten represent the key dimensions of quality tapped by the full ECERS-R: Provisions for Learning and Teaching and Interactions. These factors represent the key dimensions of quality tapped by the full ECERS-R. The short form score used here is calculated by taking the mean of this subset of items.

Table B.11a. Associations between selected characteristics and observed ECERS-R short form categories: FACES 2009 – 2014

Characteristics	Teaching and Interactions Estimate (standard error)			Provisions for Learning Estimate (standard error)		
	Minimal	Inadequate	Good or Excellent	Minimal	Inadequate	Good or Excellent
Controls						
Age of children in classroom	0.061 (0.062)	0.005 (0.023)	-0.013 (0.024)	0.004 (0.047)	-0.002 (0.023)	0.002 (0.024)
Proportion of DLLs in classroom	0.025 (0.013)	0.002 (0.004)	-0.004 (0.004)	0.001 (0.008)	-0.004 (0.004)	0.005 (0.005)
Proportion of children below poverty in classroom	0.004 (0.009)	-0.013 (0.004)*	0.012 (0.005)*	-0.010 (0.009)	0.004 (0.005)	-0.001 (0.006)
Mediators						
Teacher receives no mentoring (referent)						
Teacher receives mentoring from education coordinator	-0.724 (0.732)	-0.188 (0.238)	0.258 (0.242)	-0.262 (0.384)	-0.014 (0.363)	0.149 (0.485)
Teacher receives mentoring from center/program director	-0.947 (0.668)	-0.235 (0.261)	0.386 (0.276)	0.057 (0.592)	0.321 (0.355)	-0.454 (0.367)
Teacher receives mentoring from another teacher or someone else	-0.611 (0.983)	-0.100 (0.316)	0.175 (0.321)	-0.075 (0.458)	-0.019 (0.357)	0.006 (0.428)
Teacher has BA+	1.060 (0.528)*	-0.499 (0.273)	0.355 (0.277)	-0.322 (0.385)	0.112 (0.254)	0.043 (0.332)

Source: Spring 2007, 2010, and 2015 FACES Classroom Observation.

Notes: Statistics are weighted to represent all Head Start classrooms in the cohort year.

This table includes regression coefficients and associated standard errors.

*Asterisk indicates that the association between the characteristic and observed quality is statistically significant at the $p \leq .05$ level.

Researchers in other large scale studies have derived alternative dimensions of quality using a subset of items from the ECERS-R. Two factors reported in the Multi-State Study of Prekindergarten represent the key dimensions of quality tapped by the full ECERS-R: Provisions for Learning and Teaching and Interactions. These factors represent the key dimensions of quality tapped by the full ECERS-R. The short form score used here is calculated by taking the mean of this subset of items.

Table B.11b. Associations between selected characteristics and observed CLASS domain categories: FACES 2009 – 2014

Characteristics	Instructional Support Estimate (standard error)			Emotional Support Estimate (standard error)			Classroom Organization Estimate (standard error)		
	Low	Mid	High	Low	Mid	High	Low	Mid	High
Controls									
Age of children in classroom	-0.002 (0.025)	0.001 (0.025)	14.148 (0.116)*	0.026 (0.100)	0.022 (0.039)	-0.021 (0.039)	n.a.	n.a.	n.a.
Proportion of DLLs in classroom	-0.009 (0.005)	0.009 (0.005)	3.210 (0.005)*	-0.044 (0.014)	0.001 (0.007)	0.001 (0.007)	n.a.	n.a.	n.a.
Proportion of children below poverty in classroom	-0.003 (0.007)	0.003 (0.007)	-9.240 (0.013)*	-0.003 (0.010)	0.003 (0.011)	-0.003 (0.012)	n.a.	n.a.	n.a.
Mediators									
Teacher receives no mentoring (referent)									
Teacher receives mentoring from education coordinator	-0.237 (0.317)	0.248 (0.317)	-75.380 (0.283)*	-17.370 (0.000)	-0.451 (0.448)	0.531 (0.457)	n.a.	n.a.	n.a.
Teacher receives mentoring from center/program director	-0.182 (0.433)	0.185 (0.433)	-131.010 (0.248)*	-0.834 (0.000)	-2.270 (0.864)*	2.290 (0.869)*	n.a.	n.a.	n.a.
Teacher receives mentoring from another teacher or someone else	-0.913 (0.457)*	0.917 (0.457)*	-87.980 (0.222)*	-0.138 (0.000)	-0.532 (0.524)	0.548 (0.529)	n.a.	n.a.	n.a.
Teacher has BA+	0.027 (0.341)	-0.020 (0.3342)	36.560 (0.306)*	-2.670 (1.750)*	-1.510 (0.501)*	1.630 (0.505)*	n.a.	n.a.	n.a.

Source: Spring 2007, 2010, and 2015 FACES Classroom Observation.

Notes: Statistics are weighted to represent all Head Start classrooms in the cohort year.

This table includes regression coefficients and associated standard errors.

*Asterisk indicates that the association between the characteristic and observed quality is statistically significant at the $p \leq .05$ level.

n.a. = not applicable. Analysis not conducted between cohorts due to data availability, analytic priority, or limited variability.

This page has been left blank for double-sided copying.